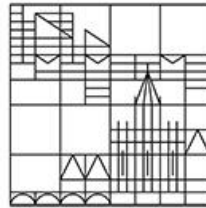


Evaluation eines Trainings in Narrativer Expositionstherapie in Kolumbien

Wissenschaftliche Arbeit
zur Erlangung des akademischen Grades
„Master of Science“ (M. Sc.)
im Fach Psychologie

Fachbereich Psychologie
an der

Universität
Konstanz



Vorgelegt von:
Christian Schart
Steinstraße 17a
78467 Konstanz

Erstgutachter: Prof. Dr. Thomas Elbert
Zweitgutachter: Dr. Michael Odenwald

Konstanz, im September 2014

Danksagungen

Zu Beginn möchte ich mich bei Herrn Professor Dr. Elbert und Herrn Dr. Odenwald für die Anregung und Annahme dieser Masterarbeit bedanken. Nur durch die Möglichkeit,, eigenständig an die Fragestellung heranzugehen, aber stets Rückhalt bei aufkommenden Unsicherheiten zu haben, konnte diese Arbeit so entstehen, wie sie hier vorliegt.

Ein ganz besonderer Dank geht auch an Claudia Bueno M. Sc., die während des gesamten Entstehungsprozesses mit Rat und Tat zur Seite stand. Ihre umfassenden Kenntnisse des kolumbianischen Kontextes mit all seinen Facetten bildeten sowohl den Grundstein für diese Untersuchung, als auch die Möglichkeit, die Ergebnisse angemessen zu interpretieren.

También quiero decir muchas gracias a los participantes del entrenamiento y terapeutas que en su trabajo diario ayudan a que Colombia se convierta en un país mas pacífico. Sin ustedes ni el entrenamiento, ni este trabajo hubiera sido posible.

Ein großer Dank geht auch an Lea, Lisa und Lukas für die Korrekturen und Anregungen in den letzten Wochen. Durch ihre Kritik habe ich zwar nochmals einiges zu tun gehabt, doch nur so konnte diese Masterarbeit den letzten Schliff erhalten.

Zusammenfassung

Nach über 50 Jahren gewaltvoller Auseinandersetzungen in Kolumbien ist die Reintegration der ehemaligen Mitglieder der illegalen Gruppierungen von Guerillas und Paramilitärs in die Zivilgesellschaft ein zentraler Moment für eine friedlichere Zukunft. Für eine gelungene Reintegration muss dabei ein besonderes Augenmerk auf die Behandlung psychischer Krankheiten und vor allem der Posttraumatischen Belastungsstörung (PTBS) gelegt werden (Maedl, Schauer, Odenwald & Elbert, 2010). Wie Bueno Castellanos (2014) an einer Stichprobe ehemaliger Kämpferinnen und Kämpfer in Kolumbien zeigen konnte, ist die Narrative Expositionstherapie (NET) eine effektive Methode zur Behandlung der PTBS-Symptomatik.

In dieser Studie wurde nun untersucht, ob durch ein zehntägiges Training in Kolumbien den teilnehmenden Psychologen und Psychologinnen die nötigen Kenntnisse und Fertigkeiten der NET vermittelt werden können, damit diese die Therapie effektiv durchführen können. Dazu wurde die Hypothese getestet, dass sich durch das Training sowohl das Wissen zu PTBS, die Selbstwirksamkeit als auch die Einstellung gegenüber evidenzbasierten Therapien ansteigen bzw. positiv verändern. Zudem wurde postuliert, dass die Teilnehmenden nach dem Training die grundlegenden Konzepte der NET umsetzen können und in einem letzten Schritt wurden die Zusammenhänge der verschiedenen Komponenten überprüft.

Insgesamt nahmen 29 Psychologinnen und Psychologen an dem Training teil, wobei 25 Personen anschließend als Therapeuten und Therapeutinnen und 4 Personen in der Leitung des Reintegrationsprogramms arbeiteten. Die Teilnehmenden hatten im Schnitt vor dem Training gut 8 Jahre als Psychologen und Psychologinnen gearbeitet und über 65% gaben an, bereits mindestens eine Fort- bzw. Ausbildung in einer anderen Therapieform erhalten zu haben. Knapp 90% der Teilnehmenden gab an, bereits vor dem Training Erfahrungen mit der Zielpopulation gehabt zu haben (im Schnitt 3.7 Jahre).

Die Datenerhebung erfolgte sowohl direkt vor und nach dem Training, als auch zu einem Follow-Up nach sieben Monaten mittels eines eigens zusammengestellten Fragebogens zu den Bereichen Reaktionen auf das Training, Selbstwirksamkeit, Einstellungen und Wissen. Die Umsetzung der NET wurde mithilfe eines Verhaltensbeobachtungsbogens während der Supervisionssitzung sieben Monate nach dem Training überprüft.

In allen Bereichen des Fragebogens zeigten sich durchweg hohe Werte. Jedoch erreichten lediglich die Veränderungen für das Wissen ein signifikantes Niveau, so dass dieses durch das Training hypothesenkonform gesteigert wurde. Ebenso zeigte sich wie postuliert, dass die Therapeuten und Therapeutinnen die grundlegenden Konzepte der NET angemessen umsetzen können. Im Hinblick auf die Zusammenhänge der verschiedenen Komponenten ergaben sich nur vereinzelt signifikante Ergebnisse.

Zusammenfassend kann gesagt werden, dass das durchgeführte Training geeignet ist, die Teilnehmenden auf ihre Arbeit als Traumatherapeutinnen und Traumatherapeuten vorzubereiten und somit einen wichtigen Teil zur Reintegration ehemaliger Kämpfer und Kämpferinnen beizutragen. Für die weitere Forschung sollten die Erkenntnisse sowohl an einer größeren Stichprobe verifiziert werden als auch spezifischer untersucht werden, welche Trainingskomponenten besonders wichtig für den Trainingserfolg sind.

Resumen

Después de más que 50 años de conflicto armado en Colombia la reintegración de los miembros de los grupos armados ilegales de las guerrillas y paramilitares a la vida civil es clave para un futuro pacífico. Para una reintegración efectiva el tratamiento de los trastornos mentales, y en especial del trastorno por estrés postraumático (TEPT) es muy importante (Maedl, Schauer, Odenwald & Elbert, 2010). Como Bueno Castellanos (2014) mostró en una muestra de excombatientes en Colombia, la Terapia Narrativa de Exposición (NET; por sus siglas en alemán) es un método efectivo en el tratamiento del TEPT.

El presente trabajo busca identificar, si un entrenamiento de diez días en Colombia puede facilitar los conocimientos y habilidades necesarias a los psicólogos y psicólogas participantes para ejecutar la terapia NET efectivamente. Para esto se puso a prueba la hipótesis que mediante el entrenamiento aumentaban o mejoraban los conocimientos del TEPT, como también la autoeficacia y la actitud frente a terapias basadas en la evidencia. Además se postuló que debido al entrenamiento los participantes pueden aplicar los conceptos básicos de la NET y se examinó las relaciones entre los diferentes componentes.

En total 29 psicólogas y psicólogos participaron del entrenamiento. De los cuales 25 trabajaban como terapeutas y 4 pertenecían a nivel central de la Agencia Colombiana para la Reintegración (ACR). En promedio los participantes habían trabajado 8 años como psicólogos o psicólogas y más del 65% afirmaron que tenían mínimo algún tipo de formación en terapia. Antes del entrenamiento casi el 90% había trabajado con excombatientes ($M = 3.7$ años).

La recopilación de datos fue realizada tanto directamente antes y después del entrenamiento como también en un seguimiento después de 7 meses a través de un cuestionario que evaluaba diferentes reacciones al entrenamiento como autoeficacia, actitudes y conocimientos. La aplicación de la NET fue revisada mediante protocolos de observación durante la supervisión 7 meses después del entrenamiento.

En todos los áreas del cuestionario se obtuvieron valores altos para los diferentes momentos evaluados. Sin embargo diferencias significativas a lo largo del tiempo se presentaron solo en el área del conocimiento, el cual aumento, según lo postulado, por medio del entrenamiento. Así mismo los conceptos básicos de la NET fueron aplicados adecuadamente por los terapeutas. Con relación a los diferentes componentes examinados se obtuvieron resultados significativos solamente en algunos.

En resumen se puede decir, que el entrenamiento ejecutado es apropiado para preparar los participantes para trabajar como terapeutas de estrés postraumático. Y por lo tanto aportar positivamente a la reintegración de los excombatientes en Colombia. Para ensayos siguientes se debería profundizar estos resultados con una muestra mas grande y además analizar qué componentes del entrenamiento son especialmente importante para el éxito del mismo.

Inhaltsverzeichnis

Tabellenverzeichnis	VI
Abbildungsverzeichnis	VI
Abkürzungsverzeichnis	VII
1 Einleitung	1
1.1 <i>Trainingsevaluation</i>	2
1.1.1 Integrated Model of Training Evaluation and Effectiveness	6
1.1.2 Therapist Training Evaluation Outcomes Framework	9
1.1.2.1 Reaktionen	10
1.1.2.2 Einstellungen	11
1.1.2.3 Wissen	11
1.1.2.4 Fertigkeiten	11
1.1.2.5 Verhaltensveränderungen in der Praxis	12
1.1.2.6 Therapieerfolg	12
1.2 <i>Bedarfsanalyse</i>	13
1.2.1 Bewaffneter Kampf in Kolumbien	13
1.2.2 Demobilisierungsprozess	14
1.2.3 Narrative Expositionstherapie	17
1.3 <i>Hypothesen</i>	18
2 Methoden	20
2.1 <i>Messinstrumente</i>	20
2.1.1 Fragebogen Reaktion-Einstellung-Wissen	20
2.1.2 Beurteilungsbogen Diagnostik-Lifeline-Exposition-NET	23
2.2 <i>Durchführung</i>	24
2.2.1 Training	25
2.2.2 Datenerhebung	25
2.3 <i>Stichprobenbeschreibung</i>	26
2.4 <i>Statistische Auswertung</i>	30

3	Ergebnisse	32
3.1	<i>Deskriptive Statistik und Veränderungen</i>	32
3.1.1	Reaktionen	32
3.1.2	Einstellungen	33
3.1.3	Wissen	36
3.1.4	Verhalten in der Praxis	39
3.2	<i>Zusammenhangsmaße</i>	41
4	Diskussion	45
4.1	<i>Einstellungen</i>	45
4.2	<i>Wissen</i>	47
4.3	<i>Verhalten in der Praxis</i>	50
4.4	<i>Zusammenhänge der Konstrukte</i>	51
4.5	<i>Reaktionen</i>	53
4.6	<i>Limitationen</i>	54
4.7	<i>Schlussfolgerungen und Ausblick</i>	58
5	Literaturverzeichnis	61
6	Anhang	61
A1	<i>Fragebogen pre</i>	61
A2	<i>Fragebogen post</i>	67
A3	<i>Beurteilungsbogen DLE-NET</i>	72

Tabellenverzeichnis

Tabelle 1	Therapist training evaluation outcomes framework.....	10
Tabelle 2	Stichprobenbeschreibung nach Beschäftigungsstatus.....	27
Tabelle 3	Absolute und relative Häufigkeiten der angegebenen Fort- und Weiterbildungen.....	28
Tabelle 4	Deskriptive Statistik zu den drei Erhebungszeitpunkten.....	32
Tabelle 5	Cronbachs α der EBPAS-Skalen zu den drei Erhebungszeitpunkten	34
Tabelle 6	Mittelwerte und Standardabweichungen der DLE-NET Items.....	40
Tabelle 7	Korrelationstabelle der einzelnen Evaluationskomponenten.....	43

Abbildungsverzeichnis

Abbildung 1	Integrated Model of Training Evaluation and Effectiveness nach Alvarez et al. (2004).....	7
Abbildung 2	Übersicht Niederlassungen der ACR sowie Anzahl der Personen im Prozess der Reintegration (PPR)	16
Abbildung 3	Abgewandeltes Modell der Trainingsevaluation und -effektivität basierend auf Alvarez et al. (2004).....	19
Abbildung 4	Flussdiagramm der Teilnehmenden über die verschiedenen Zeitpunkte der Evaluation nach Beschäftigungsstatus.	29
Abbildung 5	Vergleich EBPAS Items und Skalen zwischen Therapeut_innen und Koordinator_innen direkt nach dem Training	35
Abbildung 6	Boxplot EBPAS-Skalen der Therapeut_innen zu den drei Erhebungszeitpunkten mit p-Werten der signifikanten post-hoc berechneten Wilcoxon- Vorzeichenrangtests	36
Abbildung 7	Anteil richtiger Antworten nach Antwortmöglichkeiten zu den drei Messzeitpunkten.	38
Abbildung 8	Boxplot der Wissenscores zu den drei Messzeitpunkten mit p-Werten der post-hoc berechneten Wilcoxon Vorzeichenrangtests	39

Abkürzungsverzeichnis

ACR – Agencia Colombiana para la Reintegración

DLE-NET – Diagnostik-Lifeline-Exposition-NET

EBPAS – Evidence-Based Practice Attitude Scale

FORNET – Narrative Exposure Therapy for Forensic Offender Rehabilitation

IMTEE - Integrated Model of Training Evaluation and Effectiveness

NET – Narrative Expositionstherapie

PTBS – Posttraumatische Belastungsstörung

REW – Reaktion-Einstellung-Wissen

TTEOF – Therapist Training Evaluation Outcomes Framework

1 Einleitung

Nach über 50 Jahren des bewaffneten Konflikts in Kolumbien ist es notwendig, die Bestrebungen für eine friedlichere Zukunft zu unterstützen. Neben vielen anderen nationalen und internationalen Akteuren nimmt die Agencia Colombiana para la Reintegración (ACR) eine zentrale Rolle in der Reintegration ehemaliger Kämpfer und Kämpferinnen der illegalen Gruppen von Guerilla und Paramilitär ein (Agencia Colombiana para la Reintegración, 2014b). Damit eine Reintegration in die Zivilgesellschaft möglich ist, müssen auch die psychischen Voraussetzungen gegeben sein. Neben Depression und Substanzmissbrauch bedeutet für die Betroffenen vor allem das Vorliegen einer Posttraumatischen Belastungsstörung (PTBS) besondere Probleme für die Reintegration (Maedl et al., 2010). Daher ist eine erfolgreiche Reduktion der PTBS-Symptomatik oftmals Wegbereiter für den weiteren Verlauf. Oder wie der ehemalige Paramilitär Carlos Cano es ausdrückt: „en realidad uno para cambiar tiene que demobilizar el corazon“¹ (Reintegración ACR, 2010). Die international als effektiv und effizient bewährte Kurzzeittherapie Narrative Expositionstherapie (NET) erscheint als geeignete Form, diese Aufgabe zu erfüllen. Der fundierten Therapieevaluation steht bislang lediglich eine Studie (Pfeiffer, 2011) gegenüber, die die Evaluation der Ausbildung von Therapeuten und Therapeutinnen beinhaltet. Pfeiffer (2011) konnte anhand einer internationalen Stichprobe aus Deutschland und Norwegen zeigen, dass sowohl der Inhalt als auch die Umsetzung des zehntägigen NET-Trainings sehr positiv beurteilt wurden. Neben dieser Komponente wurde allerdings bisher noch nicht untersucht, inwiefern tatsächlich neues Wissen durch das Training vermittelt wird und ob das Training zu Veränderungen im Verhalten in der Therapiepraxis der Therapeuten und Therapeutinnen führt. Um diese Lücke zu schließen wurde im Rahmen dieser Studie untersucht, ob das zehntägige Training in NET den Teilnehmenden kolumbianischen Psychologen und Psychologinnen es ermöglichte, die notwendigen Kenntnisse und Fähigkeiten zu erwerben, um die Therapie erfolgreich anzuwenden. Hierzu wurden mittels eines Fragebogens die Veränderungen hinsichtlich Einstellungen und Wissen von vor dem Training auf nach dem Training und zu einer Follow-Up-Messung sieben Monate später

¹ sinngemäß: „In Wahrheit: um sich zu ändern muss man zuerst das Herz demobilisieren“

erhoben und verglichen. Zusätzlich wurde die Umsetzung der Therapie sieben Monate nach dem Training beobachtet.

Bevor auf die genauen Hypothesen eingegangen wird, soll zunächst der theoretische Hintergrund näher beleuchtet werden.

1.1 Trainingsevaluation

Zu Beginn soll eine Klärung der Bestandteile des Begriffs *Trainingsevaluation* stattfinden. *Training* bezeichnet "the systematic acquisition of skills, rules, concepts or attitudes that result in improved performance in another environment" (Goldstein & Ford, 2002, S. 1) und *Evaluation* ist der Prozess zur Bestimmung des Wertes und der Leistung von Dingen (Scriven, 1991). Bezogen auf den Trainingskontext definieren Goldstein und Ford (2002):

Evaluation is the systematic collection of descriptive and judgmental information necessary to make effective training decisions related to the selection, adoption, value, and modification of various training activities. (Goldstein & Ford, 2002, S. 138)

Hinsichtlich der Ziele, Zwecke und Funktionen von Evaluation können nach Stockmann (2000) fünf zentrale Bereiche differenziert werden. Grundlegend nennt Stockmann zuerst die Erkenntnisfunktion, sprich Erkenntnisse über Eigenschaften und Wirkungen von Interventionen zu erfassen. Des Weiteren kann es von Bedeutung sein, Stärken und Schwächen bestimmter Interventionen herauszuarbeiten und zu ermitteln, wie diese beibehalten bzw. abgebaut werden können (Optimierungsfunktion). Die Kontrollfunktion von Evaluation beinhaltet die Überprüfung der Effektivität und Effizienz der Intervention(en) und kann als Grundlage der Entscheidungsfunktion dienen, im Hinblick auf die Frage, welche Interventionen weitergeführt und welche abgesetzt werden sollen. Nicht zuletzt hat Evaluation oftmals auch eine Legitimationsfunktion, zum Beispiel wenn es um die Verwendung öffentlicher Gelder geht (Bortz & Döring, 2006).

Eine vielfach getroffene Unterscheidung ist zudem die zwischen summativer und formativer Evaluation. Formative Evaluation wird normalerweise während der Entwicklung oder Verbesserung eines Programms durchgeführt und meist für den internen Gebrauch eingesetzt (Scriven, 1991). Im Gegensatz dazu wird nach Scriven (1991) summative Evaluation immer nach Abschluss oder Stabilisation eines Programms durchgeführt und ist

an eine externe Zuhörerschaft oder an Entscheidungsträger gerichtet, die aufgrund der Ergebnisse über das weitere Vorgehen entscheiden. Auf den Punkt gebracht beschreibt Robert Stake (zitiert nach Scriven, 1991) den Unterschied folgendermaßen: „When the cook tastes the soup, that’s formative; when the guests taste the soup, that’s summative.“ Scriven (1991) weist explizit darauf hin, dass summative Evaluation nicht mit Ergebnisevaluation verwechselt werden sollte. Ergebnisevaluation sei lediglich eine Evaluation, die auf die Ergebnisse fokussiert, dies könne allerdings summativ oder formativ sein. Damit weicht Scriven (1991) vom Verständnis dieser Konzepte, wie sie beispielsweise bei Goldstein und Ford (2002) oder Kanning und Staufenbiel (2012) beschrieben werden, ab. Da die begriffliche Trennung bei Scriven (1991) allerdings elaborierter erscheint, sollen die Konzepte im Folgenden in der hier beschriebenen Art verstanden werden. In diesem Sinne handelt es sich bei der vorliegenden Studie schwerpunktmäßig um eine formative Evaluation, die der weiteren Verbesserung des angebotenen Trainings dienen soll.

Ein erstes Modell zur Evaluation von Trainings beschrieb Kirkpatrick in einer Artikelserie vor nunmehr über 50 Jahren (Kirkpatrick, 1959a, 1959b, 1960a, 1960b). Die in den einzelnen Artikeln näher vorgestellten vier Schritte werden aufgrund ihrer Systematik und einfachen Verständlichkeit noch immer in vielen Lehrbüchern gelehrt (z. B. Landy & Conte, 2010, Kanning & Staufenbiel, 2012), in Evaluationen angewandt und vielfach zitiert (Alvarez, Salas & Garofano, 2004). In einem ersten Schritt werden die *Reaktionen* der Trainingsteilnehmenden auf das Training systematisch erhoben. Der nächste Schritt wird als *Lernen* bezeichnet und umfasst erworbenes Wissen, verbesserte Fertigkeiten und veränderte Einstellungen auf Grund des Trainings (Kirkpatrick, 1996). Schritt drei, *Verhalten*, umfasst die Verhaltensveränderung nach dem Training in der realen Arbeitsumgebung und wird oftmals auch als Trainingstransfer bezeichnet (Kirkpatrick, 1998). Im vierten und letzten Schritt werden die *Resultate* des Trainings evaluiert. Dies können alle möglichen gewünschten Ergebnisse wie gesteigerte Produktivität, geringere Kosten, höhere Qualität (Kirkpatrick, 1996) oder im Kontext eines Therapietrainings verringerte Symptomatik oder erhöhte Lebensqualität bei den Patienten und Patientinnen sein.

Lange wurden die vier Schritte aus dem Modell von Kirkpatrick als Hierarchie gesehen, bei der die verschiedenen Ebenen sowohl miteinander korreliert sind als auch kausal zusammenhängen, sodass beispielsweise positive Reaktionen zu Lernen, Lernen zu

verändertem Verhalten und dieses zu den erwünschten Resultaten führt und damit letztlich positive Ergebnisse auf der Reaktionsebene zu positiven Resultaten (Landy & Conte, 2010). Da noch dazu mit jedem weiteren Schritt die Komplexität und Kosten steigen (Kirkpatrick, 1996), werden nach einer Umfrage der American Society for Training and Development aus dem Jahr 2005 zu 91 Prozent Reaktionen, 54 Prozent Lernen, 23 Prozent Verhalten und lediglich 8 Prozent Resultate evaluiert (Sugrue & Rivera, 2005). Alliger und Janak (1989) zeigen allerdings sowohl auf theoretischer Ebene, wie in einem Literaturreview, dass diese Annahmen nur zum Teil haltbar sind. So fanden sie für Level 1 mit den anderen Levels nur sehr geringe mittlere Korrelationen ($r_{1,2} = .07$, $r_{1,3} = .05$ und $r_{1,4} = .48$, basierend auf nur einer Studie). Die mittleren Korrelationen zwischen Level 2, 3 und 4 fallen in den untersuchten Studien etwas höher aus ($r_{2,3} = .13$, $r_{2,4} = .40$ und $r_{3,4} = .19$). Anzumerken ist, dass in die Analyse lediglich 12 Studien und 26 Korrelationen einfließen, weshalb bei einer Generalisierung der Ergebnisse Vorsicht geboten ist. In einer weithin rezipierten Metaanalyse von Alliger, Tannenbaum, Bennett, Traver und Shotland (1997) zeigte sich jedoch ebenfalls, dass die einzelnen Schritte mit mittleren Korrelationen zwischen $r = .02$ bis $r = .26$ nur relativ gering miteinander assoziiert sind. Für diese zweite Untersuchung konnten die Autoren auf 34 Studien mit insgesamt 115 Korrelationen zurückgreifen. Wie diese Ergebnisse zeigen, reicht es nicht aus, lediglich ein oder zwei Schritte des Modells zur Evaluation anzuwenden, um eine Aussage über die anderen Schritte zu machen. Oder wie Alliger et al. (1997) es ausdrücken „liking does not equate to learning or to performing“ (S. 353). Nach Möglichkeit sollten also stets alle Ebenen erfasst werden.

Ein weiteres Modell stammt von Tannenbaum, Cannon-Bowers, Salas und Mathieu (1993). Es umfasst sowohl die vier Schritte von Kirkpatrick, wobei Verhalten in *Trainings-* und *Transferleistung* unterteilt wird, als auch eine Vielzahl weiterer Komponenten. So gehen die Autoren näher auf die Rolle von individuellen Charakteristika (Fähigkeiten, Einstellungen, Selbstwirksamkeit, etc.), von Organisations- und Situationsvariablen (Organisationsklima, Trainee Auswahl, etc.) und von Erwartungen, Trainingsmotivation sowie Trainingscharakteristika (Methode, Inhalt, etc.) ein. Auf den Ergebnissen von Alliger und Janak (1989) aufbauend nehmen Tannenbaum et al. (1993) in ihrem Modell keine Verbindung zwischen Reaktionen und Posttraining-Einstellungen zu anderen Bestandteilen des Modells an – wohl aber eine hierarchische Verbindung von Lernen zu Trainingsleistung, von Trainingsleistung zu

Transferleistung und schließlich von Transferleistung zu Resultaten. Sie weisen aber ebenso darauf hin, dass sie das jeweils Vorangehende als notwendig, nicht aber als hinreichend für das jeweils Folgende sehen.

Eine umfassende Kritik erhält das Modell von Kirkpatrick (1996) in einem Artikel von Holton (1996), in welchem Holton darauf hinweist, dass die vier Schritte oder Ebenen von Kirkpatrick nicht den Anforderungen eines Modells genügen, sondern vielmehr eine Taxonomie seien. Da Taxonomien nicht alle Konstrukte des interessierenden Phänomens identifizierten, sei eine Validierung unmöglich und die Ergebnisse von Alliger und Janak (1989) somit nicht weiter überraschend. In seinem eigenen Modell schlägt Holton (1996) nun drei primäre Ergebnismaße vor: *Lernen*, *individuelle Leistung* und *Organisationsresultate*. Aufbauend auf den Ergebnissen von Mathieu, Tannenbaum und Salas (1992) werden Reaktionen in diesem Modell als Moderator bzw. Mediator zwischen Lernmotivation und Lernen gesehen und nicht mehr als Outcomevariable. Holton (1996) nimmt an, dass Lernen unter gewissen Bedingungen die individuelle Leistung positiv beeinflusst und diese wiederum die Organisationsresultate. Die Stärke von diesem Modell ist darin zu sehen, dass es versucht, diese Bedingungen zu isolieren. Für jedes der drei Ergebnismaße schlägt Holton (1996) Einflussfaktoren auf den Ebenen von Motivation (z. B. Lernmotivation), Umweltfaktoren (z. B. Transferklima) und Fähigkeiten bzw. ermöglichende Elemente (z. B. Transferdesign) vor. Damit ist dieses Modell nicht nur zur Ergebnisevaluation eines Trainings geeignet, sondern auch zur analytischen Evaluation und zur Beantwortung der Frage, wieso das Training das gewünschte Ergebnis erzielen konnte bzw. wieso nicht.

Basierend auf den bereits dargestellten Kritikpunkten an Kirkpatrick's Taxonomie schlägt Kraiger (2002) ebenfalls ein eigenes Modell mit drei Zielbereichen vor. Dieses Modell soll vor allem auch der Konfundierung von Datenerhebungsmethode und Evaluationsziel vorbeugen, indem diese klar unterschieden werden. In dem ersten Bereich *Trainingsinhalt und -design* werden das Design, die Validität sowie die Art der Überlieferung z. B. über Experten-Urteile oder Reaktionen der Teilnehmenden überprüft. Diese Evaluation dient vor allem als Feedback für die Trainer und Trainerinnen. Der zweite Bereich *Veränderungen in den Lernenden* erfasst kognitive, affektive und behaviorale Veränderungen über Interviews, Umfragen, Arbeitsproben etc. Die Ergebnisse dienen als Feedback für die Teilnehmenden sowie als Grundlage für weitere Entscheidungen bezüglich kommender Trainings. Transfer,

Resultate und Leistung werden unter dem Begriff *Organisationales Payoff* in dem dritten Zielbereich untersucht. Dabei schlägt Kraiger (2002) zur Evaluation unter anderem Kosten-Nutzen-Analysen und Umfragen vor. Die Evaluation des organisationalen Payoffs kann nach Kraiger (2002) als Feedback für die Teilnehmenden, zur Entscheidungsfindung sowie zu Marketingzwecken verwendet werden.

Neben der reinen Evaluation, was aufgrund des Trainings geschehen ist, wird in der Literatur vorgeschlagen, dass zudem die Trainingseffektivität untersucht werden sollte. Noe und Colquitt (2002) verstehen darunter die Prozesse, die vor, während und nach dem Training auftreten und die Wahrscheinlichkeit, dass dieses einen Einfluss hat, erhöhen. Während Trainingsevaluation die Ergebnisse beschreibt, was sich verändert hat, versucht Trainingseffektivität zu erklären, warum es dazu aufgrund der spezifischen Faktoren der Teilnehmer und Teilnehmerinnen, des Trainings und der Umwelt bzw. Organisation dazu kam (Alvarez et al., 2004). Nur wenn die verschiedenen Komponenten berücksichtigt und verknüpft werden, kann ein Training effektiv sein.

Um die Annahmen der vier hier beschriebenen theoretischen Modelle zu vereinen und zusätzlich mit den Ergebnissen der Forschung in Einklang zu bringen schlagen Alvarez et al. (2004) das Integrated Model of Training Evaluation and Effectiveness (IMTEE) vor, welches im Folgenden beschrieben werden soll.

1.1.1 Integrated Model of Training Evaluation and Effectiveness

Für die Erstellung ihres Modells untersuchten Alvarez et al. (2004) die Ergebnisse von insgesamt 73 empirischen Studien, die zwischen 1993 und 2002 publiziert wurden. Es wurden lediglich diejenigen Variablen berücksichtigt, die in mindestens drei Studien signifikante Zusammenhänge ($p < .05$) in der gleichen Richtung (positiv bzw. negativ) fanden. Bei überwiegend gemischten Beziehungen über verschiedene Studien hinweg (negativ, positiv und kein Zusammenhang) wurden die entsprechenden Variablen nicht in das Modell übernommen. Das entstandene Modell und die Beziehungen der verschiedenen Konstrukte sind in Abbildung 1 zu sehen.

Das IMTEE besteht aus vier Ebenen. Auf der obersten Ebene steht die Bedarfsanalyse. Hierbei werden sowohl eine Analyse der Organisation, mitsamt den langfristigen Ziele und einzigartigen Umstände, als auch der spezifischen Aufgaben, die im Training trainiert werden

sollen, sowie der zu trainierenden Personen empfohlen (Kraiger, 2002). Trainingsangebote sollten dabei so entworfen werden, dass Trainingsinhalt & -Design positive Veränderungen in den Lernenden und dem organisationalen Payoff hervorrufen (Kraiger, 2002).

Die folgenden zwei Ebenen des Modells vereinen laut Alvarez et al. (2004) die vier Trainingsevaluationsmodelle von Kraiger (2002), Kirkpatrick (1996), Holton (1996) und Tannenbaum et al. (1993). Dabei dienen die drei Konstrukte von Kraiger (2002) als analytische Oberkategorien für die sechs Konstrukte aus den Evaluationsmodellen der anderen drei Autoren. Trainingsinhalt und -Design können evaluiert werden, indem die Reaktionen auf das Training gemessen werden (Alvarez et al., 2004). Die Veränderungen in den Lernenden werden durch die *Einstellungen nach dem Training*, das *kognitive Lernen* sowie die *Trainingsleistung* erfasst. Das Konzept organisationales Payoff wird operationalisiert durch die Messung der Konstrukte Transferleistung und Resultate. Die Autoren des IMTEE weisen darauf hin, dass diese sechs Konstrukte keineswegs die einzigen Operationalisierungen für die drei Bereiche von Kraiger (2002) sind, sondern vielmehr den Stand der Forschung widerspiegeln. Weitere Evaluationsmethoden sind nach Alvarez et al. (2004) bereits vorgeschlagen worden, doch habe es zum Zeitpunkt der Erstellung des IMTEE noch zu wenige Studien gegeben, die diese Methoden angewandt haben.

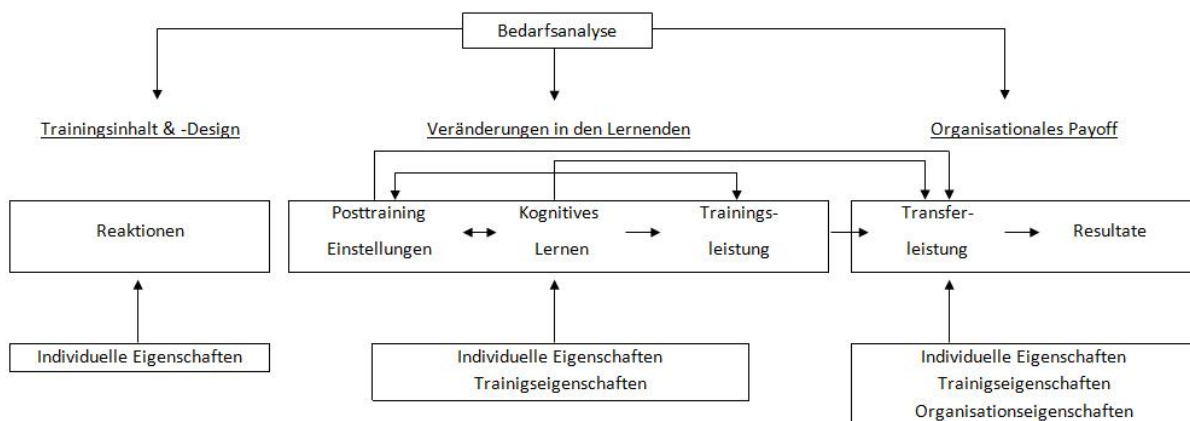


Abbildung 1. Integrated Model of Training Evaluation and Effectiveness nach Alvarez et al. (2004)

Da nach dem Stand der Forschung positive affektive Reaktionen keinen Zusammenhang mit Lernen oder anderen Ergebnismaßen haben (Alliger et al., 1997; Holton, 1996; Tannenbaum et al., 1993), wohl aber Reaktionen in Bezug auf Nützlichkeit des Trainings mit

Lernen ($r = .26$) bzw. Transfer ($r = .18$), werden diese im IMTEE dafür verwendet, die Angemessenheit des Traininginhalts- und Designs zu evaluieren (Alvarez et al., 2004).

Obwohl verschiedene Posttraining-Einstellungen in der Literatur vorgeschlagen wurden (z. B. von Tannenbaum et al., 1993 und Holton, 1996) erfüllten nur die Ergebnisse der Studien zu Posttraining-Selbstwirksamkeit die Inklusionskriterien. Dabei zeigten sich positive Korrelationen zwischen Posttraining-Selbstwirksamkeit und kognitivem Lernen, Trainingsleistung sowie Transferleistung. Da es sich lediglich um Korrelationen handelt, könnte die Beziehung zwischen Posttraining-Selbstwirksamkeit und kognitivem Lernen sowie Posttraining-Selbstwirksamkeit und Trainingsleistung reziprok sein. Eine Richtung ist daher im IMTEE nicht näher spezifiziert.

Kognitives Lernen sehen Alvarez et al. (2004) als kognitive Aneignung von Wissen, das nach den Ergebnissen von Alliger et al. (1997) und Tannenbaum et al. (1993) neben Posttraining-Selbstwirksamkeit auch mit Trainingsleistung und Transferleistung in Zusammenhang steht.

Trainingsleistung ist die Fähigkeit eine neu erworbene Fertigkeit am Ende des Trainings und noch vor dem Transfer in die tatsächliche Arbeitsumgebung durchführen zu können (Alvarez et al., 2004). Nach dem IMTEE wird sie beeinflusst von dem Ausmaß des erworbenen Wissens sowie der Posttraining-Selbstwirksamkeit.

Die Transferleistung, definiert als „behavioral changes on the job as a result of training“ (Alvarez et al., 2004, S. 397), ist in dem IMTEE abhängig von Posttraining-Selbstwirksamkeit, dem kognitiven Lernen sowie der Trainingsleistung.

Das letzte Ergebnismaß in dem IMTEE sind die Resultate, die nach Tannenbaum et al. (1993) als „quantifiable changes in related outcomes as a result of trainees’s behavioral changes“ (S. 24) gesehen werden sollten und in positiver und kausaler Beziehung zur Transferleistung stehen. Alvarez et al. (2004) weisen aber auch darauf hin, dass diese Beziehung durch die Qualität der Bedarfsanalyse moderiert wird. Kurz: wenn das Training mit den Organisationszielen im Einklang steht, ist ein Anstieg der organisationalen Effektivität wahrscheinlich (Alvarez et al., 2004).

Die letzte Ebene des IMTEE umfasst die drei breiten Konzepte zur Effektivität des Trainings. Individuelle Eigenschaften sind u. a. Selbstwirksamkeit, Motivation, Erfahrung, Alter und Zielorientierung (Noe & Colquitt, 2002). Bezüglich des Trainings selbst werden z. B. Lernprinzipien, Schwierigkeit und Eigenschaften der Trainer bzw. Trainerinnen angesprochen (Tannenbaum et al., 1993). Zu den Eigenschaften der Umwelt bzw. Organisation in der später das Erlernete umgesetzt werden soll, werden von Noe und Colquitt (2002) unter anderem Organisationsklima, organisationale Gerechtigkeit und die Möglichkeit, das Erlernete überhaupt zu zeigen (opportunity to perform) angeführt. Alvarez et al. (2004) weisen darauf hin, dass es hierzu insgesamt noch zu wenig Forschung gibt, um alle beeinflussenden Eigenschaften und deren Interaktionen und Beziehungen aufzuzeigen.

Die Ebene der Ergebnismaße des IMTEE wird unter etwas anderen Bezeichnungen von Decker, Jameson und Naugle (2011) in ihrem Therapist Training Evaluation Outcomes Framework (TTEOF) für den Kontext der Therapietrainingsevaluation spezifiziert und näher untersucht. Die für diese Arbeit relevanten Erkenntnisse sollen im Folgenden kurz dargestellt werden.

1.1.2 Therapist Training Evaluation Outcomes Framework

Das TTEOF baut wie viele andere Evaluationsmodelle auf der Taxonomie von Kirkpatrick (1996) auf. Decker et al. (2011) spezifizieren jedoch die zweite Ebene näher, indem sie Lernen aufspalten in Veränderungen in *Einstellungen*, *Wissen* und *Fertigkeiten*. Zudem wird die Ebene Verhalten umbenannt in *Verhaltensveränderung in der Praxis* und Resultate in *Therapieerfolg*². Damit hat das TTEOF sechs verschiedene Ebenen (siehe Tabelle 1). Konzeptuell treffen sie damit, wie bereits erwähnt, ziemlich genau die Ebenen von Alvarez et al. (2004).

Ein großer Teil des Beitrags von Decker et al. (2011) ist darin zu sehen, dass sie die Methoden und Ergebnisse von 34 Trainingstherapieevaluationsstudien zusammengetragen haben, die zwischen 1997 und 2010 publiziert wurden. Einige dieser Methoden und Ergebnisse werden im Folgenden entsprechend der Ebenen näher besprochen. Hierbei ist auffallend, dass keine der Studien alle Ebenen evaluiert, vier Studien berücksichtigen fünf Ebenen, eine Studie vier Ebenen, fünf Studien drei Ebenen, 13 Studien zwei Ebenen und

² Im Englischen Original *client outcome*.

neun Studien lediglich die Veränderungen auf einer Ebene. Der Zusammenhang zwischen dem Therapieerfolg und anderen Ebenen des Modells wird von lediglich einer Studie untersucht.

Tabelle 1

Therapist training evaluation outcomes framework (nach Decker et al., 2011)

Evaluationsebene	Kurz- oder Langzeit	Methoden
1 – Reaktionen zum Training	Kurz	Therapeuten-Selbstbericht
2a – Einstellungen	Kurz	Therapeuten-Selbstbericht
2b – Wissen	Kurz	Fragebogen Lösen von Fallbeispielen
2c – Fertigkeiten	Kurz	Beobachter bewertetes Rollenspiel
3 – Verhaltensveränderung in der Praxis	Lang	Therapeuten selbstberichtete Anwendung Therapeuten selbstberichtete Manualtreue Chart Review Klienten bewertete Manualtreue Beobachter bewertete Manualtreue und Kompetenz
4 – Therapieerfolg	Lang	Klientenbericht über Symptome oder Funktionsniveau Therapeutenbericht des Funktionsniveaus

1.1.2.1 Reaktionen

Bei dem Review von Decker et al. (2011) zeigte sich, dass für die Beurteilung der Reaktionen lediglich Selbstberichtsmaße der Teilnehmenden verwendet wurden. Da es sich bei den Reaktionen um die Zufriedenheit mit dem Training sowie die wahrgenommene Nützlichkeit des Trainings (Alliger et al., 1997) und damit um subjektseitig definierte Konzepte handelt, ist dies nicht weiter verwunderlich. Die Erhebung erfolgte in den meisten Studien über Skalen mit einem Item pro Facette. So unterscheiden Seng, Prinz und Sanders (2006) beispielsweise zwischen Zufriedenheit mit dem Trainingsinhalt: „How would you rate the content of this course?“ und Gesamtzufriedenheit: „In an overall sense, how satisfied were you with the course?“. Manche Studien verwenden allerdings auch umfangreichere Skalen zur Erfassung der Zufriedenheit mit dem Training (z. B. Herschell et al., 2009).

1.1.2.2 Einstellungen

In den untersuchten Studien wurden die Einstellungen der Therapeuten und Therapeutinnen als Einstellungen gegenüber bestimmten Interventionen, evidenzbasierten Verfahren generell oder die Fähigkeit eine neu gelernte Intervention anwenden zu können konzeptualisiert (Decker et al., 2011). Hierfür wurden verschiedene Skalen eingesetzt. Für die Einstellungen gegenüber evidenzbasierten Verfahren sind dies beispielsweise die *Evidence-Based Practice Attitude Scale* von Aarons (2004) oder die *Attitudes Toward Treatment Manuals Scale* von Addis und Krasnow (2000). Zusammenfassend empfehlen Decker et al. (2011) die relativ kurzen Skalen sowohl vor und nach dem Training als auch bei Follow-Up-Messungen einzusetzen, da Einstellungsveränderungen über die Zeit auftreten können.

Ebenso zählen Decker et al. (2011) die wahrgenommene eigene Kompetenz bzw. Selbstwirksamkeit in der Implementierung der neu erlernten Techniken zu den Einstellungen. Insgesamt berichten nur wenige der untersuchten Studien Zusammenhänge zwischen wahrgenommener Kompetenz und anderen Variablen. Seng et al. (2006) zeigten einen geringen Anstieg von vor dem Training auf nach dem Training doch ergaben sich keinerlei signifikante Unterschiede zwischen denjenigen, die das Training abgeschlossen und denjenigen, die es abgebrochen haben.

1.1.2.3 Wissen

Zur Überprüfung des Wissens der Trainingsteilnehmerinnen und -teilnehmer wurden in den untersuchten Studien sowohl Fragebögen als auch Fallbeispiele verwendet. Bei beiden Methoden zeigte sich eine erwartungsgemäß signifikante Steigerung des Wissens von vor dem Training auf nach dem Training. Positive Zusammenhänge von Wissen und Fertigkeiten wurden von einigen der untersuchten Studien berichtet, jedoch nicht statistisch berechnet (Decker et al., 2011).

1.1.2.4 Fertigkeiten

Um die Fertigkeiten der Trainingsteilnehmerinnen und -teilnehmer zu evaluieren, wurden in den untersuchten Studien Rollenspiele aufgezeichnet (Audio bzw. Video) und anschließend mit entsprechenden Kodiersystemen ausgewertet. Zwar überprüften die eingegangenen Studien nicht, inwiefern die in Rollenspielen gezeigten Fertigkeiten auch dem Verhalten in der Praxis entsprechen, dennoch empfehlen Decker et al. (2011) diese einzusetzen, da Rollenspiele unter anderem die Vorteile mit sich bringen, dass keine

Einwilligung eines Patienten oder einer Patientin vorliegen muss und dass die Beispiele so gestaltet werden können, dass die Teilnehmerinnen und Teilnehmer die volle Bandbreite des im Training Vermittelten zeigen können.

1.1.2.5 Verhaltensveränderungen in der Praxis

Zur Erfassung der Verhaltensveränderung in der Praxis gibt es eine Vielzahl von Annäherungsversuchen. Eine der einfachsten Möglichkeiten ist, die Therapeutinnen und Therapeuten zu fragen, ob sie die neu erlernte Intervention überhaupt anwenden (wie beispielsweise bei Seng et al., 2006). Die Reliabilität und Validität dieser Vorgehensweise wurde allerdings nicht überprüft und so wird weitere Forschung benötigt, um diesen Ansatz empfehlen zu können (Decker et al., 2011).

Eine weitere Möglichkeit ist die Überprüfung der Manualtreue der Therapierenden. In der Literaturübersicht von Decker et al. (2011) zeigte sich, dass in den verschiedenen Studien dazu entweder Selbstberichte, Beobachterratings oder Klientenratings verwendet wurden. Bei einer Untersuchung von Carroll, Nich und Rounsaville (1998) zeigte sich allerdings, dass die Übereinstimmung von Selbstbericht und Beobachterrating lediglich schlecht bis ausreichend war. Zu ähnlichen Erkenntnissen gelangen auch Martino, Ball, Nich, Frankforter und Carroll (2009). Zwar zeigte sich bei der Beurteilung der grundlegenden Implementierung der trainierten Motivational Enhancement Therapy eine gute bis sehr gute Übereinstimmung (74-97%) zwischen Therapeut_innen, Supervisor_innen und unabhängigen Beobachter_innen, bei der Hinzunahme von fortgeschrittenen Techniken verringerte sich die Übereinstimmung jedoch auf unter 65 Prozent. Aufgrund dieser Ergebnisse empfehlen Decker et al. (2011) für die Bewertung der Verhaltensveränderung in der Praxis nicht allein auf Selbstberichte zu bauen. In verschiedenen Studien zeigte sich für die Beobachterratings jedoch, dass durch gutes Beobachtertraining (20 Stunden bei Morgenstern, Morgan, McCrady, Keller & Carroll, 2001; 44 Stunden bei Martino et al., 2009) eine hohe Inter-Rater-Reliabilität erreicht werden kann. Zudem zeigten sich signifikante Korrelationen zwischen der Manualtreue und Kompetenz der Therapeut_innen und der Motivationsveränderung der Klient_innen ($.13 < r < .22$).

1.1.2.6 Therapieerfolg

Instrumente zur Erhebung der Symptome und des Funktionsniveaus der Klienten und Klientinnen sind für viele Bereiche in der klinischen Psychologie in großem Ausmaß

vorhanden und bekannt. Für den Bereich der Posttraumatischen Belastungsstörung seien an dieser Stelle nur einige exemplarisch genannt. Die Posttraumatic Stress Diagnostic Scale (PDS) von Foa (1995) ist beispielsweise ein häufig eingesetztes Selbstratinginstrument und die Clinician Administered PTSD Scale (CAPS) gilt als aktueller Gold-Standard in der PTSD Diagnose (Schauer, Neuner & Elbert, 2011).

Abseits von Symptomen und dem Funktionsniveau sollten aber auch Bereiche untersucht werden, die für die Klienten und Klientinnen von großer Bedeutung sind, wie etwa Beschäftigungsstatus, Stabilität im Haushalt, Erreichen von persönlichen Zielen etc. (Decker et al., 2011).

Decker et al. (2011) schließen mit der Beobachtung, dass es für den Bereich der Evaluation von Therapietrainings bereits eine Fülle von Herangehensweisen gibt und plädieren wie Beidas und Kendall (2010) dafür, dass für ein umfassendes Bild stets verschiedene Methoden und Informationsquellen zur Evaluation herangezogen werden sollten. Darüber hinaus empfehlen sie zudem, dass möglichst viele Ebenen des TTEOF berücksichtigt werden sollten.

1.2 Bedarfsanalyse

Damit ein Training effektiv sein kann, sollte zunächst entschieden werden, was und wer trainiert werden soll (Salas & Cannon-Bowers, 2001). In der Phase der Bedarfsanalyse wird hierfür untersucht, welche spezifischen Anforderungen das Organisationsklima und die Organisationsziele, die Eigenschaften der Trainingsteilnehmenden sowie die zu lernenden Aufgaben und Fertigkeiten mit sich bringen (Alvarez et al., 2004, Goldstein & Ford, 2002).

Das in dieser Arbeit untersuchte Training bezieht sich auf Traumatherapie bei ehemaligen Mitgliedern von Guerilla und Paramilitär in Kolumbien, die von Therapeuten und Therapeutinnen der Agencia Colombiana para la Reintegración (ACR) durchgeführt wird. Wie es zu dem Bedarf an diesem Training kam, soll anhand eines kurzen historischen Abrisses dargelegt werden.

1.2.1 Bewaffneter Kampf in Kolumbien

Der bewaffnete Kampf in Kolumbien ist einer der am längsten anhaltenden Konflikte und mit mindestens 220 000 Toten auch einer der blutigsten (Grupo de Memoria Histórica, 2013). Der Charakter des Konflikts ist vielseitig und hat sich über die Jahre hinweg immer

wieder gewandelt. Die Autoren der Grupo de Memoria Histórica (2013) unterteilen den Konflikt in vier Phasen. Die erste Phase (1958 - 1982) ist gekennzeichnet durch die Verbreitung der Guerillas. Während der zweiten Phase (1982 - 1996) kommt es zu einer weiteren Expansion der Guerillas, dem Aufkommen der Paramilitärs, einer Verbindung mit dem Drogenhandel von beiden Seiten, einem teilweisen Zusammenbruch des kolumbianischen Staates sowie zu Friedensprozessen und demokratischen Reformen. In der dritten Phase (1996 - 2005) verhärtet sich der Konflikt zwischen Guerillas, Paramilitärs und dem Staat und die politische und öffentliche Meinung verlangt nach einer militärischen Lösung des Konflikts. Die vierte Phase (2005 - 2012) umfasst die militärischen Offensiven des Staates gegen die Guerilla sowie den Prozess der Demobilisierung der paramilitärischen Gruppen. Aktuell befindet sich der Konflikt vielleicht im Beginn einer weiteren Phase, die durch die Friedensverhandlungen der kolumbianischen Regierung mit der größten noch existierenden Guerilla Fuerzas Armadas Revolucionarias de Colombia (FARC) gekennzeichnet ist.

1.2.2 Demobilisierungsprozess

So unterschiedlich die Phasen des Konfliktes, so unterschiedlich waren auch die Lösungsversuche. Immer wieder wechselte die Strategie der Regierung von Agrarreformen, über Friedensverhandlungen, zu militärischen Offensiven. Die Demobilisierung der Guerrillas Movimiento 19 de Abril (M-19 im Jahr 1990) und Ejército Popular de Liberación (EPL im Jahr 1991) führten zur Gründung des Consejo Nacional de Normalización, das die Reintegrationsprogramme koordinieren und überwachen sollte (Contreras Rolón, Jaramillo de Botero & Picón Amaya, 2006).

Im Jahr 2003 wurde von dem Innen- und Justizministerium Kolumbiens das Programa para la Reincorporación de la Vida Civil ins Leben gerufen. Dieses Programm sollte zur Wiedereingliederung der Personen dienen, die sich individuell für eine Demobilisierung entschlossen hatten. Dafür erhielten die Personen psychosoziale und finanzielle Unterstützung, Fortbildungen sowie Zugang zum nationalen Gesundheitssystem. Dieses auf kurzfristige Erfolge ausgerichtete Programm funktionierte bis 2006. Als die Anzahl der Demobilisierten durch die kollektive Demobilisierung des paramilitärischen Dachverbandes Autodefensas Unidas de Colombia (AUC) aber weiter anstieg, ergaben sich administrative und konzeptuelle Probleme (Agencia Colombiana para la Reintegración, 2014d). Daher

wurde im September 2006 die Alta Consejería Presidencial para la Reintegración zur „Abrüstung, Demobilisierung und Reintegration“ (Desarme, Desmovilización y Reintegración, DDR) ins Leben gerufen. Weiterhin änderte sich nun auch die Strategie von einer kurzfristigen hin zu einer langfristigen. Im September 2011 wurde die Alta Consejería Presidencial para la Reintegración in die Agencia Colombiana para la Reintegración umgewandelt. Dieser institutionelle Wechsel erhöhte die administrative und finanzielle Autonomie und sollte zur Kontinuität und Bestärkung der Reintegrationspolitik in Kolumbien beitragen (Agencia Colombiana para la Reintegración, 2014d).

Heute ist das Reintegrationsprogramm in Kolumbien mit über 56.000 registrierten Demobilisierten eines der größten und am besten entwickelten der Welt (Eder, 2014). Dazu gehört auch das besondere Augenmerk auf die psychosoziale Unterstützung, die aufgrund der besonderen psychischen Belastung der ehemaligen Kämpfer und Kämpferinnen zentral für das Gelingen der Reintegration ist (Maedl et al., 2010; Morgenstein, 2008; Moor, 2007; Odenwald, Hinkel & Schauer, 2007). Während eine qualitative Untersuchung der Perspektive der Demobilisierten aus den Anfangstagen des Programms zu dem Schluss kommt, dass dieser Aspekt noch ungenügend implementiert war (Anaya, 2007), so zeigen die Statistiken der ACR für das Jahr 2013, dass von den 30.612 betreuten Personen 30.594 (99.9%) psychosoziale Unterstützung in Anspruch nahmen (Agencia Colombiana para la Reintegración, 2014c). Bei einer näheren Untersuchung, was unter den Bereich der psychosozialen Unterstützung fällt, so findet man in der entsprechenden Resolution, dass es darum gehe, die Kapazitäten zu entwickeln, ein eigenes Lebensprojekt zu verfolgen und selbstständig bürgerlichen Aufgaben nachzukommen (Agencia Colombiana para la Reintegración, 2013). Nach dem Modelo de Atención Psicosocial para la Paz (MAPAZ) der ACR aus dem Jahr 2011 wird weiter zwischen individuellen Aktivitäten (z. B. Hausbesuche und persönliche Beratung) und Gruppenaktivitäten (z. B. Workshops und Familienberatung) unterschieden. Dabei gehe es darum, die Demobilisierten bei der Identifikation und Lösung von Konflikten in ihrem sozialen Kontext zu unterstützen, wobei explizit darauf hingewiesen wird, dass „el servicio psicosocial para la ACR no es un programa de atención en psicología clínica³“ (Agencia Colombiana para la Reintegración, 2011, S. 11). Im weiteren Verlauf der

³ Eigene Übersetzung: Das Angebot der psychosozialen Unterstützung ist für die ACR kein Angebot in klinischer Psychologie bzw. Therapie.

Arbeit der ACR zeigte sich allerdings, dass die bisherigen Angebote der psychosozialen Unterstützung nicht ausreichend waren, weshalb im Jahr 2013 zusätzlich eine spezifisch klinisch-psychologische Unterstützung mit in das Angebot für die Demobilisierten aufgenommen wurde (Agencia Colombiana para la Reintegración, 2014a).

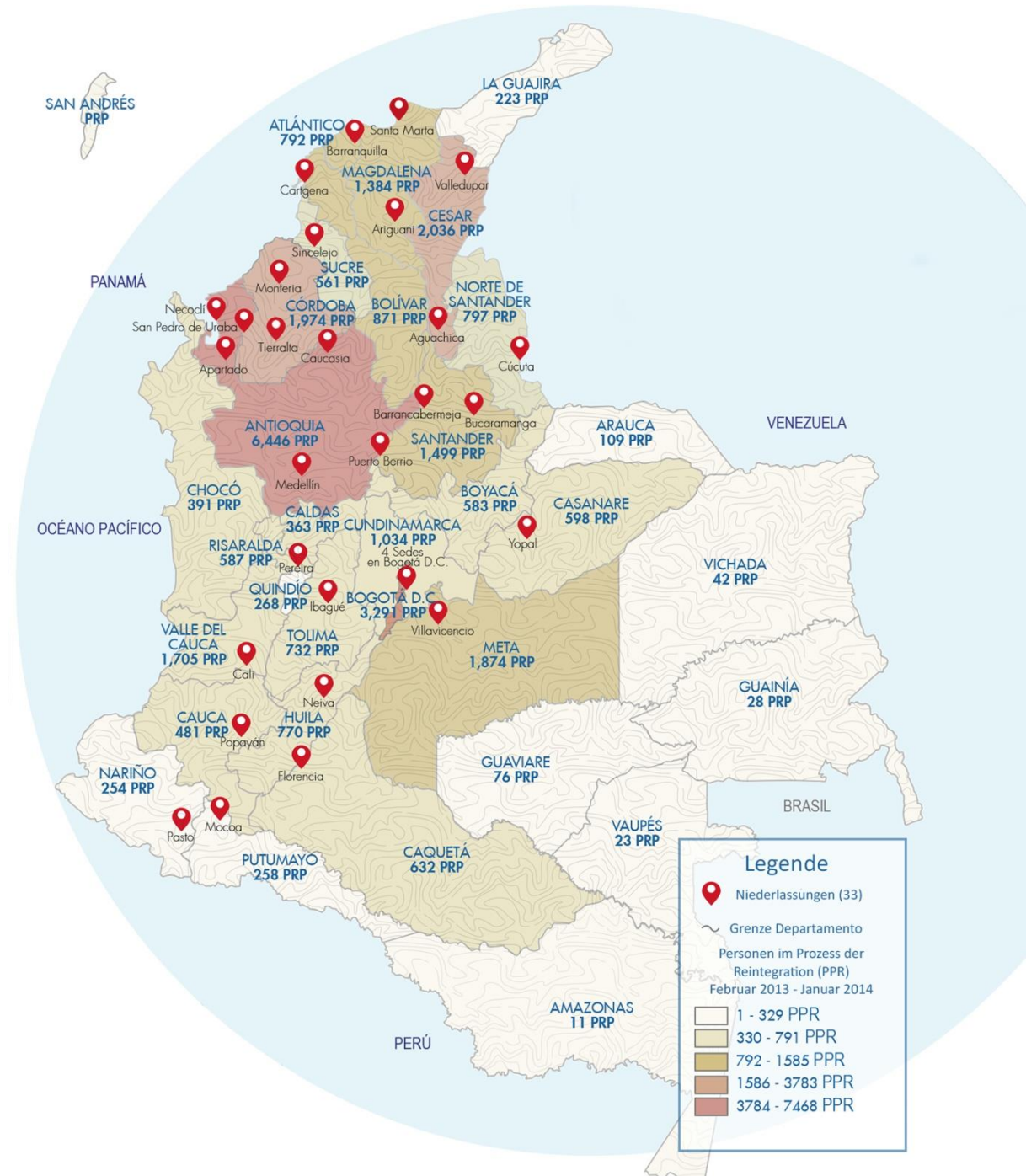


Abbildung 2. Übersicht Niederlassungen der ACR sowie Anzahl der Personen im Prozess der Reintegration (PPR). Adaptiert nach Agencia Colombiana para la Reintegración (2014c).

In jeder der insgesamt 33 landesweiten Niederlassungen der ACR (Abbildung 2) ist daher seit 2013 ein klinischer Psychologe bzw. eine klinische Psychologin vorgesehen. Diese als *Reintegrador III* bezeichnete Stelle erweitert das Spektrum der psychosozialen Unterstützung durch Psycholog_innen und Sozialarbeiter_innen (*Reintegrador II*) um die Behandlung psychischer Probleme (Agencia Colombiana para la Reintegración, 2014a).

Nach Angaben der ACR (2014a) nimmt die Posttraumatische Belastungsstörung dabei eine besondere Stellung ein. Auch eine Studie von Bueno Castellanos (2013) zeigte, dass unter den Demobilisierten 24 % die Diagnosekriterien einer PTBS erfüllen. Um den besonderen Anforderungen dieser Betroffenenengruppe gerecht zu werden, gab es innerhalb der ACR den Bedarf an einer sowohl effektiven, als auch effizienten Therapie. Diese beiden Eigenschaften wurden von der Narrativen Expositionstherapie (NET) generell über verschiedene Länder und Zielgruppen hinweg immer wieder erfüllt (vgl. Bichescu, Neuner, Schauer & Elbert, 2007; Ruf et al., 2010; Schaal, Elbert & Neuner, 2009) und auch spezifisch für die Zielgruppe in Kolumbien zeigte sich bei Bueno Castellanos (2014) eine erfolgreiche Reduktion der PTBS Symptomatik durch die Anwendung der NET.

1.2.3 Narrative Expositionstherapie

Basierend auf der Erfahrung, dass in Krisen- und Kriegsgebieten finanzielle und zeitliche Ressourcen beschränkt sind und eine mögliche Therapie auch lokalen Helfern mit wenig oder keinem medizinischen bzw. psychologischen Hintergrund vermittelt werden können soll, wurde die NET auf Basis der Theorie des Furcht/Trauma-Netzwerkes, kognitiv-verhaltenstherapeutischen Prinzipien und der Testimony-Methode entwickelt (Schauer et al., 2011).

Die grundlegende Idee ist, dass für eine erfolgreiche Reduktion der PTBS-Symptome sowohl eine Habituation der emotionalen Reaktion mittels Exposition, sowie eine Rekonstruktion des autobiographischen Gedächtnisses und damit des Narrativs der traumatischen Erinnerungen erfolgen müssen (Schauer et al., 2011). Hierfür werden zu den identifizierten traumatischen Erlebnissen die sensorischen, emotionalen und kognitiven Erinnerungen in Raum und Zeit verortet (Elbert, Hermenau, Hecker, Weierstall & Schauer, 2012). In der NET werden dabei drei verschiedene Phasen durchschritten. Nach der ersten Phase der Diagnostik und Psychoedukation folgt in der zweiten Phase bzw. Therapiesitzung das Erstellen der Lebenslinie des Klienten bzw. der Klientin. Sie soll einen ersten Überblick

über das Leben der Klient_innen geben, indem in chronologischer Reihenfolge Blumen für positive und Steine für angstbesetzte und traumatische Erlebnisse auf einer Linie (z. B. eine Schnur) niedergelegt und mit einem kurzen Titel versehen werden (Schauer et al., 2011). Für die Arbeit mit Tätern und Täterinnen werden in der FORNET (Narrative Exposure Therapy for Forensic Offender Rehabilitation) zusätzlich Stöckchen als Symbol für eigene gewalttätige Handlungen eingesetzt (Elbert et al., 2012). Ab der dritten Sitzung beginnt die dritte Phase, die eigentliche Exposition mit den traumatischen Erfahrungen. Die NET ist als Kurzzeittherapie konzipiert mit gewöhnlicher Weise zehn Sitzungen à 90-120 Minuten.

1.3 Hypothesen

Aufgrund der Ergebnisse von Bueno Castellanos (2014) wird angenommen, dass die NET eine geeignete Therapie zur Behandlung der PTBS-Symptomatik auch in dem gegebenen Kontext darstellt und somit zu den Organisationszielen der ACR beiträgt. Damit die NET jedoch erfolgreich praktiziert werden kann, ist ein Training der Therapeut_innen vonnöten, welches diesen die notwendigen Grundlagen und Fertigkeiten vermittelt, um die Traumatherapie durchzuführen. Um die Qualität und Wirksamkeit des Trainings zu überprüfen, werden basierend auf dem IMTEE (Alvarez et al., 2004) und TTEOF (Decker et al., 2011) im Rahmen dieser Arbeit die folgenden Hypothesen aufgestellt und überprüft (siehe auch Abbildung 3).

Grundlegend wird angenommen, dass sich durch das Training Veränderungen in den Lernenden ergeben, so dass sich über die drei Erhebungszeitpunkte vor dem Training, direkt nach dem Training und zu einer Follow-Up-Erhebung signifikante Unterschiede in den Einstellungen und dem Wissen der Teilnehmenden zeigen (Hypothese 1). Spezifischer wird erwartet, dass sowohl die Selbstwirksamkeit bzw. die Zuversicht, NET effektiv anwenden zu können (Hypothese 1a), die Einstellung gegenüber evidenzbasierten Verfahren (Hypothese 1b) als auch das Wissen zu Diagnose, Symptomen und Therapie der PTBS (Hypothese 1c) durch das Training signifikant steigen und zum Follow-Up erhalten bleiben.

Zudem wird davon ausgegangen, dass die teilnehmenden Psychologinnen und Psychologen nach dem Training in der Lage sind, die grundsätzlichen Konzepte der NET umzusetzen, bzw. anzuwenden (Hypothese 2), so dass im Schnitt eine gute Umsetzung der entsprechenden Bereiche zu erwarten ist.

Im Hinblick auf die Zusammenhänge der verschiedenen Evaluationskomponenten werden die im IMTEE angenommenen Korrelationen postuliert (Abbildung 3). Für die in dieser Studie untersuchten Bereiche werden deshalb signifikante Korrelationen zwischen:

- a) den Einstellungen nach dem Training und dem Wissen über Diagnose, Symptome und Therapie von PTBS;
- b) der Selbstwirksamkeit und dem Verhalten in der Praxis;
- c) den Einstellungen gegenüber evidenzbasierten Therapien und dem Verhalten in der Praxis;
- d) dem Wissen über Diagnose, Symptome und Therapie von PTBS und dem Verhalten in der Praxis;
- e) den bisherigen Erfahrungen (=individuellen Eigenschaften) und den Reaktionen auf das Training;
- f) den bisherigen Erfahrungen und den Einstellungen nach dem Training;
- g) den bisherigen Erfahrungen und dem Wissen über Diagnose, Symptome und Therapie von PTBS, sowie
- h) den bisherigen Erfahrungen und dem Verhalten in der Praxis

erwartet (Hypothese 3).

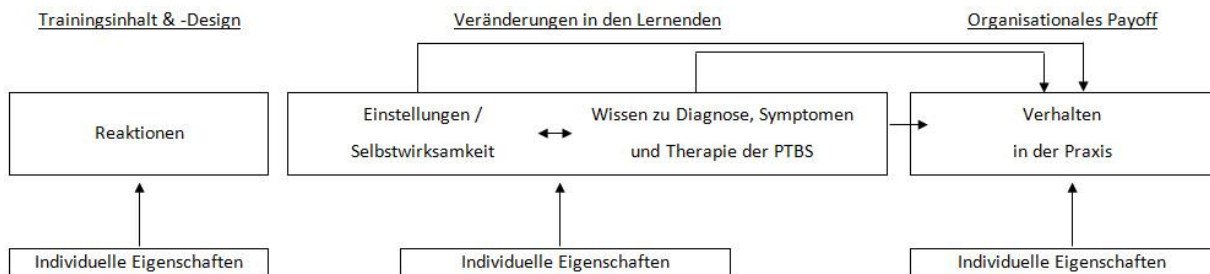


Abbildung 3. Abgewandeltes Modell der Trainingsevaluation und -effektivität basierend auf Alvarez et al. (2004). Die Pfeile deuten die angenommenen Korrelationen der einzelnen Komponenten an.

2 Methoden

2.1 Messinstrumente

Für die vorliegende Studie wurden sowohl bestehende Instrumente verwendet als auch neue Instrumente konstruiert. Diese werden im Folgenden näher beschrieben.

2.1.1 Fragebogen Reaktion-Einstellung-Wissen

Um die Ebenen 1 (Reaktion zum Training) und 2 (Einstellung und Wissen) des Therapist training evaluation outcomes framework (TTEOF) vor und nach dem Training evaluieren zu können, wurden zwei Fragebögen (siehe Anhang A1 und A2) mit verschiedenen Elementen erstellt. Die beiden Versionen unterscheiden sich dadurch, dass nur in der Version für die Messung vor dem Training (REW-pre) die soziodemographischen Daten erhoben wurden und nur in der Version für die Erhebung nach dem Training (REW-post) die Reaktion zum Training erfasst wurden. Alle anderen Bereiche (Einleitung, Einstellung und Wissen) werden sowohl vor als auch nach dem Training erhoben.

Die Einleitung des Fragebogens informiert über das Ziel der Evaluation und wie die Analyse der Daten erfolgt. Um die Tendenz, sozial erwünscht zu antworten, zu verringern, wurde wie in der Literatur empfohlen (Jonkisz, Moosbrugger & Brandt, 2012, S. 59) versichert, dass es weder richtige noch falsche Antworten gibt und die Daten anonym und vertraulich behandelt werden. Zudem wurde unterstrichen, dass die Teilnahme freiwillig ist und jederzeit beendet werden kann. Zuletzt wurden Kontaktinformationen für eventuelle Fragen oder Kommentare zur Verfügung gestellt.

Zu Beginn des Hauptteils des Fragebogens wurden die Teilnehmenden gebeten, einen Kode (Pseudonym, Metschke & Wellbrock, 2002, S. 23) zur späteren anonymen Verknüpfung der Daten anzugeben, bestehend aus den ersten beiden Buchstaben des Vornamens des Vaters, den ersten beiden Buchstaben des Vornamens der Mutter und dem eigenen Geburtstag (ebenfalls zweistellig).

In der Version REW-post folgt anschließend die Erhebung der Reaktion auf das Training. Für eine umfassendere Erhebung dieses Konstrukts wurde dazu mit einem Item erfasst, ob man das Training einem Kollegen oder einer Kollegin empfehlen würde (übernommen von Morgenstern et al., 2001) und in Anlehnung an Seng et al. (2006) die

Zufriedenheit mit den Inhalten sowie die Gesamtbeurteilung des Trainings erhoben. Anders als bei Seng et al. (2006) wurde hierbei allerdings eine fünf- und keine siebenstufige Likert-Skala verwendet um die Abstufungen innerhalb des Fragebogens konstant zu halten. Auf eine umfassende Skala wie sie beispielsweise Herschell et al. (2009) eingesetzten, wurde hier bewusst verzichtet, da die Reaktionsmaße bisher wenig Zusammenhang mit anderen Maßen gezeigt haben (Alliger et al., 1997) und daher an dieser Stelle die Länge des Fragebogens minimiert werden konnte (vgl. Kirkpatrick, 1998, S. 26) .

Anschließend folgt ein Abschnitt zu den Einstellungen der Teilnehmenden. Dazu wurde im Sinne der Selbstwirksamkeit sowohl erfragt wie zuversichtlich die Psychologinnen und Psychologen sind, die Narrative Expositionstherapie effektiv einsetzen zu können (ein Item, vgl. Morgenstern et al., 2001) als auch die generelle Einstellung gegenüber evidenzbasierten Verfahren. Die hierfür verwendete *Evidence-Based Practice Attitude Scale* (EBPAS) von Aarons (2004) besteht aus insgesamt 15 Items die nach den Faktorenanalysen des Autors zu den vier Skalen *Anordnung* (Requirement, $\alpha = .90$), *Anreiz* (Appeal, $\alpha = .80$), *Offenheit* (Openness, $\alpha = .78$) und *Abweichung* (Divergence, $\alpha = .59$) zusammengefasst werden können. Die Skala Anordnung umfasst dabei die Einstellung, dass der Therapeut oder die Therapeutin eine neue Therapie oder Intervention einsetzen würde, wenn dies von einer Organisation, einem Vorgesetzten oder dem Staat gefordert werden würde. Anreiz bedeutet, inwiefern die neue Intervention intuitiv ansprechend wirkt, Sinn zu ergeben scheint, man das Gefühl hat, gut genug ausgebildet zu sein, um sie korrekt anwenden zu können sowie ob sie von Kollegen oder Kolleginnen eingesetzt wird, die glücklich mit der Intervention sind. Auf der Skala Offenheit erzielen diejenigen hohen Werte, die generell offen sind, neue Arten von Interventionen und Therapien anzuwenden. Die Skala Abweichung wird auch beschrieben als „perceived Divergence of EBP⁴ with usual practice“ (Aarons, 2004, S. 64) und umfasst die Sichtweise, dass evidenzbasierte Verfahren nicht klinisch nützlich und weniger wichtig als klinische Erfahrung sind. Hierbei ist zu beachten, dass im Gegensatz zu den anderen drei Skalen ein hoher Wert auf der Skala Abweichung bedeutet, dass die entsprechende Person weniger wahrscheinlich evidenzbasierte Verfahren anwenden wird.

⁴ Evidence-based practice

Zusätzlich zur Auswertung der einzelnen Skalen kann ein Gesamtwert berechnet werden, indem zunächst alle Itemwerte der Skala Abweichung invertiert und anschließend die Werte aller Items summiert werden. Die interne Konsistenz der Gesamtskala liegt bei $\alpha = .77$ und damit in einem akzeptablen Bereich. Nach Aarons (2004) kann das Modell der durchgeführten explorativen Faktorenanalyse 63% der Varianz erklären.

Zur Evaluation des Wissens über Posttraumatische Belastungsstörung (PTBS) wurde mittels Literaturrecherche nach bereits existierenden Erhebungsinstrumenten gesucht und die Autoren zweier als geeignet scheinenden Studien angeschrieben. Da eine Antwort allerdings ausblieb, wurde ein neues Instrument erstellt. Ausgehend von der Überlegung, dass die zukünftigen Therapeutinnen und Therapeuten sowohl eine PTBS richtig diagnostizieren, spezifische Symptome erkennen und im Anschluss therapieren können müssen, wurden Items für diese drei Bereiche konstruiert. Als Grundlage hierfür dienten das Diagnostic and Statistical Manual of Mental Disorders in der vierten Auflage (American Psychiatric Association, 2000) sowie Vorlesungs- und Seminarunterlagen des Lehrstuhls für Klinische Psychologie der Universität Konstanz. Die Antworten zu den insgesamt elf Fragen wurden in einem geschlossenen Multiple-Choice-Format gegeben, wobei bei jeder Frage keine, eine oder mehrere Antworten richtig sein konnten („pick any of n“-Format, Jonkisz et al., 2012, S. 49). Es wurde versucht, vier Antwortmöglichkeiten pro Frage zu bieten. Insgesamt variierten die Antwortmöglichkeiten pro Frage jedoch zwischen drei und acht. Für die Auswertung der Wissensfragen wird für jedes richtig gesetzte und richtig nicht gesetzte Kreuz ein Punkt vergeben und für jedes falsch gesetzte bzw. falsch nicht gesetzte Kreuz ein Punkt abgezogen. Durch die Summation der Punkte errechnet sich der Wissensindex. Dieser bewegt sich zwischen minimal -48 für den Fall, dass alle Kreuze falsch gesetzt und maximal 48, wenn alle Kreuze richtig gesetzt wurden. Werden alle Antwortmöglichkeiten angekreuzt, so ist die Summe der Punkte -2, wird kein einziges Kreuz gesetzt ist die Summe 2.

In der Version REW-pre folgt anschließend die Abfrage der soziodemografischen Daten. Neben dem Alter in Jahren wurde auch das Geschlecht erfasst. Dabei wurde eine dreifache Antwortmöglichkeit geben. Neben „männlich“ und „weiblich“ konnten die Teilnehmenden „keine Angabe“ wählen. Da sich gezeigt hat, dass Erfahrung mit den Veränderungen in den Lernenden assoziiert ist (vgl. Alvarez et al., 2004) wurde zusätzlich in einem offenen Format erfragt, welche anderen Fort- oder Ausbildungen in anderen

Therapieformen die Teilnehmenden bereits absolviert hatten und seit wie vielen Jahren sie als Psychologe bzw. Psychologin gearbeitet haben. Des Weiteren wurde erfragt seit wie vielen Monaten sie ggf. schon für die Agencia Colombiana para la Reintegración arbeiten und für den Fall, dass sie bereits Erfahrung in der Arbeit mit Opfern und / oder Tätern des kolumbianischen Konflikts hatten, ebenfalls die Anzahl der Monate in dieser Arbeit und in einem offenen Format die Möglichkeit, ihre Arbeit näher zu konkretisieren. Zuletzt wurde der persönliche Bezug zu dem kolumbianischen Konflikt erfragt. Ebenfalls in einem offenen Format wurden die Teilnehmenden gebeten, dass wenn sie oder nahe Angehörige direkten Kontakt zu Guerillas oder Paramilitärs hatten, diesen kurz zu beschreiben.

Die Übersetzung und Rückübersetzung des Fragebogens erfolgte wie durch einschlägige Literatur empfohlen (z. B. Schmitt & Eid, 2007) durch unabhängige bilinguale Psychologinnen. Um die Passung der spanischen Version für eventuelle sprachliche oder kulturelle Besonderheiten in Kolumbien zu garantieren, wurde die Übersetzung von Deutsch nach Spanisch von einer kolumbianischen Psychologin übernommen. Bei Unstimmigkeiten wurden die entsprechenden Textstellen mit einer dritten, unabhängigen bilingualen Psychologin geklärt und die spanische Version dementsprechend angepasst.

2.1.2 Beurteilungsbogen Diagnostik-Lifeline-Exposition-NET

Der Beurteilungsbogen Diagnostik-Lifeline-Exposition-NET (DLE-NET) soll eine möglichst vollständige und gleichzeitig ökonomische Beurteilung der therapeutischen Fertigkeiten in der NET bieten. Als Basis diente der „Beurteilungsbogen für NET-Sitzungen“ von Ertl (2013). Dieser Beurteilungsbogen umfasst acht spezifische Items zu zentralen Aspekten in der Expositionsphase der NET sowie drei Items, die das Verständnis für die Welt des Patienten oder der Patientin, die Interaktionswirksamkeit sowie das Interesse und die Empathie der Therapierenden für die Patientinnen und Patienten erfragen. Jedes der elf Items ist gekennzeichnet durch eine Überschrift und einen kurzen beschreibenden Text. Die Bewertung erfolgt auf einer dreistufigen Skala von 0 = *trifft nicht zu*, über 1 = *trifft teilweise zu* und schließlich 2 = *trifft voll zu*. Die Auswertung erfolgt über eine einfache Summation der Itemwerte, wobei Ertl (2013) einen Gesamtwert von 11 als befriedigendes Ergebnis angibt.

Da der Beurteilungsbogen von Ertl (2013) zu Beginn dieser Studie noch nicht erprobt wurde und weder statistische Gütekriterien noch sonstige Erfahrungswerte mit dem Instrument bekannt waren, wurde eine erste Erprobung in Kolumbien vorgenommen. Die

Übersetzung des Beurteilungsbogens auf Spanisch erfolgte durch eine bilinguale Psychologin. Die Adäquatheit der Übersetzung wurde durch Rückübersetzung überprüft.

Auf Grundlage von insgesamt 12 Beobachtungen in den Sitzungen von neun verschiedenen Therapeut_innen in acht verschiedenen Städten in Kolumbien zeigte sich, dass der Beurteilungsbogen einige zentrale Bereiche der NET adäquat erfassen kann, andere jedoch nicht berücksichtigt. Als erstes ist zu nennen, dass der Beurteilungsbogen lediglich für die Expositionsphase konzipiert wurde, in der Evaluationspraxis aber ebenfalls Fertigkeiten in Diagnostik und Lifeline evaluiert werden sollten. Zudem fehlte beispielsweise ein Item zur angemessenen Auswahl des behandelten Lebensereignisses (handelt es sich um ein traumatisches Ereignis nach den gängigen Diagnosekriterien oder nicht). Der Beobachtungsbogen wurde um die entsprechenden Elemente ergänzt und in vier Bereiche (Diagnostik, Lifeline, Exposition und therapeutische Fähigkeiten) aufgeteilt. Es wird angenommen, dass die therapeutischen Fähigkeiten unabhängig von der jeweiligen Phase der Therapie beurteilt werden können und daher sollten diese stets zusätzlich zu dem jeweiligen Bereich der aktuellen Phase bewertet werden. Zudem wurde die Skala abgeändert und erweitert, da es im ersten Testlauf Missverständnisse mit dem Anker „trifft nicht zu“ gab. Teilweise wurde diese Option gewählt, wenn eine gewisse Verhaltensweise von Seiten des Therapeuten nicht durchgeführt wurde (z. B. Vermeidung vorzubeugen), teilweise aber auch um auszurücken, dass man keine Angabe machen kann, da beispielsweise gar keine Vermeidung von Seiten des Patienten eintrat. Die Skala wurden deshalb umgeändert in 0 = *schlecht*, 1 = *gut*, 2 = *sehr gut* und X = *keine Angabe*. Zudem gibt es die Möglichkeit, einen kurzen Kommentar zu verfassen, warum keine Angabe möglich war.

Der Beurteilungsbogen wurde wiederum auf Spanisch übersetzt und von einer unabhängigen Psychologin rückübersetzt. Da bei der Rückübersetzung keine Unstimmigkeiten auftraten, erfolgte keine weitere Korrektur der spanischen Version.

2.2 Durchführung

Aufgrund der Möglichkeiten vor Ort konnte kein zwei Gruppen-Plan mit einer Kontrollgruppe durchgeführt werden. Die Auswirkungen auf die interne Validität mussten in Kauf genommen werden und ganz nach der Devise von Cronbach (1982) wird Evaluation

als die Kunst des Möglichen gesehen. Um dennoch ein möglichst aussagekräftiges Design zu erhalten, wurde die Untersuchung als Eingruppen-Pretest-Posttest-Studie mit zusätzlicher Follow-Up-Erhebung geplant.

2.2.1 Training

Das Training wurde von Mitarbeitern und Mitarbeiterinnen der Arbeitsgruppe Klinische Psychologie der Universität Konstanz geplant und durchgeführt. Im Zeitraum vom 12. August bis zum 23. August 2013 wurden den 29 Teilnehmenden der ACR in Bogotá an zehn aufeinanderfolgenden Tagen theoretische und praktische Kenntnisse zu Trauma und Traumatherapie (NET) vermittelt. Dabei wurde spezifisch auf den theoretischen Hintergrund der PTBS und Diagnostik von PTBS und komorbiden Psychischen Störungen eingegangen sowie Demonstrationen der verschiedenen NET-Elemente (vgl. Schauer et al., 2011; Elbert et al., 2012) und praktische Gruppen-Übungen durchgeführt.

Im Oktober 2013 erfolgte eine erste Supervision aller Einsatzstellen von einer kolumbianischen Mitarbeiterin der Arbeitsgruppe in Konstanz bei der die Therapeuten und Therapeutinnen weitere praktische Erfahrungen sammeln konnten. Dabei arbeitete immer ein Therapeut mit einem Co-Therapeuten gemeinsam mit einem Patienten unter der Anwesenheit der Supervisorin. Anschließend wurden in einer weiteren Therapiesitzung die Rollen von Therapeutin und Co-Therapeutin getauscht. Nach Abschluss der Supervisions-sitzung erhielten die Teilnehmenden ein ausführliches Feedback über ihre Arbeit.

Bis zur zweiten Supervisionsrunde in den Einsatzstellen im März 2014 gab es alle zwei Wochen die Möglichkeit einer persönlichen Supervision über Skype, die auch in hohem Maße in Anspruch genommen wurde (C. Bueno, persönl. Mitteilung, 03.02.2014).

2.2.2 Datenerhebung

Die Daten für diese Arbeit wurden im Zeitraum von August 2013 bis März 2014 erhoben. Der Fragebogen REW wurde direkt vor (REW-pre) und nach (REW-post) dem Training in den Trainingsräumen in Bogotá von den Teilnehmenden ausgefüllt. Zudem gab es, wie beispielsweise bei Goldstein und Ford (2002) empfohlen, eine Follow-Up-Erhebung (REW-follow-up; identischer Fragebogen zu REW-post) im März 2014, sieben Monate nach dem Training. Zu allen Zeitpunkten wurde der Fragebogen als Papier und Bleistift Version eingesetzt. Die Übertragung der Werte von REW-pre und REW-post in das Programm SPSS

erfolgte per Hand von den Originalfragebögen. Unleserliche Passagen in den offenen Antwortformaten wurden durch Fragezeichen gekennzeichnet. Die Fragebögen der Follow-Up-Messung wurden von den Therapeuten und Therapeutinnen in ihren jeweiligen Einsatzorten in Kolumbien im März 2014 ausgefüllt und eingescannt. Die weitere Verarbeitung erfolgte wie oben beschrieben.

Der Beurteilungsbogen DLE-NET als Papier und Bleistift Version wurde von der Supervisorin des Programms in Kolumbien in den Supervisionssitzungen im März 2014 eingesetzt. Die Evaluation erfolgte während bzw. direkt nach der jeweiligen Sitzung. Um die Objektivität der Beurteilung zu erhöhen bzw. zu bestimmen wäre es wünschenswert, Beurteilungen von verschiedenen Ratern einzuholen und eine Interrater-Reliabilität zu berechnen. Da es allerdings aus Datenschutzgründen nicht möglich war, Aufzeichnungen der Sitzungen anzufertigen und der Ablauf der Therapie nicht durch eine weitere externe Person beeinflusst werden sollte, wurde darauf bewusst verzichtet. Die Übertragung der Beurteilungsbögen in SPSS erfolgte von Hand.

2.3 Stichprobenbeschreibung

An dem Training nahmen insgesamt $N = 29$ Psychologen und Psychologinnen im Alter zwischen 27 und 46 Jahren teil ($M = 33.82$, $SD = 5.44$), von denen $n = 25$ direkt therapeutisch mit den ehemaligen Mitgliedern von Guerilla und Paramilitärs (im Folgenden bezeichnet als Therapeut_innen) und $n = 4$ in der programmatischen Leitung in der ACR (im Folgenden bezeichnet als Koordinator_innen) arbeiten sollten. Da sich für die beiden Gruppen in der Mehrzahl der Variablen keine signifikanten Unterschiede ergaben, wurden die Ergebnisse zusammengefasst, um eine größere Stichprobe zu erhalten. Im Falle von signifikanten Unterschieden zwischen den Gruppen wird darauf explizit hingewiesen. Eine Übersicht über die wichtigsten demographischen Variablen getrennt nach Beschäftigungsstatus ist in Tabelle 2 abgebildet.

Die Auswahl der Therapeuten und Therapeutinnen erfolgte bereits Anfang 2013 indem sich Psychologen und Psychologinnen auf die neu geschaffenen Stellen für klinische Psychologen (Reintegrador III) bewerben konnten. Ein Großteil der Bewerber_innen setzte sich aus den Psycholog_innen zusammen, die bereits im Rahmen der psychosozialen Unterstützung (Reintegrador II) in der ACR gearbeitet hatten. Somit müssen diese

Teilnehmenden als besondere Auswahl gesehen werden und als nicht repräsentativ für die Gesamtheit der kolumbianischen Psychologen und Psychologinnen.

Tabelle 2

Stichprobenbeschreibung nach Beschäftigungsstatus

	Therapeut_innen		Koordinator_innen	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Alter in Jahren	34.00	5.77	32.75	3.59
Tätigkeitsjahre ^a	8.13	3.22	7.33	1.53
Tätigkeitsjahre ^b	2.67	2.61	5.75	0.66
Erfahrungen mit Opfern ^c	3.37	2.98	5.40	1.66

Anmerkungen. Es zeigten sich keine signifikanten Unterschiede zwischen den Gruppen.

M = Mittelwert; *SD* = Standardabweichung

^a als Psycholog_in

^b bei der ACR

^c in Jahren

Von den Teilnehmenden wählten bei der Frage nach ihrem Geschlecht fünf Personen (17.2%) die Option „männlich“, 17 Personen (58.5%) „weiblich“, sechs Personen „keine Angabe“ (20.7%) und eine Person (3.4%) wählte keine dieser drei möglichen Optionen. Im Schnitt waren die Teilnehmer und Teilnehmerinnen bereits seit $M = 8.04$ Jahren als Psychologen und Psychologinnen tätig ($SD = 3.07$), wobei die Werte zwischen minimal 3.5 und maximal 17 Jahren schwankten. 19 Personen (65.52%) gaben an, bereits mindestens eine Fort- oder Ausbildung in anderen Therapieformen erhalten zu haben. Eine Übersicht über die absoluten und relativen Häufigkeiten, mit denen die verschiedenen Therapieformen angegeben wurden, ist in Tabelle 3 dargestellt.

Von den teilnehmenden Psychologinnen und Psychologen gaben 23 (79.3%) an, bereits vor dem Training bei der ACR gearbeitet zu haben. Dabei schwankte die bisherige Arbeitszeit zwischen zwei und 108 Monaten ($M = 36.85$, $SD = 31.86$). Ein ähnliches Bild zeigte sich bei den bisherigen Erfahrungen der Teilnehmenden bezüglich ihrer Arbeit mit den Opfern und / oder Tätern des bewaffneten kolumbianischen Konflikts. Hierzu gaben 24 (82.7%) Personen an, zwischen 3 und 120 Monate Erfahrung zu haben ($M = 44.54$, $SD = 34.57$). Zwei weitere Personen machten zwar keine Angaben über die Zeit ihrer bisherigen Erfahrungen, beschrieben diese jedoch nichtsdestotrotz. Womit davon auszugehen ist, dass insgesamt 26

Personen (89.7%) über einschlägige Erfahrungen verfügen. Eine Auswertung der angegebenen Erfahrungen nach der Inhaltsanalyse nach Mayring (2003) ergab folgendes Bild. Die große Mehrzahl der Teilnehmenden ($n = 24$) gab an, bisher über die psychosoziale Arbeit mit den Demobilisierten mit den Betroffenen in Kontakt gewesen zu sein. Fünf Personen gaben an bereits klinisch-therapeutisch mit Opfern oder Tätern des Konflikts gearbeitet zu haben. Je eine Person berichtete, dass sie in der Verhinderung der Rekrutierung für illegale Gruppen gearbeitet bzw. an einem Buch über den bewaffneten Konflikt mitgewirkt hat.

Tabelle 3

Absolute und relative Häufigkeiten der angegebenen Fort- und Weiterbildungen

Therapieform	Absolute Häufigkeit	Relative Häufigkeit
Analytische Therapie	4	13.79
Gestalttherapie	2	6.90
(Kognitive) Verhaltenstherapie	9	31.03
Rational-Emotive Therapie	2	6.90
Systemische Therapie	3	10.34
Sonstige Therapie	7	24.14

Anmerkungen. Mehrfachnennungen möglich; „Sonstige Therapie“⁵ umfasst: Emotional Freedom Techniques, Cybernetic Transposition, Hypnotherapie, Neuro-Linguistisches Programmieren, Neuropsychologie für Kinder, Psychodynamik, Terapía Energetica Psicocorporal, Terapía Regresiva Reconstructiva

Insgesamt sieben Teilnehmende (24.1%) gaben an, dass entweder sie oder nahe Angehörige direkten Kontakt zu Guerrillas oder Paramilitärs gehabt zu haben. Bei der Auswertung mittels Inhaltsanalyse ergaben sich die vier Kategorien „Freund(e) als Mitglied illegaler einer Gruppe“, „selbst Mitglied einer Gruppe gewesen“, „in einer Konfliktzone gelebt“ und „Familienmitglied entführt bzw. ermordet worden“. Pro Kategorie gaben je zwei Teilnehmende an, diese spezifischen Erfahrungen gemacht zu haben.

Drei der Therapeuten und Therapeutinnen arbeiteten nur bis Ende 2013 in der ACR und die Koordinator_innen wurden nur zur Pre- und Post-Erhebung befragt. Für eine

⁵ Insofern keine passende Übersetzung gefunden wurde, wurde der spanische Originalwortlaut verwendet, um einer Verfälschung vorzubeugen.

Übersicht über den Fluss der Teilnehmenden siehe Abbildung 4. Obwohl zum Follow-Up-Zeitpunkt von allen verbleibenden Teilnehmenden Fragebögen als auch der Beobachtungsbogen DLE-NET vorliegt, konnte für vier Personen aufgrund fehlender oder unstimmgiger Codes keine Verknüpfung der beiden Datenquellen erfolgen.

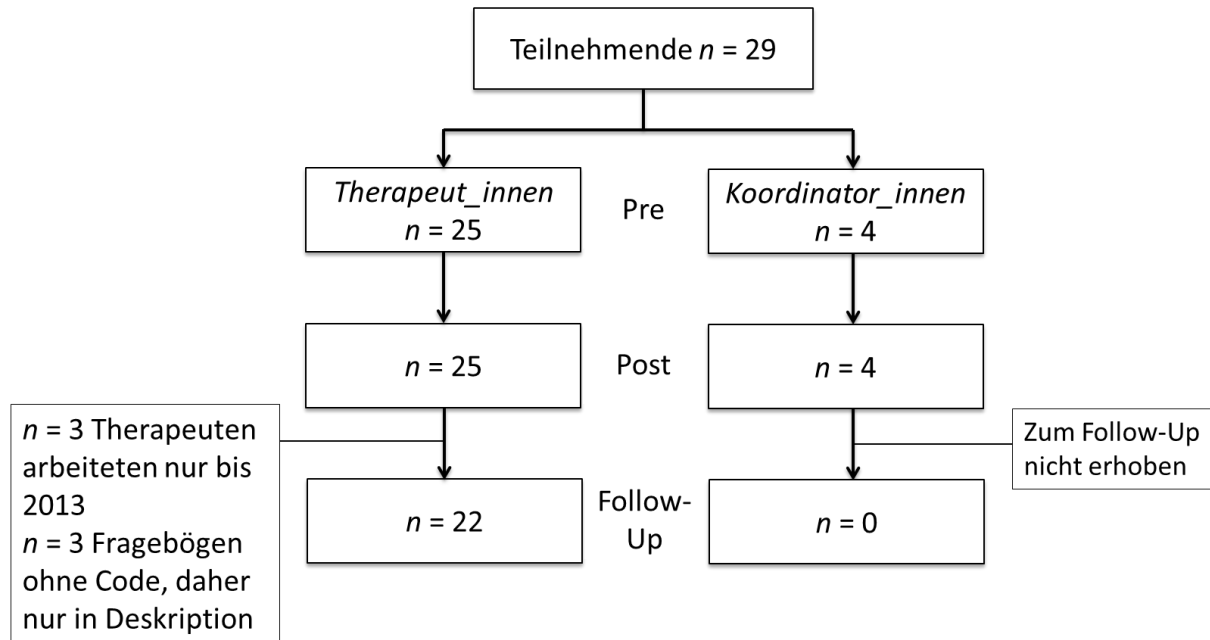


Abbildung 4. Flussdiagramm der Teilnehmenden über die verschiedenen Zeitpunkte der Evaluation nach Beschäftigungsstatus.

Beim Vergleich der Werte der drei Therapeut_innen, die 2014 nicht mehr bei der ACR arbeiteten und demnach nicht an der Follow-Up-Erhebung teilnahmen mit den Werten der übrigen Therapeut_innen zeigten sich einige signifikante Unterschiede. So waren die drei Therapeut_innen im Schnitt signifikant weniger zufrieden mit dem Inhalt des Trainings ($Mdn = 4.00$ vs. 5.00 ; $U = 10.00$, $z = 2.44$, $p = .02$), zeigten vor dem Training eine größere Ablehnung evidenzbasierter Verfahren ($Mdn = 2.75$ vs. 2.00 ; $U = 20.00$, $z = 2.37$, $p = .02$), waren aber gleichzeitig nach dem Training eher bereit, diese auf Anordnung anzuwenden ($Mdn = 5.00$ vs. 4.00 ; $U = 22.00$, $z = 2.44$, $p = .03$). Zudem erreichten sie vor dem Training durchschnittlich geringere Wissensscores ($Mdn = 23.00$ vs. 32.00 ; $U = 22.00$, $z = 2.26$, $p = .02$) und hatten insgesamt weniger Erfahrung in der Arbeit mit Opfern und/oder Tätern des kolumbianischen Konflikts ($Mdn = 4.00$ vs. 47.00 ; $U = 3.50$, $z = 2.35$, $p = .02$).

2.4 Statistische Auswertung

Die Auswertung der Daten erfolgte mit dem Programm IBM SPSS Statistics in der Version 21. Zur Überprüfung der Qualität des eingesetzten Frage- sowie Beurteilungsbogens wurde die Verteilung der Rohwerte, die Itemschwierigkeit (p_i), die Trennschärfe (r_{it}) sowie die Homogenität (Cronbachs α) untersucht.

Vor der Entscheidung, welche Verfahren zur statistischen Auswertung der Daten verwendet werden soll, muss überprüft werden, welche Verfahren überhaupt zulässig sind. Für die parametrischen Verfahren sind in der Regel normalverteilte statistische Kennwerte sowie Varianzhomogenität der zu vergleichenden Gruppen notwendig (Bortz, Lienert & Boehnke, 2008). Zur Überprüfung der Normalverteilungsannahme sollte dabei nicht nur auf einschlägige Tests (Kolmogoroff-Smirnov-Test, Bartlett-Test) zurückgegriffen werden, sondern zusätzlich eine optische Überprüfung der Verteilung anhand von Histogrammen mit Normalverteilungskurve vorgenommen werden. Für die Mehrzahl der zu überprüfenden Variablen ergaben sich sowohl signifikante Ergebnisse ($p < .05$) des Kolmogoroff-Smirnov-Tests mit Lilliefors-Korrektur als auch grobe Abweichungen in der optischen Überprüfung der Q-Q-Diagramme und Histogramme von der Normalverteilungsannahme. Daher wurde an dieser Stelle entschieden, wie von Bortz et al. (2008) empfohlen, wo möglich auf parametrische Verfahren zu verzichten und stattdessen verteilungsfreie Verfahren zu verwenden.

Zur Überprüfung, ob die Teilnehmenden zu den verschiedenen Zeitpunkten signifikant unterschiedliche Ergebnisse in den abhängigen Variablen erreichten, wurde *Friedman's ANOVA* (Field, 2009) berechnet. Dieser Test überprüft die Nullhypothese, dass die verschiedenen Stufen eines Faktors (hier Messzeitpunkte) aus Populationen mit gleicher zentraler Tendenz stammen (Bortz et al., 2008).

Analog zu der Rangvarianzanalyse von Friedman zur Überprüfung von Unterschieden hinsichtlich der zentralen Tendenz mehrerer verbundener Stichproben kann für mehrere unverbundene Stichproben der *H-Test* von Kruskal und Wallis berechnet werden. Dieser nimmt ebenfalls an, dass die Merkmalsverteilungen in den verschiedenen Populationen unter Gültigkeit der Nullhypothese formgleich sind und sich bei Gültigkeit der Alternativhypothese mindestens zwei Medianwerte unterscheiden (Bortz & Lienert, 2008).

Um neben dem generellen Effekt auch die spezifischen Gruppenunterschiede zu überprüfen, wurden im Anschluss im Falle von verbundenen Stichproben *Vorzeichenrangtests* nach Wilcoxon bzw. bei unverbundenen Stichproben *U-Tests* nach Mann-Whitney berechnet. Beide Tests prüfen, ob sich die Stichproben hinsichtlich ihrer zentralen Tendenz unterscheiden. Um eine Inflation des Fehlers erster Art zu vermeiden, muss nach Field (2009) hierbei allerdings eine Bonferroni-Korrektur durchgeführt werden, sodass als neues α -Niveau $\alpha^* = \alpha / (\text{Anzahl der Vergleiche})$ gilt. Bei drei Messzeitpunkten bedeutet dies bei einem gängigen Signifikanzniveau von $\alpha = .05$, dass das korrigierte Signifikanzniveau für die entsprechenden Tests bei $\alpha^* = .0167$ liegt.

Die Berechnung der Effektgrößen für den Wilcoxon-Vorzeichenrangtest, sowie den U-Test erfolgt nach Field (2009) mit der Formel $r = \frac{z}{\sqrt{N}}$, wobei N für die Anzahl der Beobachtungen steht, die bei unverbundenen Stichproben der Anzahl der Teilnehmenden entspricht, bei verbundenen Stichproben allerdings der Anzahl der Teilnehmenden multipliziert mit 2.

Zur Bestimmung von monotonen Zusammenhängen zweier mindestens ordinal-skalierter Merkmale kann nach Bortz et al. (2008) Kendalls τ_b berechnet werden. Bei einem perfekt monotonen Zusammenhang ergibt sich ein Wert von ± 1 für die Rangkorrelation. Falls eine der Variablen lediglich eine dichotome Ausprägung aufweist, wird die punktbiserial Korrelation berichtet (vgl. Bortz & Döring, 2006, S. 171).

3 Ergebnisse

Analog zu der Reihenfolge der Hypothesen finden für die verschiedenen Konstrukte des IMTEE zunächst stets eine statistische Beschreibung der Variablen und anschließend die statistische Überprüfung mit den unterstellten Veränderungen bzw. Zusammenhängen statt.

3.1 Deskriptive Statistik und Veränderungen

3.1.1 Reaktionen

Die Reaktionen auf das Training waren durchweg sehr positiv. Sowohl direkt nach dem Training als auch zur Follow-Up-Erhebung gaben alle Teilnehmenden an, dass sie das Training einem Kollegen weiterempfehlen würden. Die Zufriedenheit mit dem Inhalt als auch die Gesamtzufriedenheit mit dem Training bewegte sich ebenfalls zu den beiden Messzeitpunkten im Bereich zwischen *zufrieden* bis *sehr zufrieden* (vgl. Tabelle 4). Es zeigte sich dabei zudem eine Tendenz, dass zum Follow-Up eine noch höhere Zufriedenheit berichtet wurde, als direkt nach dem Training. Dieser Unterschied war allerdings nur für die Gesamtzufriedenheit signifikant ($T = 0, p = .05, r = .35$).

Tabelle 4

Deskriptive Statistik zu den drei Erhebungszeitpunkten

	Pre		Post		Follow-Up	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Reaktionen						
Zufriedenheit mit dem Inhalt			4.62	0.92	4.95	0.22
Gesamtzufriedenheit			4.54	0.86	4.95	0.22
Selbstwirksamkeit	3.81	1.09	4.32	0.69	4.43	0.60
EBPAS-Skalen						
Anordnung	3.86	1.10	4.11	1.00	3.85	0.97
Anreiz	3.96	0.62	4.15	0.69	4.14	0.73
Offenheit	4.37	0.60	4.57	0.51	4.61	0.45
Abweichung	2.16	0.70	2.09	0.72	2.13	0.76
Gesamt	3.99	0.45	4.21	0.44	4.14	0.40
Wissen	28.55	5.55	34.62	5.33	33.00	5.12

Anmerkungen. Niedrige Werte bei Skala Abweichung bedeuten hohe Zustimmung zu evidenzbasierten Verfahren. *M* = Mittelwert; *SD* = Standardabweichung; EBPAS = Evidence-Based Practice Attitude Scale

3.1.2 Einstellungen

Hinsichtlich der Einstellungen zeigte sich dass die Teilnehmenden im Mittel bereits vor dem Training *mittelmäßig* bis *sehr zuversichtlich* waren ($M = 3.81$), die NET effektiv anwenden zu können (vgl. Tabelle 4). Diese Zuversicht stieg im Laufe der Zeit weiter an ($M_{\text{post}} = 4.32$; $M_{\text{Follow-Up}} = 4.43$). Es zeigte sich allerdings keine signifikante Veränderung über die drei Zeitpunkte hinweg ($\chi^2(2, N = 14) = 4.06, p = .13$). Betrachtet man die Post-Hoc-Vergleiche zwischen den Werten der einzelnen Zeitpunkte für sich, so zeigte sich sowohl zwischen vor und direkt nach dem Training ($T = 10, p = .03, r = .33$) als auch zwischen vor dem Training und Follow-Up ($T = 9, p = .03, r = .39$) eine Veränderung. Da wie oben angesprochen allerdings eine Bonferroni-Korrektur durchgeführt wurde, kann in diesem Fall nicht von einer signifikanten Veränderung der Werte zu den einzelnen Zeitpunkten ausgegangen werden. Ebenso war die Veränderung zwischen direkt nach dem Training und Follow-Up nicht signifikant ($T = 7, r = .22$).

Zur Beurteilung der EBPAS-Skalen wurde zunächst im Rahmen der Itemanalyse die Schwierigkeit der einzelnen Items vor dem Training berechnet, da zu diesem Zeitpunkt die meisten Werte vorlagen. Die durchschnittliche Schwierigkeit aller Items betrug dabei $p_i = .60$ und reichte von $.46$ bis $.73$. Damit liegen alle Items in einem angemessenen mittleren Bereich (Bortz & Döring, 2006). Das schwierigste Item war dabei Item 9 (würde Intervention anwenden, wenn „sie intuitiv ansprechend wäre“), das leichteste Item 15 (würde Intervention anwenden, wenn „Sie das Gefühl hätten, genügend Training erhalten zu haben, um diese richtig anzuwenden“).

Wie die Itemschwierigkeit wurde auch die Trennschärfe lediglich für die Werte vor dem Training berechnet. Dabei erreichten die Items der Subskalen Trennschärfen zwischen $r_{it} = -.02$ (Item 6) bis $r_{it} = .82$ (Item 13). In der Gesamtskala variierten die Trennschärfen zwischen $r_{it} = -.22$ (Item 3) und $r_{it} = .58$ (Item 1). Auffallend ist dabei, dass drei der vier Items der Skala Abweichung, die mit invertierten Werten in die Gesamtskala eingehen, eine negative Trennschärfe aufweisen. Lediglich Item 5 ist positiv mit dem Gesamttestwert der Skala korreliert, wenn auch nur gering ($r_{it} = .16$). Nach Bortz und Döring (2006) sollten Items mit geringer Trennschärfe $r_{it} < .01$ unter Umständen aus dem Test entfernt werden. Aufgrund der geringen Stichprobengröße wurde die Skala allerdings wie von Aarons (2004) konzipiert eingesetzt.

Die Homogenität der EBPAS-Skalen wurde für alle drei Zeitpunkte mit Cronbachs α bestimmt. Wie in Tabelle 5 zu sehen ist erreicht lediglich die Skala Anordnung durchgehend akzeptable bis gute Werte in der internen Konsistenz. Am kritischsten sind die Werte der Skala Abweichung, die sich zu allen Zeitpunkten unter .5 und damit in einem inakzeptablen Bereich befanden.

Tabelle 5

Cronbachs α der EBPAS-Skalen zu den drei Erhebungszeitpunkten

Skala	Pre	Post	Follow-Up
Anordnung	.802	.716	.785
Anreiz	.384	.592	.766
Offenheit	.692	.868	.658
Abweichung	.353	.456	.414
EBPAS Gesamt	.549	.687	.643

Anmerkungen. EBPAS = Evidence-Based Practice Attitude Scale

Bezüglich der Einstellungen gegenüber evidenzbasierten Verfahren zeigten die Teilnehmenden im Mittel eine *große* bis *sehr große* Zustimmung (Tabelle 4). Dies gilt sowohl für die Gesamtauswertung als auch für die einzelnen Unterskalen. Dabei ist zu beachten, dass die Skala *Abweichung* eine kritische Sicht auf evidenzbasierte Verfahren repräsentiert und somit niedrigere Werte einer höheren Zustimmung entsprechen.

Bei den Einstellungen gegenüber evidenzbasierten Verfahren zeigten sich die einzigen signifikanten Gruppenunterschiede zwischen Therapeut_innen und Koordinator_innen. Wobei sich dies vor dem Training lediglich so formulierte, dass die Therapeut_innen ($Mdn = 2.50$) eher als Mitglieder der administrativen Ebene ($Mdn = 1.50$) die Einstellung vertraten, besser als akademische Forscher zu wissen, wie sie für ihre Klienten sorgen können (Item 3; $U = 18.00$, $z = 2.06$, $p = .04$, $r = .39$).

Direkt nach dem Training ergaben sich deutlich mehr signifikante Unterschiede zwischen den beiden Gruppen (vgl. Abbildung 5). Auf der Gesamtskala ergab sich für die Koordinator_innen im Schnitt ($Mdn = 3.90$) ein signifikant geringerer Wert als für die Therapeut_innen ($Mdn = 4.40$; $U = 17.00$, $z = 1.98$, $p = .05$, $r = .38$) und damit eine etwas geringere positive Einstellung gegenüber evidenzbasierten Verfahren. So waren die Koordinator_innen weniger bereit ($Mdn = 3.33$), eine evidenzbasierte Intervention auf Anordnung anzuwenden, als die Therapeuten und Therapeutinnen ($Mdn = 4.67$; $U = 15.00$,

$z = 2.18$, $p = .03$, $r = .42$). Ebenso war für die Koordinatoren und Koordinatorinnen ($Mdn = 3.25$) der Anreiz einer neuer Therapieformen nicht so zentral für die Entscheidung, diese einzusetzen, wie für die Therapeut_innen ($Mdn = 4.25$; $U = 10.00$, $z = 2.48$, $p = .01$, $r = .44$) und die Offenheit für neue Therapieformen und Interventionen der Koordinator_innen im Schnitt ($Mdn = 4.00$) signifikant geringer als die der Therapeut_innen ($Mdn = 5.00$; $U = 9.00$, $z = 2.48$, $p < .01$, $r = .51$). Andererseits war allerdings auch die abweichende Haltung gegenüber evidenzbasierten Interventionen unter den Koordinatoren und Koordinatorinnen ($Mdn = 1.25$) signifikant geringer als unter den Therapeuten und Therapeutinnen ($Mdn = 2.00$; $U = 15.00$, $z = 2.13$, $p = .03$, $r = .41$).

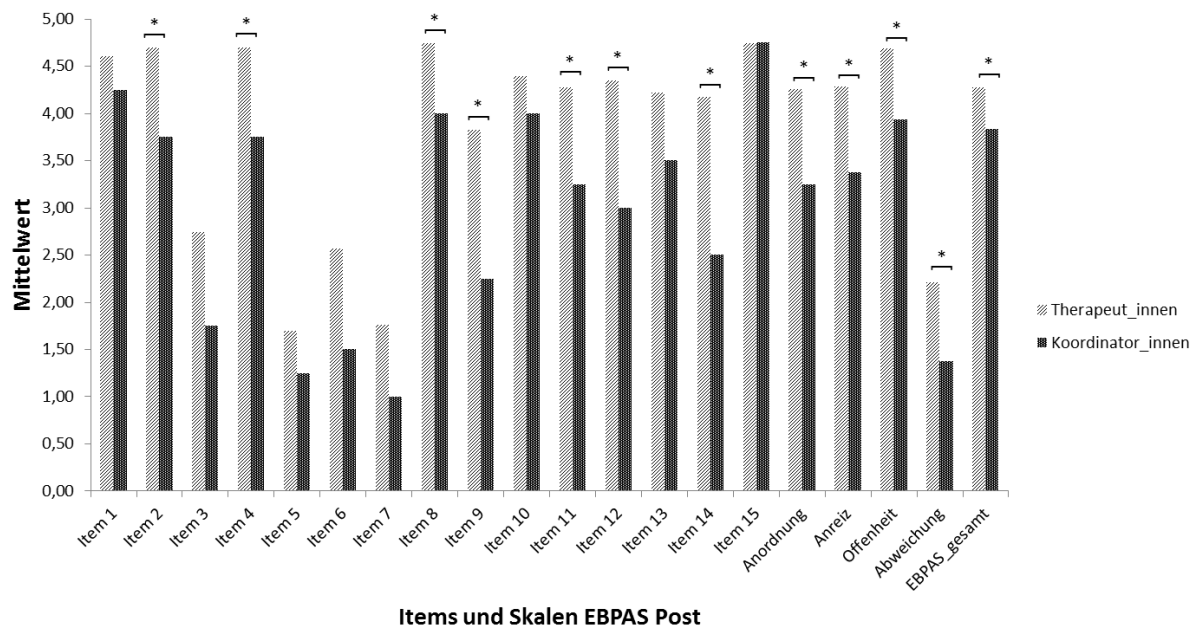


Abbildung 5. Vergleich EBPAS Items und Skalen zwischen Therapeut_innen und Koordinator_innen direkt nach dem Training. Mit Sternchen markiert sind signifikante Gruppenunterschiede mit $p < .05$.

Aufgrund der beschriebenen Gruppenunterschiede in den Werten der EBPAS-Skalen wurde die Hypothese, dass sich die Einstellungen gegenüber evidenzbasierten Verfahren zu den verschiedenen Zeitpunkten signifikant unterscheiden, getrennt für die beiden Gruppen untersucht. Dabei zeigten sich für die Koordinator_innen keinerlei signifikante Veränderungen von vor dem Training auf nach dem Training. Auch für die Therapeuten und Therapeutinnen ergaben sich keine signifikanten Ergebnisse über die drei Zeitpunkte hinweg. Allerdings gab es vereinzelte signifikante Post-Hoc-Vergleiche (Abbildung 6). Für die Gesamtskala zeigte sich ein signifikanter Unterschied von vor auf direkt nach dem Training,

so dass die Therapeuten und Therapeut_innen nach dem Training im Mittel ($Mdn = 4.40$) evidenzbasierten Verfahren positiver eingestellt waren als vor dem Training ($Mdn = 4.13$; $T = 55.00$, $z = 2.53$, $p < .01$, $r = .37$). Zudem zeigte sich, dass für die Therapeut_innen nach dem Training ($Mdn = 4.25$) die Wahrscheinlichkeit, ein evidenzbasiertes Verfahren bei großem Anreiz einzusetzen, signifikant höher war als noch vor dem Training ($Mdn = 4.00$; $T = 48.00$, $z = 2.38$, $p = .01$, $r = .35$). Zudem waren sie nach dem Training noch offener für neue Therapieformen ($Mdn = 5.00$) als zuvor ($Mdn = 4.50$; $T = 13.00$, $z = 2.53$, $p = .01$, $r = .37$).

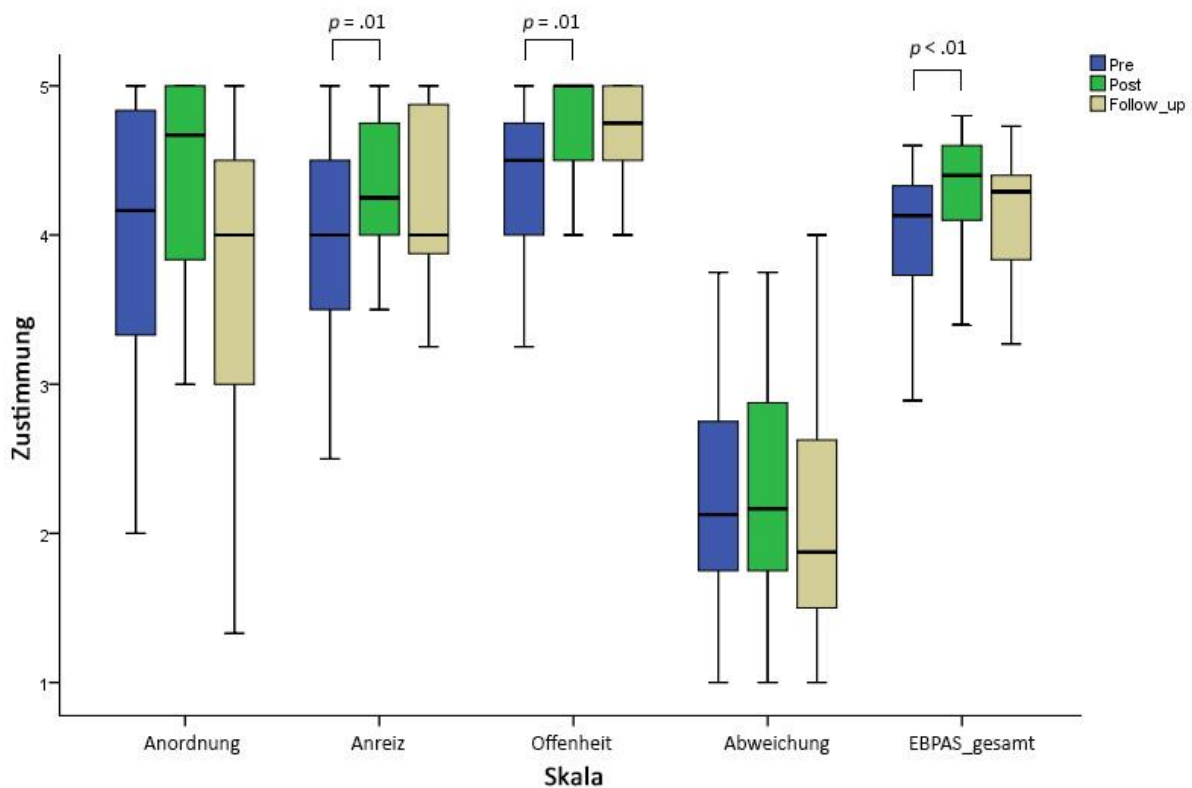


Abbildung 6. Boxplot EBPAS-Skalen der Therapeut_innen zu den drei Erhebungszeitpunkten mit p-Werten der signifikanten post-hoc berechneten Wilcoxon-Vorzeichenrangtests. Niedrige Werte bei Abweichung bedeuten hohe Zustimmung zu evidenzbasierten Verfahren

3.1.3 Wissen

Wie für die EBPAS-Skalen wurde auch für die Wissensfragen eine Itemanalyse für die Werte vor dem Training gerechnet. Bei der Itemschwierigkeit zeigte sich ein Durchschnitt von $p_i = .79$. Das schwierigste Item war mit $p_i = .55$ Item 9 „Wie viele Symptome des B-Kriteriums müssen vorhanden sein, um es als zutreffend festzustellen?“ und das leichteste Item mit $p_i = 1$ Item 5 „Wer kann eine PTBS entwickeln?“. Betrachtet man den Anteil der richtigen Antworten aufgeteilt nach den Antwortoptionen (Abbildung 7) ist ebenso festzustellen, dass

eine Vielzahl von Antwortoptionen durchgehend von einer großen Zahl von Teilnehmenden richtig beantwortet werden konnte, es allerdings auch einige Optionen gab, hinsichtlich derer sich auch durch das Training nur geringfügig höhere Lösungsraten ergeben haben.

Die Trennschärfe der Items variierte zwischen $r_{it} = -.26$ und $.53$ und die Homogenität der Gesamtskala Wissen betrug $\alpha = .359$. Folgt man Bortz und Döring (2006) und eliminiert schrittweise die Items mit negativer Trennschärfe, so erreichte das beste Modell eine interne Konsistenz von $\alpha = .61$, mit Trennschärfen von $r_{it} = .09$ bis $.58$. In diese Skala gingen lediglich die Items 1 und 4 sowie 6-9 ein. Da für die verbleibenden Items keine sinnvolle Kombination zu einer weiteren Sub-Skala zu finden war, wurde trotz der geringen Itemkennwerte für die Gesamtskala der Gesamtscore verwendet, um alle in den Items enthaltenen Informationen mit einzubeziehen. Bei der Interpretation der Ergebnisse muss allerdings besondere Vorsicht gelten.

Hinsichtlich des abgefragten Wissens zu Diagnose, Symptomen und Therapie der PTBS zeigten die Teilnehmenden bereits vor dem Training relativ hohe Werte (Tabelle 4 und Abbildung 8). So wurden im Schnitt 79.20 % der Aufgaben von den Teilnehmenden richtig gelöst, wobei sich der Anteil der richtig gelösten Aufgaben von 65.2 % bis 90.2 % erstreckte. Es zeigte sich eine signifikante Veränderung des Wissens zu den drei erhobenen Zeitpunkten ($\chi^2(2, N = 17) = 11.38, p < .01$), so dass nach dem Training mit 85.6 % richtig gelöster Aufgaben und einem mittleren Punktwert von $Mdn = 34.00$ signifikant mehr Punkte erreicht wurden also noch vor dem Training ($Mdn = 28.00; T = 45, p < .01, r = .46$). Die Spannweite der richtig gelösten Aufgaben variierte nach dem Training zwischen 74.1 % und 95.4 %. Der Unterschied zwischen vor dem Training ($Mdn = 28.00$) und der Follow-Up-Erhebung ($Mdn = 32.00$) wurde mit $T = 35, p = .05, r = .32$ nicht auf dem geforderten Signifikanzniveau von $.0167$ signifikant. Ebenso zeigte sich keine signifikante Veränderung von direkt nach dem Training zum Follow-Up-Zeitpunkt ($T = 36, p = .17, r = -.23$).

Weder vor ($U = 49.50, z = 0.03, p = .96, r = .01$) noch direkt nach dem Training ($U = 20.00, z = 1.18, p = .24, r = .26$) zeigten sich signifikante Gruppenunterschiede zwischen Therapeut_innen und Koordinator_innen.

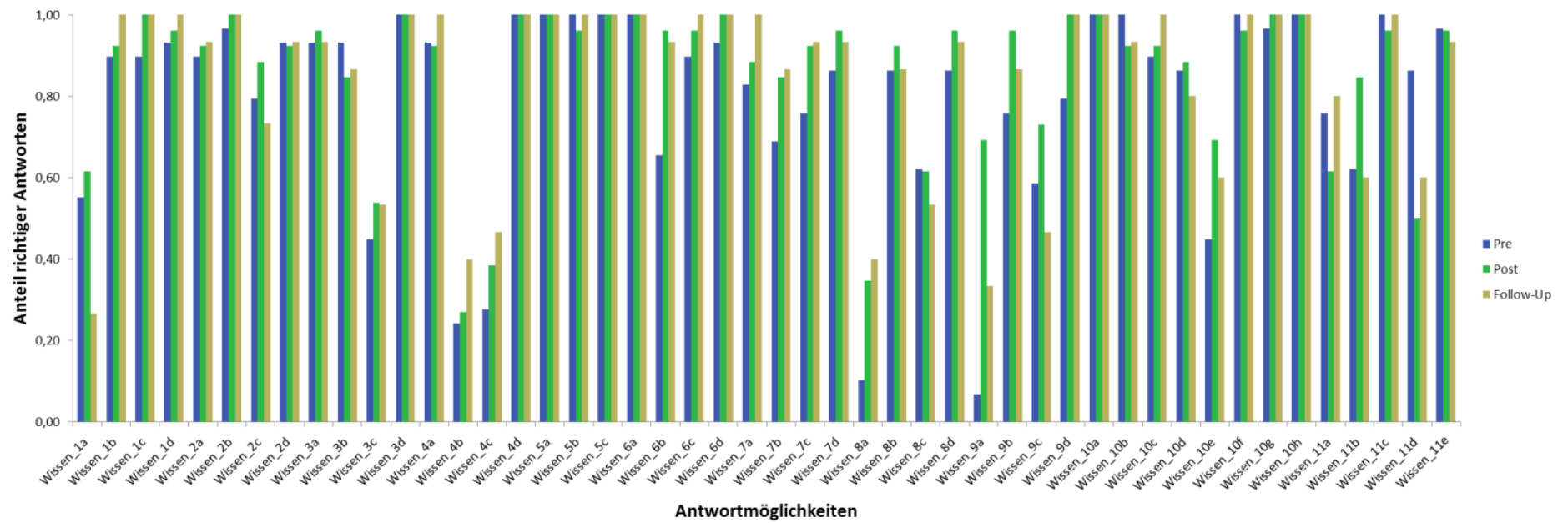


Abbildung 7. Anteil richtiger Antworten nach Antwortmöglichkeiten zu den drei Messzeitpunkten.

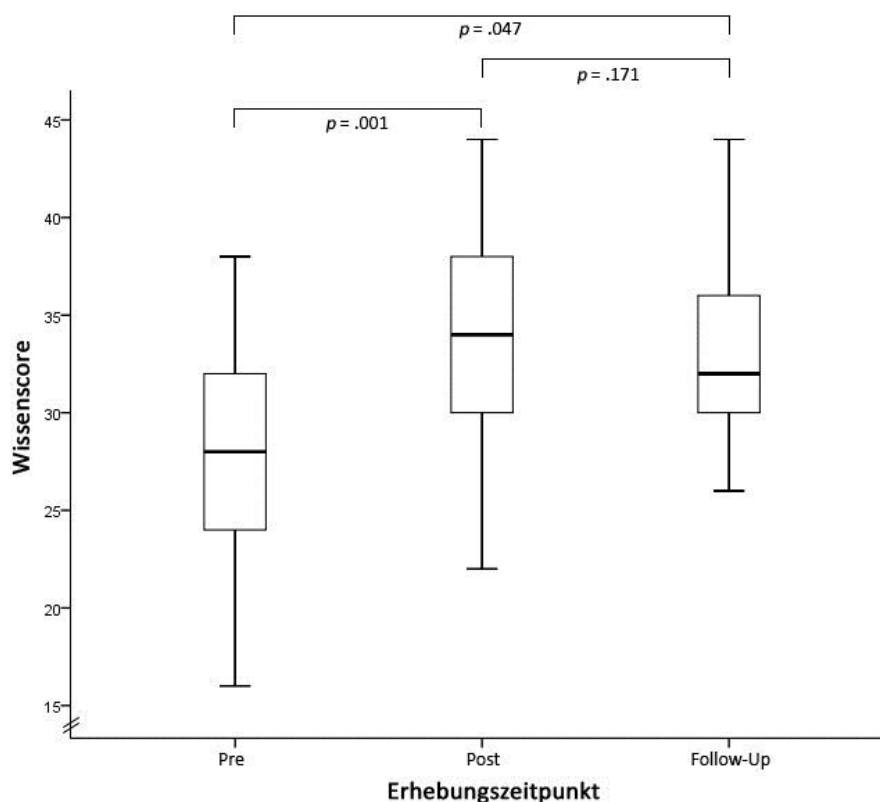


Abbildung 8. Boxplot der Wissenscores zu den drei Messzeitpunkten mit p-Werten der post-hoc berechneten Wilcoxon Vorzeichenrangtests. Mögliche Werte variieren zwischen ± 48 .

Im Hinblick auf die prozentual richtig gelösten Items des Wissenstests zeigten sich keinerlei signifikante Unterschiede über die drei Messzeitpunkte hinweg. Bei den Post-Hoc-Vergleichen der einzelnen Zeitpunkte zeigte sich lediglich für das Item 9 („Wie viele Symptome des B-Kriteriums müssen vorhanden sein, um es als zutreffend festzustellen?“) eine signifikante Veränderung von vor dem Training ($Mdn = .50$) auf direkt nach dem Training ($Mdn = 1.00$; $T = 18.00$, $z = 3.62$, $p < .01$, $r = .51$).

3.1.4 Verhalten in der Praxis

Das tatsächlich gezeigte Verhalten der Therapeutinnen und Therapeuten in den Therapiesitzungen mit den Klientinnen und Klienten wurde mit Hilfe des Beurteilungsbogens DLE-NET (Anhang A3) erfasst. Dabei zeigten sich für die Skalen Diagnostik ($\alpha = .923$; $.608 < r_{it} < .929$) und Exposition ($\alpha = .954$; $.698 < r_{it} < .910$) hohe Trennschärfen und exzellente interne Konsistenzen. Für die Skala Lifeline lag lediglich ein Datensatz vor, weshalb keine Kennwerte berechnet werden konnten.

Die gemittelten Fähigkeiten der Therapeuten und Therapeutinnen in den von den einzelnen Items erfassten Bereichen sind in Tabelle 6 dargestellt. Auffallend ist dabei, dass

die Fähigkeiten in der Diagnostik am niedrigsten ausfielen, wobei die „Kenntnis der Diagnoseinstrumente“ (Item 1) mit $M = 0.58$ ($SD = 0.67$) noch etwas geringer sind als der als „Diagnostisches Interview“ (Item 2) beschriebene Interviewstil ($M = 0.83$, $SD = 0.76$). In der Expositionsphase zeigten sich besondere Schwierigkeiten bei dem „Innervieren des emotionalen Netzwerkes („heißes“ Gedächtnis)“ (Item 10; $M = 0.78$) sowie der Verknüpfung von „kaltem“ und „heißem“ Gedächtnis (Item 11; $M = 0.78$). Die Erfragung der „Kontextuelle[n] Einbettung („kaltes“ Gedächtnis)“ (Item 9; $M = 1.22$) allein gehörte allerdings gemeinsam mit der Beendigung der Sitzung zu einem relativ sicheren Punkt (Item 14; $M = 1.22$) zu den Bereichen, die von den Therapeut_innen am besten beherrscht wurden. Insgesamt zeigte sich bei keinem der 18 Items eine signifikante Abweichung des Mittelwerts von 1, womit alle erfassten Bereiche als *gut* umgesetzt betrachtet werden können.

Tabelle 6

Mittelwerte und Standardabweichungen der DLE-NET Items der einzelnen Phasen

Item	Diagnostik ^a		Lifeline ^b		Exposition ^c		Gesamt	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	0.58	0.67						
2	0.83	0.72						
3			0.00	0.00				
4			1.00	0.00				
5			1.00	0.00				
6			1.00	0.00				
7					0.86	0.69		
8					1.13	0.35		
9					1.22	0.67		
10					0.78	0.67		
11					0.78	0.67		
12					0.89	0.60		
13					0.89	0.93		
14					1.22	0.44		
15	1.00	0.74	1.00	0.00	1.33	0.50	1.14	0.64
16	0.83	0.84	1.00	0.00	1.22	0.67	1.00	0.76
17	1.17	0.58	1.00	0.00	1.22	0.44	1.18	0.50
18	1.25	0.45	1.00	0.00	1.22	0.67	1.23	0.53

Anmerkungen. Itemwortlaut der phasenspezifischen Items siehe DLE-NET (Anhang A3).

Skalenanker: 0 = schlecht, 1 = gut, 2 = sehr gut. DLE = Diagnostik-Lifeline-Exposition;

M = Mittelwert; *SD* = Standardabweichung.

^a $n = 12$; ^b $n = 1$; ^c $n = 9$

Bei der Betrachtung des Mittelwertes der einzelnen Phasen zeigte sich sowohl für die Sitzungen der Diagnostik- ($M = 0.94$, $SD = 0.57$) als auch der Expositionsphase ($M = 1.06$, $SD = 0.49$) eine *gute* Umsetzung der erfassten Inhalte, sodass weder für Diagnostik ($t(11) = -0.34$, $p = .74$) noch für Exposition ($t(8) = 0.37$, $p = .72$) die Mittelwerte signifikant von 1 abwichen. Da für die Lifelinephase ($M = 0.87$) nur eine Beobachtung vorlag, konnte kein Signifikanztest durchgeführt werden.

3.2 Zusammenhangsmaße

In Tabelle 7 sind die Korrelationen aller für Hypothese 3a und 3e-g relevanten Variablen abgetragen, auf die im Folgenden der Reihe nach näher eingegangen werden soll. In Bezug auf Hypothese 3a zeigte sich kein Zusammenhang zwischen den Einstellungen nach dem Training und dem Wissensscore. Dies gilt sowohl für die Erhebung direkt nach dem Training als auch nach sieben Monaten.

Ebenso ergaben sich nur wenige signifikanten Korrelationen zwischen dem Verhalten in der Praxis und den anderen Variablen (Hypothese 3b: Selbstwirksamkeit; 3c: Einstellungen zu evidenzbasierten Verfahren; 3d: Wissen zu PTBS und 3h: individuelle Eigenschaften), weshalb diese Werte nicht in Tabelle 7 zu finden sind. Aufgrund der geringen Fallzahlen und zur explorativen Untersuchung wurde deshalb zudem überprüft, welche Korrelationen mit $p < .10$ marginal signifikant wurden. Für die Therapeut_innen, die in der Diagnostikphase ($n = 9$) beurteilt wurden, ergab sich dabei sowohl für den DLE-Gesamtwert ($\tau_b = .57$, $p = .06$) als auch für den Bereich der therapeutischen Fähigkeiten eine positive Korrelation mit den Tätigkeitsjahren als Psycholog_in ($\tau_b = .57$, $p = .07$). Dies weist darauf hin, dass diejenigen Personen, die bereits länger als Psychologen gearbeitet haben, tendenziell auch bessere therapeutische Fähigkeiten zeigen. Zudem zeigte sich, dass je höher der Zuwachs im Wissensscore von vor auf direkt nach dem Training war, desto höhere Werte wurden von den Therapeut_innen im Bereich der Diagnostik erreicht ($\tau_b = .64$, $p = .06$). Betrachtet man lediglich die Therapeuten und Therapeutinnen in der Expositionsphase ($n = 9$), ergaben sich ebenfalls Zusammenhänge mit dem Wissensscore. So zeigte sich sowohl vor dem Training als auch zum Follow-Up ein positiver Zusammenhang mit dem DLE-Gesamtwert ($\tau_b = .57$, $p = .04$ bzw. $\tau_b = .65$, $p = .05$). Dabei war der Wissensscore vor dem Training allerdings eher mit den therapeutischen Fähigkeiten verknüpft ($\tau_b = .73$, $p = .02$) und zum Follow-Up mit den Werten aus dem Bereich der Exposition des DLE ($\tau_b = .65$, $p = .05$), nicht aber mit den

therapeutischen Fähigkeiten. Da nur eine Person in der Lifelinephase beurteilt wurde, konnten hierfür keine Korrelationen berechnet werden.

Im Bezug der bisherigen Erfahrungen und Personenvariablen der Teilnehmenden mit den untersuchten Konstrukten zeigten sich einige signifikante Zusammenhänge, so dass sich die Hypothesen 3e-g zum Teil bewähren konnten. So war das Wissen direkt nach dem Training positiv mit dem angegebenen Geschlecht verbunden, so dass Frauen signifikant höhere Punktzahlen erreichten als Männer ($r_{pb} = .53, p = .02$). Vor dem Training und zum Follow-Up kehrte sich dieser Zusammenhang tendenziell um, erreichte aber keine Signifikanz. Sowohl direkt nach dem Training ($r_{pb} = .49, p = .03$) als auch zum Follow-Up ($r_{pb} = .73, p < .01$) ergaben sich zudem signifikante Korrelationen zwischen dem Geschlecht und der Bereitschaft, ein evidenzbasiertes Verfahren auf Anordnung anzuwenden. Dabei waren es wiederum die Frauen, die höhere Werte zeigten als die Männer. Zudem ergaben sich zur Follow-Up-Erhebung weitere signifikant positive Korrelationen für das Geschlecht mit der Skala Anreiz ($r_{pb} = .81, p < .01$) und dem EBPAS-Gesamtwert ($r_{pb} = .73, p < .01$), wobei weiterhin die Angabe des weiblichen Geschlechts mit höheren Skalenwerten assoziiert war.

Des Weiteren zeigte sich vor dem Training eine hoch signifikante Korrelation zwischen den Tätigkeitsjahren als Psychologe oder Psychologin und der Zuversicht, NET effektiv anwenden zu können ($\tau_b = .46, p = .01$). Dieser Zusammenhang verschwand allerdings nach dem Training nahezu vollkommen ($\tau_b = .08, p = .68$ bzw. $\tau_b = .15, p = .50$).

Im Hinblick auf die Erfahrungen der Teilnehmenden in der Arbeit mit Opfern bzw. Tätern des innerkolumbianischen Konfliktes zeigte sich, dass diejenigen mit längerer Erfahrung direkt nach dem Training tendenziell eine höhere Gesamtzufriedenheit mit dem Training angaben ($\tau_b = .39, p = .04$). Zudem zeigte sich eine signifikant positive Korrelation zwischen der Länge der Erfahrungen und der Selbstwirksamkeit zum Follow-Up ($\tau_b = .49, p = .03$), nicht aber zu den anderen beiden Erhebungszeitpunkten.

Tabelle 7

Korrelationstabelle der einzelnen Evaluationskomponenten

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1. Inhaltszufriedenheit	—																							
2. Gesamtzufriedenheit	.896**	—																						
Selbstwirksamkeit																								
3. Pre	.494*	.592**	—																					
4. Post	.314	.406*	.579**	—																				
5. Follow-Up	.101	.431	.359	.230	—																			
Wissensscore																								
6. Pre	.229	.123	-.030	-.261	.088	—																		
7. Post	.243	.124	-.105	-.089	.125	.211	—																	
8. Follow-Up	.224	.236	.138	-.044	.181	.361*	-.076	—																
EBPAS Post																								
9. Anordnung	-.092	.068	.263	.147	.593**	-.280	.187	.075	—															
10. Anreiz	.136	.216	.257	.249	.451*	.075	.223	.198	.569**	—														
11. Offenheit	.192	.391*	.359*	.337	.493*	.020	.145	-.019	.438**	.428**	—													
12. Abweichung	-.024	-.018	-.067	.172	.024	-.226	.245	-.187	.152	.192	.209	—												
13. EBPAS-Gesamt	.121	.207	.346*	.149	.580**	.045	.170	.055	.661**	.636**	.506**	-.079	—											
EBPAS Follow-Up																								
14. Anordnung	-.141	.087	-.067	.000	.418*	-.296	.363	.076	.500*	.104	.355	.256	.310	—										
15. Anreiz	.116	.308	-.064	.326	.377	-.408*	.414*	-.088	.342	.109	.415*	.382*	.080	.504**	—									
16. Offenheit	-.118	.191	-.132	.266	.279	-.138	.063	.138	.474*	.255	.506*	-.104	.449*	.325	.315	—								
17. Abweichung	.490*	.234	-.130	-.043	.046	-.185	.318	.158	.316	.139	.201	.403*	.116	.304	.154	.020	—							
18. EBPAS-Gesamt	-.244	.189	-.128	.269	.414*	-.169	.248	-.047	.336	.137	.436*	.128	.275	.514**	.601**	.605**	-.157	—						
19. Alter in Jahren	.064	.050	.220	.150	-.042	-.165	-.274	.014	.098	-.007	.138	-.040	.093	.170	-.137	-.123	.021	-.116	—					
20. Geschlecht ^a	.344	.335	-.073	.174	.408	-.302	.529*	-.257	.485*	.318	.235	.327	.310	.733**	.809**	.177	.249	.727**	.039	—				
21. Ausbildungsanzahl	-.267	-.139	-.025	.195	-.011	-.009	-.016	.193	.007	.035	-.154	-.102	-.020	-.055	-.045	.257	-.241	.180	-.003	-.013	—			
22. Tätigkeitsjahre ^b	.272	.076	.458**	.075	.146	.022	-.137	-.050	.074	.004	.239	.000	.183	.000	-.168	-.374	-.021	-.235	.559**	-.024	-.306	—		
23. Tätigkeitsmonate ^c	.288	.143	.124	.027	.076	.075	-.117	.223	-.136	-.188	-.122	-.153	-.124	-.180	-.169	-.156	.190	-.297	.102	-.106	-.373*	.392*	—	
24. Erfahrungen ^d	.418	.389*	.241	.129	.490*	.109	.015	.318	-.019	-.074	.098	-.037	.045	.355	.164	.243	.196	.200	.294	-.067	-.021	.289	.306	—

Anmerkungen. Inhalts- und Gesamtzufriedenheit nur für Post-Zeitpunkt, da zum Follow-Up bis auf einen Fall konstant und sich daher für die meisten Fälle keine Korrelation berechnen lässt. Signifikante Korrelationen sind durch Fettdruck hervorgehoben; die für die Hypothesen relevanten durch hellblaue Hinterlegung. EBPAS = Evidence-Based Practice Attitude Scale. $13 < n < 27$.

^a Kodierung: 0 = männlich, 1 = weiblich

^b als Psychologe

^c bei der ACR

^d mit Tätern und/oder Opfern in Monaten

* $p < .05$

** $p < .01$

4 Diskussion

Zur Evaluation des Trainings in Narrativer Expositionstherapie wurden die Reaktionen auf das Training sowie die Veränderungen in Wissen und Einstellungen durch das Training und das Verhalten in der Praxis untersucht. Dabei zeigte sich eine hohe Zufriedenheit mit dem Training, signifikante Veränderungen in Wissen und Einstellungen sowie eine gute Umsetzung der gelehrt Konzepte in der Praxis. Der Aufbau der Diskussion folgt zunächst den postulierten Hypothesen, um schließlich über weitere Erkenntnisse und Limitationen zu den Schlussfolgerungen und dem Ausblick zu kommen.

4.1 Einstellungen

In Bezug auf die Selbstwirksamkeit zeigten sich weit weniger signifikante Veränderungen durch das Training als erwartet. Daher die Hypothese 1a, dass sich die Selbstwirksamkeit über die drei Zeitpunkte hinweg verändert, nicht als bewährt angesehen werden. Da der p -Wert und damit der Bewährungsgrad einer Hypothese in erster Linie von dem Stichprobenumfang abhängig ist, sollte darüber hinaus stets die entsprechende Effektgröße berücksichtigt werden (Deutsche Gesellschaft für Psychologie, 2007). So zeigte sich sowohl für die Veränderung der Zuversicht, NET effektiv anwenden zu können von vor auf nach dem Training, als auch von vor dem Training zum Follow-Up-Zeitpunkt nach sieben Monaten eine mittlere Effektgröße, wobei die mittlere Zuversicht tendenziell immer weiter anstieg. Der Zuversicht-Zuwachs von direkt nach dem Training zur Follow-Up-Erhebung entspricht einer geringen Effektgröße. Dies legt den Schluss nahe, dass das Verfehlen des geforderten Signifikanzniveaus auf den geringen Stichprobenumfang zurückgeführt werden kann, es aber substantielle Veränderungen in der Zuversicht gab, die vor allem mit dem Training, aber auch mit der Zeit der danach folgenden Praxis in Zusammenhang zu bringen sind. Überdies ist zu beachten, dass die Zuversicht durchweg hoch ausgefallen ist. So bewegt sich diese im Mittel sowohl direkt nach dem Training als auch zum Follow-Up zwischen den beiden obersten Skalenwerten *sehr* und *sehr stark zuversichtlich*. Eine weitere Steigerung könnte gegebenenfalls über die fortlaufende Supervision und Feedback zu dem gezeigten Verhalten möglich sein. Im Hinblick auf die Itemformulierung könnte zudem überprüft werden, ob die potenzielle oder die aktuelle Selbstwirksamkeit erfasst wird. Zusammenfassend sprechen die gefundenen Ergebnisse aber durchweg für das durchgeführte Training.

Die Hypothese 1b, dass die Einstellungen gegenüber evidenzbasierten Verfahren sich über die drei Zeitpunkte hinweg verändern, konnte sich zum Teil bewähren. Zwar erreichte keine der gerechneten Friedman-Rangvarianzanalysen ein statistisch signifikantes Niveau, doch gab es für die Therapeut_innen bei den Skalen Anreiz, Offenheit sowie EBPAS-gesamt signifikante Steigerungen der Werte von vor dem Training auf direkt nach dem Training. Somit ist davon auszugehen, dass zumindest für diejenigen Teilnehmenden, die als Therapeut_innen arbeiten sollten, das Training einen positiven Einfluss auf die Einstellung gegenüber evidenzbasierten Verfahren hatte. Für die Teilstichprobe der Koordinator_innen kann aufgrund der geringen Stichprobengröße von $n = 4$ keine Aussage getroffen werden.

Insgesamt zeigten sich im Mittel durchweg hohe Werte hinsichtlich der Einstellungen gegenüber evidenzbasierten Verfahren. So waren die Teilnehmenden in *moderatem* bis *großem* Ausmaß bereit, eine evidenzbasierte Therapie anzuwenden sowohl wenn sie dazu angeordnet würden, als auch wenn sie den Anreiz des Verfahrens als hoch empfänden. Die Offenheit für neue Interventionen und Therapieformen variierte sogar zwischen *großem* und *sehr großem* Ausmaß. Aber auch eine Abweichung von evidenzbasierten Verfahren lag zumindest *in gewisser Weise* vor. Verglichen mit den Werten von Aarons (2004) und Aarons et al. (2010) zeigte sich, dass die Mittelwerte damit im Schnitt einen bis zwei Skalenpunkte (für die Skala Offenheit) nach oben von diesen Werten bzw. Normen⁶ abweichen. Die Teilnehmenden des Trainings zeigten also eine vergleichsweise generell positive Einstellung gegenüber evidenzbasierten Verfahren, die durch eine besondere Offenheit, aber auch eine gewisse Kritik, die sich in der Skala Abweichung zeigte, gekennzeichnet war. Selbstverständlich muss dabei beachtet werden, dass die von Aarons und Kollegen ermittelten Werte für die USA gelten. Überraschend ist allerdings, dass die in der Studie von Aarons et al. (2010) untersuchten Personen, die als ethnische Zugehörigkeit „hispanisch“ angaben signifikant geringere EBPAS-Werte als „Weiße“ erzielten. Ob diese Einstellungen also generell auf die Psychologen und Psychologinnen in Kolumbien zutreffen, oder nur spezifisch für die Teilnehmenden des Trainings, kann aufgrund fehlender Vergleichswerte nicht festgestellt werden. Festgehalten werden kann jedoch, dass die lateinamerikanische Psychologie und psychologische Forschung durch das europäische Erbe stark empirisch,

⁶ Aarons et al. (2010) ermittelten EBPAS-Werte für über 1000 Beschäftigte im Gesundheitswesen in 26 Staaten der USA.

objektiv und quantitativ ausgerichtet ist (Alarcón, 1999), was eine positive Einstellung gegenüber evidenzbasierten Verfahren zumindest nicht unwahrscheinlich macht. Es kann angenommen werden, dass Teilnehmende offener für neue Verfahren sind, wenn sie noch nicht das Gefühl haben, bereits über eine geeignete Intervention zur Begegnung der konfrontierenden Probleme zu verfügen. So zeigte sich bei Aarons (2004), dass Offenheit positiv mit dem professionellen Status assoziiert war, so dass Praktikanten und Praktikantinnen offener waren als festangestellte Mitarbeiter und bei Aarons et al. (2010) eine negative Korrelation zwischen Offenheit und Arbeitserfahrung in Jahren.

Während sich bei den Einstellungen gegenüber evidenzbasierten Verfahren vor dem Training nahezu keine signifikanten Unterschiede zwischen Koordinator_innen und den Therapeut_innen zeigten, ergaben sich direkt nach dem Training für eine Großzahl der Items und für alle Skalen signifikante Gruppenunterschiede. Dabei erreichten die Koordinatoren und Koordinatorinnen durchgehend niedrigere Werte als die Therapeuten und Therapeutinnen. Dies ist darauf zurück zu führen, dass die Werte der Koordinator_innen tendenziell gefallen, während die Werte der Therapeut_innen tendenziell gestiegen sind. Da sich in der Gruppe der Koordinator_innen allerdings lediglich vier Personen befanden, ist eine inhaltliche Interpretation dieser Unterschiede fraglich und es soll an dieser Stelle genügen, auf den gefundenen Unterschied hinzuweisen.

4.2 Wissen

In Bezug auf das Wissen zu PTBS zeigte sich hypothesenkonform eine signifikante Veränderung über die drei Messzeitpunkte hinweg (Hypothese 1c), so dass der mittlere Wissensscore von vor dem Training auf direkt nach dem Training signifikant anstieg und dieser Zuwachs zum Follow-Up erhalten blieb.

Inhaltlich muss zunächst festgestellt werden, dass unabhängig vom Zeitpunkt hohe Wissenswerte von den Teilnehmenden erreicht wurden. Aufgrund fehlender Vergleichswerte für die eingesetzte Skala ist eine Interpretation dieses Ergebnisses zunächst erschwert. Eine Möglichkeit besteht darin, diejenigen Antwortmöglichkeiten näher zu betrachten, die lediglich von einem vergleichsweise geringen Teil der Teilnehmenden richtig gelöst werden konnten. So ist auffallend, dass vor und direkt nach dem Training immerhin über die Hälfte der Teilnehmenden ankreuzte, dass mittels der CAPS eine PTBS diagnostiziert werden kann,

zum Follow-Up-Zeitpunkt jedoch nur noch gut ein Viertel. Eine mögliche Erklärung dafür könnte sein, dass die CAPS im Programm der ACR nicht zur Diagnostik verwendet wird, sondern die andere richtige Alternativantwort bei Frage 1, die PDS. Ob dies aber tatsächlich der entscheidende Faktor war, kann nicht sicher beantwortet werden.

In Bezug auf die Frage, welche Eigenschaften von Ereignissen besonders stark mit der Entwicklung von PTBS assoziiert sind, kreuzten vor dem Training weniger als die Hälfte die richtige Option 3c (impotencia / Hilflosigkeit) an. Dies mag auf die Mehrdeutigkeit des spanischen Begriffs „impotencia“ zurückzuführen sein. Während in der vorliegenden Rückübersetzung „Machtlosigkeit“ verwendet wurde und im Wörterbuch zudem das inhaltlich passende Konzept „Ohnmacht“ erscheint, umfasst der Begriff im medizinischen Sinne auch „Impotenz“ (*Langenscheidt Diccionario Moderno Alemán*, 2009). Für den weiteren Einsatz der Skala im Spanischen sollte daher statt „impotencia“ der Begriff aus dem spanischen DSM IV (Pichot, López-Ibor Aliño, Juan J. & Valdés Miyar, 1995) „desesperanza“ verwendet werden.

Hinsichtlich der Dauer, wie lange eine Person die PTBS-Symptomatik zeigen muss, damit die Diagnose vergeben werden kann, schwankten die Antworten zwischen einem und drei Monaten (Option 4b und 4c). Nun könnte diese Unsicherheit als einfaches Nichtwissen abgetan werden. Dennoch erscheint eine nähere Betrachtung berechtigt, da mit der Zeit zwar tendenziell immer mehr Personen die korrekte Antwort wählen, nichtsdestotrotz über die Hälfte auf die falsche Option 4c (3 Monate) beharrten. Klar ist, dass im DSM 3-5 (American Psychiatric Association, 1980, 2000, 2013) durchgehend gefordert wird, dass die Symptome für mindestens einen Monat vorherrschen müssen. Möglicherweise haben die Teilnehmenden drei Monate gewählt, da bis zu dieser Zeit von einer akuten und darüber hinaus von einer chronischen PTBS gesprochen wird. Interessant ist zudem, dass von einigen Teilnehmenden die Option sechs Monate handschriftlich ergänzt wurde. Nach DSM kann die Diagnose weiter spezifiziert werden mit dem Zusatz „mit verzögertem Beginn“, wenn die volle Ausprägung der PTBS frühestens sechs Monate nach dem traumatischen Ereignis auftritt. Alles in allem führte das Training allerdings nicht dazu, dass diese Frage signifikant besser beantwortet werden konnte, weshalb in zukünftigen Trainings hierauf ein größeres Augenmerk gelegt werden könnte. Konjunktiv daher, da in der Tat in gegebenem Kontext, in dem die Klienten und Klientinnen bereits meist Jahre mit den Folgen der belastenden

Ereignissen konfrontiert sind, gefragt werden kann, ob diese doch zum Teil willkürlich erscheinenden Schwellenwerte (Stieglitz, 2008) zentral für die Therapie sind, oder ob der Schwerpunkt des Trainings vielleicht nicht doch eher auf anderen Bereichen liegen sollte.

Ähnliche Bedenken könnten bezüglich der Frage neun geäußert werden („Wie viele Symptome des B-Kriteriums müssen vorhanden sein, um es als zutreffend festzustellen?“). Zwar führte das Training zu einer signifikanten Steigerung der richtigen Antworten (9a, ein Symptom), doch zum Follow-Up wählten wiederum anteilmäßig mehr Personen die Option 9c (drei Symptome). Anders als bei der Frage vier, könnte diese Meinung in gegebenem Kontext allerdings eher dazu führen, dass Personen, die die DSM-Kriterien einer PTBS erfüllen würden, nicht erkannt werden und daher keine entsprechende Unterstützung erhalten. Ob es dazu tatsächlich kommt oder nicht, kann aufgrund der vorliegenden Datenlage nicht beurteilt werden. Eine mögliche Alternativerklärung wäre, dass die Therapeut_innen in der diagnostischen Praxis mit einem digitalen Auswertungsprogramm arbeiteten, so dass die Kenntnis über die genaue Anzahl der Symptome und die Entscheidung, ob eine PTBS vorliegt oder nicht, von diesem übernommen wurde.

Hinsichtlich der Bestandteile des A-Kriteriums, ob ein traumatisches Ereignis vorlag oder nicht, wurde von vielen Teilnehmenden richtig erkannt, dass eine subjektive Reaktion beteiligt sein muss (8c). Die Rolle der objektiven Reaktion (8a) wurde zwar durch das Training eher wahrgenommen, dennoch aber nur von einem geringen Teil der Teilnehmenden. Dass die Teilnehmenden größtenteils auch keine der anderen beiden verbleibenden Optionen wählten, lässt darauf schließen, dass sie sich nicht bewusst waren, dass sich das A-Kriterium aus zwei Bereichen zusammensetzt. Möglicherweise war die Notwendigkeit, dass eine Konfrontation mit drohendem Tod oder Gefahr der körperlichen Unversehrtheit gegeben sein muss, zu offensichtlich oder aber, die Teilnehmenden haben diese Bedeutung nicht in den Worten „objektive Reaktion“ erkannt. Ob die Teilnehmenden generell wussten, was eine traumatische Situation nach dem DMS kennzeichnet oder nicht, kann über die vorliegenden Daten nicht beurteilt werden. Möglicherweise wäre hier anzuraten, ein offenes Antwortformat zu wählen, so dass die Teilnehmenden ihr Wissen in eigenen Worten formulieren können.

Die Hypothese, dass sich das Wissen über die Messzeitpunkte hinweg verändert, konnte sich wie angesprochen bewähren. Sowohl dass das Wissen durch das Training,

insofern hier von einem kausalen Zusammenhang ausgegangen werden kann, anstieg, als auch, dass es zum Follow-Up wieder etwas abfällt, jedoch über einem Niveau von vor dem Training verweilt, konnte erwartet werden und erscheint nicht weiter verwunderlich. Die Effektivität des Trainings kann in dieser Hinsicht als gesichert angenommen werden. Unter Berücksichtigung der vorherigen Ausführungen kann zudem ergänzt werden, dass das Training auch bei bereits vorhandenem Wissen auf einem hohen Niveau zu einer weiteren Steigerung beitragen konnte, was wiederum für die Qualität des Trainings spricht. Zusammenfassend zeigten sich also die erwarteten Veränderungen trotz der Wissensdefizite in den einzeln diskutierten Bereichen, die in weiteren Trainings stärker berücksichtigt werden sollten.

4.3 Verhalten in der Praxis

Im Schnitt zeigte sich für jede der drei Phasen eine gute Umsetzung der erfassten Inhalte, weshalb die Hypothese 2 als bewährt gelten kann. Da die Teilnehmenden vorab über keine Informationen zur Umsetzung der NET verfügten, ist nicht anzunehmen, dass sie sich das Wissen über einen anderen Weg als das durchgeführte Training angeeignet haben. Daraus wird gefolgert, dass das Training geeignet ist, die grundlegenden Elemente der NET zu erlernen und in der Praxis anzuwenden.

In Bezug auf die einzelnen Bereiche zeichneten sich allerdings einige Tendenzen ab, auf die bei einer weiteren Anwendung des Trainings Rücksicht genommen werden kann. So zeigte sich, dass vor allem hinsichtlich der Kenntnis der Diagnoseinstrumente noch einige Lücken bestanden. Zwar muss hierbei auch bedacht werden, dass die Diagnostik, die in den Zentren der ACR betrieben wird, sehr umfangreich ist und eine Vielzahl von Instrumenten zum Einsatz kommen. Ungeachtet dessen könnte überlegt werden, wie die beschränkte Zeit besser genutzt werden kann. Eine Möglichkeit wäre beispielsweise, die verwendeten Instrumente den Teilnehmenden bereits vor dem Training zur Lektüre zukommen zu lassen oder mehr Gruppenarbeiten mit eigener Initiative in diesem Bereich. Letztendlich muss allerdings stets im Auge behalten werden, dass es sich bei den beobachteten Verhaltensweisen um die Vertrautheit mit dem Material sowie dem generellen Interviewstil handelte, nicht aber die korrekte oder inkorrekte Umsetzung. Daher wäre eine bessere Umsetzung der Diagnostik vielleicht wünschenswert, im Rahmen beschränkter Ressourcen und des Trainings der NET-spezifischen Kenntnisse Vorzug bietend aber hinnehmbar.

Zudem zeigten die Teilnehmenden vor allem beim Innervieren des emotionalen Gedächtnisses und der Verknüpfung von „heißem“ und „kaltem“ Gedächtnis Schwierigkeiten. Ersteres ist nicht weiter verwunderlich, da es nicht nur für den Klienten oder die Klientin schwierig ist, sich den emotionalen Aspekten des Erlebten auszusetzen, sondern auch für den Therapeuten oder die Therapeutin. Dies rührt nicht nur da her, dass die Therapeut_innen die Klient_innen vor ihrem eigenen Leid bewahren wollen. Im Kontext des anhaltenden Konflikts in Kolumbien muss auch gesehen werden, dass die Kenntnis der Erfahrungen der Klient_innen tendenziell ein Risiko für die Therapeut_innen darstellen kann (C. Bueno, persönl. Mitteilung, 19.08.2014). Somit ist nicht nur von Seiten der Klient_innen Vermeidung zu beobachten, sondern auch von Seiten der Therapeut_innen. Da allerdings aus Sicht der NET eine Aktivierung und Verknüpfung sowohl des emotionalen als auch des kontextuellen Gedächtnisses für eine Besserung notwendig ist (Schauer et al., 2011), sollten Strategien für das Training gefunden werden, die es den Teilnehmenden ermöglichen, diese Kompetenzen noch besser zu erwerben.

4.4 Zusammenhänge der Konstrukte

Im Hinblick auf die im IMTEE und Hypothese 3 angenommenen Zusammenhänge ergab sich ein gemischtes Bild, so dass sich nicht alle Teilhypothesen bewähren konnten. An erster Stelle zeigten wider Erwarten weder direkt nach dem Training noch zum Follow-Up signifikante Zusammenhänge zwischen den Einstellungen und dem Wissen (Hypothese 3a). Die in vorhergehenden Studien gefundenen Zusammenhänge von Selbstwirksamkeit und kognitivem Lernen (z. B. Bell & Kozlowski, 2002, Cannon-Bowers, Salas, Tannenbaum & Mathieu, 1995) konnten also nicht repliziert werden. Die Korrelation der Einstellungen gegenüber evidenzbasierten Verfahren mit dem Wissen wurde so bisher nicht untersucht und wies daher von Beginn an einen explorativen Charakter auf. Die durchgehend niedrigen und nicht signifikanten Korrelationen legen den Schluss nahe, dass es keinen Zusammenhang zwischen den beiden Bereichen gibt. Eine mögliche Erweiterung wäre es, nicht wie hier geschehen eine breite Palette von Wissen zu Diagnose, Symptomen und Therapie der PTBS zu erfassen, sondern stattdessen spezifischere Kenntnisse der NET deren Grundlagen.

Bei der Überprüfung der Zusammenhänge zwischen dem Verhalten in der Praxis und den weiteren Variablen (Hypothese 3b-d & 3h) wurden keine signifikanten Korrelationen in der Gesamtgruppe gefunden. Aufgrund der von Alliger et al. (1997) in ihrer Metaanalyse

gefundenen mittleren Korrelation von Lernen und dem Transfer in die Praxis von lediglich $r = .13$ wäre für das Auffinden dieses Zusammenhangs weitaus mehr Teilnehmer nötig gewesen. Da nach der Aufteilung auf die verschiedenen Phasen lediglich sehr geringe Fallzahlen vorlagen, wird an dieser Stelle bewusst auf eine tiefere Interpretation verzichtet und stattdessen die Ergebnisse nur nochmals kurz dargestellt. Für die Therapeut_innen die in der Phase der Diagnostik beurteilt wurden zeigten sich höhere Werte in den therapeutischen Fähigkeiten für diejenigen, die bereits länger als Psycholog_innen gearbeitet hatten. Für die Beurteilungen in der Expositionsphase zeigte sich dieser Zusammenhang nicht, dafür zeigte sich sowohl vor dem Training als auch zum Follow-Up ein positiver Zusammenhang des DLE-Gesamtwerts mit dem Wissen zu PTBS. Ob diese unterschiedlichen Tendenzen tatsächlich mit den untersuchten Bereichen assoziiert sind, oder ob die Zusammenhänge vielmehr als Resultat der Personenauswahl gesehen werden müssen, sollte anhand einer deutlich größeren Stichprobe näher untersucht werden.

Die Hypothesen 3e-g konnten sich zum Teil bewähren, so dass es einzelne signifikante Korrelationen zwischen den bisherigen Erfahrungen bzw. individuellen Eigenschaften und den weiteren erfassten Konstrukten gab, deren Bedeutung im Folgenden diskutiert wird. Die Zusammenhänge von Geschlecht und den Einstellungen gegenüber evidenzbasierten Verfahren spiegeln genau die Befunde von Aarons (2004) wieder. Ebenso wie in der groß angelegten Studie mit über 1000 Teilnehmenden zeigten sich hier zum Follow-Up, signifikante Korrelationen zwischen Geschlecht und den Skalen Anordnung, Anreiz sowie der EBPAS-Gesamtskala. Woran es liegt, dass Frauen evidenzbasierten Verfahren gegenüber positiver eingestellt sind, kann an dieser Stelle nicht abschließend beantwortet werden. Eine mögliche Interpretation kann von der Forschung zu Geschlechterrollen geliefert werden. So zeigte sich in einer Studie von Díaz-Loving, Díaz-Guerrero, Helmreich und Spence (1981), dass Adjektive wie *stark* und *dominant* als „männlich“ beschrieben wurden und *gehorsam* und *unterwürfig* als „weiblich“. Zudem ergab sich in der Studie, dass diese Zuschreibung in Mexiko als vergleichsweise höher ausfiel als in den USA. Auch in einer Studie aus Kolumbien von Martínez-Gómez, Guerrero-Rodríguez und Rey-Anacona (2012), zeigt sich bereits in der Skalenbenennung *sumisión* (Unterwerfung) und *machismo* (Männlichkeitskult), für zwar negative, aber dennoch typisch „weibliche“ bzw. „männliche“ Eigenschaften, eine klare

Tendenz. Vor diesem Hintergrund scheint es nicht weiter überraschend, dass Frauen eher als Männer bereit waren, sich einer Anordnung bezüglich evidenzbasierter Therapien zu fügen.

Ein weiteres interessantes Ergebnis ist der positive Zusammenhang zwischen den Tätigkeitsjahren als Psycholog_in und der Zuversicht, NET effektiv anwenden zu können. Während dieser vor dem Training noch hoch signifikant ausfiel, war er nach dem Training nicht mehr vorhanden. Eine mögliche Erklärung hierfür wäre, dass diejenigen Teilnehmenden, die bereits länger als Psychologe oder Psychologin tätig waren, vor dem Training aus dieser Tätigkeit mehr Selbstvertrauen schöpfen konnten, das Training aber dazu beitrug, dass diese Unterschiede nivelliert wurden. Dies steht im Einklang mit den Ergebnissen von Bandura (1977), dass Selbstwirksamkeit vor allem durch eigene Erfolgserlebnisse gesteigert werden kann. In wie fern allerdings der Befund dazu passt, dass lediglich beim letzten Erhebungszeitpunkt die Selbstwirksamkeit umso höher war desto länger die Therapeut_innen bereits mit Opfern bzw. Tätern des kolumbianischen Konflikts gearbeitet haben, ist unklar.

4.5 Reaktionen

Bezüglich der Reaktionen auf das Training wurden keine spezifischen Hypothesen formuliert, weshalb im Folgenden lediglich eine Kontextualisierung der Ergebnisse erfolgt. Die sehr hohe Zufriedenheit sowohl mit dem Inhalt des Trainings als auch dem Training als Ganzem, verknüpft mit der Angabe, dass alle Teilnehmenden das Training weiterempfehlen würden, darf ohne weiteres als Zeichen für eine hohe Qualität des Trainings gesehen werden und entspricht auch den Erfahrungen der NET-Trainingsevaluation von Pfeiffer (2011). Da es sich bei der Erhebung um eine anonyme Fragebogenstudie handelte, ist nicht davon auszugehen, dass die durchweg hohen Werte auf soziale Erwünschtheit zurückzuführen sind. Nichtsdestotrotz sollte das weitere Setting des Trainings bei der Interpretation beachtet werden. Die Teilnehmenden setzten sich zusammen aus Psychologen und Psychologinnen, die eine hohe Motivation zeigten, sich für die fordernde und teilweise in der Bevölkerung als gefährlich geltende Arbeit mit den ehemaligen Kämpfern und Kämpferinnen der Guerillas und Paramilitärs beworben und im Schnitt bereits über 3,5 Jahre in verschiedener Form mit dieser Population gearbeitet haben. Auch wenn etwa 65 % der Teilnehmenden bereits vor dem Training über eine psychotherapeutische Fort- bzw. Ausbildung verfügten, so waren diese meist nicht spezifisch auf die Behandlung der PTBS-Problematik ausgerichtet, die einen

Großteil des Profils dieser Therapeut_innen ausmacht. Die Deckung dieses Bedarfes weist also auf die Angemessenheit des Trainingsinhaltes hin und spiegelt sich in der hohen Zufriedenheit der Teilnehmenden wieder. Aber auch für den anderen Teil der Teilnehmenden, die Koordinator_innen, stellt das Training nicht nur einen wichtigen Meilenstein innerhalb des Programms der ACR dar, sondern trifft auch deren Interesse⁷. Zudem bot das Training die Möglichkeit sich mit Kollegen und Kolleginnen aus anderen Teilen des Landes zu vernetzen und auszutauschen. Nicht zuletzt muss angemerkt werden, dass das Training von internationalen Experten auf dem Gebiet der Traumatherapie durchgeführt wurde und somit in den Augen der Teilnehmenden eine ganz besondere Lernmöglichkeit darstellte.

Über die Angemessenheit des Trainingsinhalts- und Designs hinaus können die Reaktionen auf das Training, wie vom IMTEE angenommen (Alvarez et al., 2004) und in Metaanalysen gezeigt (Alliger et al., 1997), jedoch nicht zur Vorhersage der weiteren Evaluationskomponenten Einstellung, Wissen und Verhalten in der Praxis verwendet werden. Auch in dieser Studie zeigte sich also, dass die Evaluation der Reaktionskomponente alleine, nicht geeignet ist, Rückschlüsse auf den Erfolg des Trainings zu ziehen und stets weitere Komponenten erhoben werden müssen.

4.6 Limitationen

Eine der größten methodischen Limitationen dieser Studie ist das gewählte Design. Da aufgrund der Gegebenheiten vor Ort keine Kontrollgruppe verwendet werden konnte, kann dies die interne Validität der Studie schmälern, da Alternativerklärungen für die Veränderungen nicht ausgeschlossen werden können. Auch wenn vielen der Gefahren der internen Validität (vgl. Goldstein & Ford, 2002) bereits durch das gewählte Design adäquat begegnet wurde, sollte für zukünftige Studien dennoch eine alternative Herangehensweise gewählt werden. Als besonders geeignet erscheint die internal referencing strategy (IRS) nach Haccoun und Hamtiaux (1994). Dabei werden neben den trainingsrelevanten Items, die Wissen oder Fähigkeiten erfassen, auf die innerhalb des Trainings eingegangen wurde, zusätzlich solche Items verwendet, die zwar logisch und sachlich verwandte Inhalte erfragen,

⁷ wie beispielsweise eine Whatsapp-Statusmeldung nahelegte, in der sich eine Teilnehmende freut, ein Zertifikat in Traumatherapie zu haben

welche allerdings nicht durch das Training abgedeckt wurden (=irrelevante Items). Im Anschluss kann dann die Zeit mal relevant-irrelevant Korrelation berechnet werden, wobei davon ausgegangen wird, dass durch das Training die Werte der relevanten Items signifikant zunehmen, die der irrelevanten Items jedoch nicht. Somit kann trotz fehlender Kontrollgruppe der als „Geschichte“ bezeichneten Gefahr der internen Validität effektiv vorgebeugt, bzw. dafür kontrolliert werden (Haccoun & Hamtiaux, 1994).

Eine weitere Limitation dieser Untersuchung ist ihr Fokus. Zwar konnten einige der Evaluationsfunktionen nach Stockmann (2000) zufriedenstellend umgesetzt werden, jedoch war dies hinsichtlich der Optimierungsfunktion nur zum Teil möglich. Wie in der Betrachtung der Ergebnisse hinsichtlich der einzelnen Items geschehen, konnten bereits einige inhaltliche Optimierungsmöglichkeiten gefunden werden. Jedoch können beispielsweise keine Aussagen darüber gemacht werden, welche Methoden der Wissensvermittlung besonders effektiv waren, ob es zu viele oder zu wenig praktische Übungen und Rollenspiele gab oder welche Themen noch weiter vertieft werden sollten. In diesem Sinne wäre es sinnvoll, in zukünftigen Trainingsevaluationen noch mehr darauf zu achten, Bereiche zu erheben, die den Planern und Planerinnen sowie den Trainern und Trainerinnen nützliche Informationen zum Optimierungspotential des Trainings liefern.

Auch hinsichtlich der erfassten Konstrukte muss eine kritische Reflexion erfolgen. Besonders auffallend ist, dass zu Beginn festgestellt wurde, dass eine Evaluation aller Ebenen des IMTEE erfolgen sollte, sich im weiteren Verlauf allerdings keine „Outcome-Variablen“ zum Therapieerfolg finden. Hierzu sei angemerkt, dass die entsprechenden Variablen im Rahmen der Supervision von der ACR zwar erhoben wurden, zum Zeitpunkt der Fertigstellung dieser Arbeit jedoch noch nicht verfügbar waren. Neben dieser kompletten Ebene gibt es auf den hier untersuchten Ebenen stets eine Vielzahl weiterer Einflussfaktoren, wie Motivation oder spezifische Trainings- und Organisationscharakteristika (Tannenbaum et al., 1993), die hier, um den Rahmen nicht zu sprengen, nicht erhoben wurden, jedoch in Folgeuntersuchungen berücksichtigt werden sollten.

Neben diesen methodischen Aspekten gibt es auch für die Interpretation der Ergebnisse und die künftige Forschung einige Dinge, die in Bezug auf die eingesetzten Erhebungsinstrumente beachtet werden müssen.

Für die Skalen der EBPAS bewegte sich die Itemschwierigkeit in einem angemessenen mittleren Bereich. Die Beurteilung der Trennschärfen fiel dagegen weniger positiv aus. So ergaben sich sowohl für die einzelnen Subskalen, als auch für die Gesamtskala teils negative Trennschärfen, so dass Personen mit einem hohen Wert auf der Skala, eine niedrige Ausprägung auf dem Item aufwiesen und umgekehrt. Dies kann ein Hinweis darauf sein, dass das Item nicht wie beabsichtigt verstanden wurde. Tatsächlich wiesen drei der vier Items der Skala Abweichung, die invertiert in die Gesamtskala eingehen, eine negative Trennschärfe auf. Für den weiteren Einsatz der EBPAS sollte für diese Items überprüft werden, ob die teils negative Formulierung abgeändert und falls nicht, doch zumindest visuell hervorgehoben werden kann, z. B. durch Fettdruck oder Unterstreichung (Krohne & Hock, 2007).

Ebenso kritisch sind die internen Konsistenzen der einzelnen EBPAS-Skalen zu sehen, die vor allem für die Skala Abweichung stets in einen inakzeptablen Bereich fallen. Analog zu den Ergebnissen von Aarons (2004) und Aarons et al. (2010) zeigte sich auch in dieser Studie wieder, dass die Skala Abweichung mit Abstand die geringste Homogenität aufweist. Möglicherweise sollte in Betracht gezogen werden, dass die Abweichung von evidenzbasierten Verfahren kein so einheitliches Konstrukt darstellt, wie von den Autoren angenommen.

Die Itemschwierigkeit der Wissenskala war im Schnitt angemessen, mit der Tendenz eher zu leicht zu sein. So konnte beispielsweise jeder und jede Teilnehmende richtig zuordnen, wer eine PTBS entwickeln kann. Für den weiteren Einsatz der Wissenskala sollten diese Items im Blick behalten werden. Da es sich bei der Stichprobe, wie bereits angemerkt, um eine Auswahl besonders motivierter Psycholog_innen handelt, kann an dieser Stelle aber noch nicht empfohlen werden, gewisse Items nicht mehr zu verwenden. Dazu sollte die Itemschwierigkeit zunächst bei einer breiter gefächerten Stichprobe überprüft werden.

Bei der weiteren Itemanalyse zeigten sich durchwachsene Ergebnisse. Mit negativen bis mittelmäßigen Trennschärfen und einer ungenügenden internen Konsistenz ist es fraglich, ob von einer homogenen Skala *Wissen* ausgegangen werden kann. Auch wenn Bortz und Döring (2006) darauf hinweisen, dass extreme Itemschwierigkeiten zu niedrigeren Trennschärfen führen können, muss in diesem Falle wohl eher davon ausgegangen werden, dass ein hohes Wissen in manchen Bereichen nicht mit hohem Wissen in den übrigen erfragten Bereichen einhergehen muss. Durch schrittweise Elimination einzelner Items mit

negativer Korrelation mit dem Gesamtwert konnte eine Zusammenstellung von Items erreicht werden, die eine deutlich höhere, dennoch immer noch fragwürdige interne Konsistenz aufweist. Da zudem aus den verbleibenden Items keine weitere Subskala gebildet werden konnte und die Information dieser Items damit verloren ginge, wurde weiterhin die Gesamtskala verwendet. Für die zukünftige Verwendung dieser Wissensskala sollte allerdings anhand einer größeren Stichprobe untersucht werden, inwiefern eine homogenere Skala, gegebenenfalls durch Subskalen oder andere Items, gebildet werden kann.

In Bezug auf die Gütekriterien der Wissensskala kann davon ausgegangen werden, dass sowohl Durchführungs- als auch Auswertungsobjektivität aufgrund des geschlossenen Formats gegeben sind. Die Interpretationsobjektivität kann nicht ohne weiteres bestimmt werden, da noch keine etablierten Standards für diesen Fragebogenteil existieren. Die interne Konsistenz als Maß für die Reliabilität der Wissensskala ist ungenügend und sollte bei zukünftiger Anwendung nach Möglichkeit verbessert werden. Ob die Skala das Wissen zu PTBS valide erfasst, kann mit gutem Gewissen zum Teil verneint werden. Während die Inhaltsvalidität zumindest zum Teil gegeben ist, da die vorhandenen Items im Bezug zur PTBS stehen und somit, wie die an der Konstruktion mitwirkenden Experten bestätigen konnten, wichtige Aspekte erfassen, muss auch klar gesehen werden, dass durch die beschränkte Anzahl an Fragen, niemals repräsentativ das gesamte Wissen zu PTBS erfasst werden kann. Das erfasste Wissen konzentriert sich stark auf die Diagnostik und vernachlässigt die Therapie.

Für den Beurteilungsbogen DLE-NET sollen zunächst die hohen Trennschärfen und internen Konsistenzen betrachtet werden, die an sich für die Qualität des Instruments sprechen (Bortz & Döring, 2006). Diese Werte legen nahe, dass die ausgewählten Kriterien oder Beobachtungsbereiche, trotz der unterschiedlichen Facetten inhaltlich verbundene Komplexe bilden. Eine alternative Erklärung wäre, dass die tendenziell ähnliche Bewertung einer Person in allen Bereichen auf einen Halo-Effekt zurückgeführt werden kann, so dass für die einzelnen Bewertungen nicht mehr die spezifischen Bereiche zugrunde gelegt wurden, sondern der generelle Eindruck, dass die Person *schlecht*, *gut* oder *sehr gut* arbeitet. Aufgrund des Hintergrunds und des Trainings der Supervisorin in Verhaltensbeobachtung ist in diesem Falle allerdings nicht von dieser alternativen Sichtweise auszugehen.

Hinsichtlich der Gütekriterien des DLE-NET kann festgehalten werden, dass die Objektivität im Rahmen des Möglichen durch eine standardisierte Anleitung und Durchführung gesichert wurden. Eine weitere Steigerung der Objektivität wäre gegebenenfalls dadurch zu erreichen gewesen, dass die Verhaltensbeobachtung anhand von aufgezeichnetem Video- oder Tonmaterial durchgeführt worden wäre, was allerdings in gegebenem Kontext in Kolumbien nicht möglich war, da von Seiten der ACR keine Genehmigung erfolgte. Eine weitere Möglichkeit die Objektivität zu sichern ist die Berechnung der Interraterreliabilität. Da aber die Anwesenheit einer weiteren Person in dem Therapiesetting als unzumutbar angesehen wurde, konnte von dieser Möglichkeit kein Gebrauch gemacht werden. Angesichts dessen, dass der DLE-NET von Experten der NET mitentwickelt wurde, kann angenommen werden, dass die Inhaltsvalidität gegeben ist. Um die Konstruktvalidität zu überprüfen, wäre der geeignetste Ansatz die Multitrait-Multimethod-Methode, anhand derer die konvergente und diskriminante Validität bestimmt werden kann (Bortz & Döring, 2006). Da die Validierung des DLE-NET allerdings nicht den zentralen Teil dieser Arbeit darstellt, soll an dieser Stelle genügen, dass eine angemessene Konstruktvalidität mittels logisch-inhaltlichen Analysen (vgl. Bortz & Döring, 2006, S. 202) der Relation der Items mit dem interessierenden Konstrukt festgestellt wurde. Da die Frage, ob qualitative Forschungstechniken generell reliabel sein sollten, selbst unter Experten umstritten ist (Bortz & Döring, 2006, S. 327), wurde an dieser Stelle verzichtet, näher auf dieses Gütekriterium einzugehen.

4.7 Schlussfolgerungen und Ausblick

Zusammenfassend zeigte sich, dass das durchgeführte Training den Teilnehmenden, das nötige Wissen sowie die grundlegenden Fertigkeiten vermittelte, um die NET angemessen durchführen zu können. Basierend auf dem generellen Wirksamkeitsnachweis der NET in den verschiedensten Ländern und Kontexten (s. o.) kann damit der Schluss gezogen werden, dass auch ein relativ kurzes, aber fokussiertes NET-Training das nötige Handwerkszeug vermitteln kann, um den Herausforderungen der Reintegration von ehemaligen Kämpferinnen und Kämpfern illegaler Gruppierungen zu begegnen. In kurzer Zeit und oftmals damit verbunden vergleichsweise geringen finanziellen Aufwendungen können Psychologen und Psychologinnen zu Traumatherapeut_innen weitergebildet

werden. Nicht nur für den hier untersuchten Kontext in Kolumbien, sondern für Krisenregionen auf der ganzen Welt ist dies eine wichtige Erkenntnis.

Um die weitere Entwicklung der Kenntnisse und Fertigkeiten in der Traumatherapie nachzuverfolgen, wäre eine nachfolgende Erhebung mindestens ein Jahr nach dem Training wünschenswert. So könnte überprüft werden, ob mit zusätzlicher praktischer Erfahrung auch diejenigen Personen, die ein gutes halbes Jahr nach dem Training noch einige Defizite in der Umsetzung der Techniken aufwiesen, diese Lücken schließen konnten, oder ob es stets Personen geben wird, die dazu nicht in der Lage sind. Dies legt auch die Frage nahe, ob es stabile personenbezogene Merkmale gibt, die den späteren Erfolg in der Therapie mitbestimmen.

Die hier gefundenen Zusammenhänge können relativ wenig darüber aussagen. Neben der Erfassung weiterer Variablen und der Replikation an einer größeren Stichprobe sollte in zukünftigen Studien aber vor allem auch die Herangehensweise erweitert werden. Zwar war die Untersuchung, *warum* es zu Veränderungen kommt Teil dieser Studie, doch stand die Frage *was* sich verändert stets klar im Mittelpunkt. *Dass* es zu Veränderungen bei den Teilnehmenden kommt, konnte mit dieser Untersuchung nun gezeigt werden. Die weitere Forschung sollte sich also, neben der Überwindung der aufgezeigten Limitationen der eingesetzten Erhebungsinstrumente, vor allem auf die spezifischen Funktionsweisen und einflussnehmende personelle und kontextuelle Faktoren konzentrieren.

Die generell positiven Einstellungen gegenüber evidenzbasierten Verfahren widersprechen den Ergebnissen von Aarons et al. (2010) für den lateinamerikanischen Kulturkreis. Auch wenn aufgrund der kleinen und noch dazu sehr selektierten Stichprobe keine Überinterpretation der Ergebnisse geschehen sollte, ist es dennoch angebracht, vor kulturellem Schubladendenken zu warnen. Bei kulturübergreifenden Projekten wie diesem ist die Reflexion des eigenen Hintergrundes, vor welchem Interpretationen stattfinden, von besonderer Bedeutung. In diesem Sinne ist an dieser Stelle die enge Zusammenarbeit und der stetige Austausch mit Personen, die in beiden Kulturkreisen gelebt und gearbeitet haben, hervorzuheben. Nur so war es möglich, die gefundenen Ergebnisse in den richtigen Kontext zu stellen. Für weitere Studien wird mit Nachdruck empfohlen, diese Herangehensweise zu verfolgen und nicht der Gefahr eurozentristischer Forschung zu verfallen (Teo, 2005, S. 164).

Abschließend soll nochmals an die Worte von Carlos Cano (Reintegración ACR, 2010) erinnert werden und seine Erfahrung, dass nur wer sein Herz demobilisiert zum Frieden beitragen kann. In diesem Sinne ist diese Arbeit all denjenigen gewidmet, die in Kolumbien und anderswo unter den Folgen gewalttätiger Auseinandersetzungen leiden. In der Hoffnung, dass durch die stattgefundene Forschung ein kleiner Teil dazu beigetragen werden kann, dass alle Beteiligten einen friedlicheren Weg gehen können.

5 Literaturverzeichnis

- Aarons, G. A. (2004). Mental health provider attitudes toward adoption of evidence-based practice: The evidence-based practice attitude scale (EBPAS).
- Aarons, G. A., Glisson, C., Hoagwood, K., Kelleher, K., Landsverk, J. & Cafri, G. (2010). Psychometric properties and United States national norms of the evidence-based practice attitude scale (EBPAS). *Psychological assessment*, 22 (2), 356–365.
- Addis, M. E. & Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes toward psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology*, 68 (2), 331–339.
- Agencia Colombiana para la Reintegración. (2011). *Manual del modelo de atención psicosocial para la paz - MAPAZ. El desarrollo de competencias y habilidades en la población en proceso de reintegración y su entorno próximo.*
- Agencia Colombiana para la Reintegración. (2013). Resolución número 754.
- Agencia Colombiana para la Reintegración. (2014a). *Informe de gestión ACR. Enero - Diciembre 2013.* Zugriff am 16.04.2014. Verfügbar unter http://www.reintegracion.gov.co/Documentos_Contentido/1697_informe_gestion_2013.pdf
- Agencia Colombiana para la Reintegración. (2014b). *La corresponsabilidad de los actores externos frente a la Política de Reintegración.* Zugriff am 12.08.2014. Verfügbar unter <http://www.reintegracion.gov.co/Reintegracion/procesodereintegracion/corresponsabilidad>
- Agencia Colombiana para la Reintegración. (2014c). *Reintegración en Colombia. Hechos & Datos.* Zugriff am 14.04.2014. Verfügbar unter <http://www.reintegracion.gov.co/ACR/cifras/colombia.pdf>
- Agencia Colombiana para la Reintegración. (2014d). *Reseña histórica.* Zugriff am 14.04.2014. Verfügbar unter http://www.reintegracion.gov.co/Quienes_Somos/Historia
- Alarcón, R. (1999). La psicología en América Latina en el siglo XX. Un análisis de sus características. *Revista de Psicología de la PUCP*, XVII (2), 132–146.
- Alliger, G. M. & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology*, 42 (2), 331–342.
- Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H. & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, 50 (2), 341–358.

- Alvarez, K., Salas, E. & Garofano, C. M. (2004). An integrated model of training evaluation and effectiveness. *Human Resource Development Review*, 3 (4), 385–416.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Anaya, L. (2007). Individual demobilization and reintegration process in Colombia: implementation, challenges and former combatants' perspectives. *Intervention*, 5 (3), 179–190.
- Bandura, A. (1977). Self-efficacy. Toward a unifying theory of behavioral change. *Psychological Review*, 84 (2), 191–215.
- Beidas, R. S. & Kendall, P. C. (2010). Training therapists in evidence-based practice: A Critical Review of Studies From a Systems-Contextual Perspective.
- Bell, B. S. & Kozlowski, S. W. J. (2002). Adaptive guidance enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology*, 55, 267–306.
- Bichescu, D., Neuner, F., Schauer, M. & Elbert, T. (2007). Narrative exposure therapy of political imprisonment-related chronic posttraumatic stress disorder and depression. *Behaviour Research and Therapy*, 45 (9), 2212–2220.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation. Für Human- und Sozialwissenschaftler* (4. Aufl.). Heidelberg: Springer Medizin Verlag.
- Bortz, J. & Lienert, G. A. (2008). *Kurzgefasste Statistik für die klinische Forschung. Leitfaden für die verteilungsfreie Analyse kleiner Stichproben* (3. Aufl.). Heidelberg: Springer Medizin Verlag.
- Bortz, J., Lienert, G. A. & Boehnke, K. (2008). *Verteilungsfreie Methoden in der Biostatistik* (3. Aufl.). Heidelberg: Springer Medizin Verlag.
- Bueno Castellanos, C. P. (2013). *Erfassung der Traumaexposition und der psychischen Folgen organisierter Gewalt. Eine Epidemiologische Studie und Vergleich zwischen den*

- Gewaltakteuren und der betroffenen Bevölkerung des bewaffneten Konflikts in Kolumbien.* Unveröffentlichte Dissertation, Universität Konstanz.
- Bueno Castellanos, C. P. (2014). *Intervención psicológica a desmovilizados con trastornos de estrés postraumático.* Unpubliziertes Manuskript.
- Cannon-Bowers, J. A., Salas, E., Tannenbaum, S. I. & Mathieu, J. E. (1995). Toward theoretically based principles of training effectiveness: A model and initial empirical investigation. *Military Psychology, 7* (3), 141–164.
- Carroll, K. M., Nich, C. & Rounsaville, B. J. (1998). Utility of therapist session checklists to monitor delivery of coping skills treatment for cocaine abusers. *Psychotherapy Research, 8* (3), 307–320.
- Contreras Rolón, T., Jaramillo de Botero, M. & Picón Amaya, A. (2006). *La Política de reincorporación a la vida civil 2003 - 2006. Una Mirada Institucional* (Ministerio del Interior y de la Justicia, Hrsg.).
- Cronbach, L. J. (1982). *Designing evaluation of educational and social programs.* San Francisco: Jossey-Bass.
- Decker, S. E., Jameson, M. T. & Naugle, A. E. (2011). Therapist training in empirically supported treatments: A review of evaluation methods for short- and long-term outcomes. *Administration and Policy in Mental Health and Mental Health Services Research, 38* (4), 254–286.
- Deutsche Gesellschaft für Psychologie (Hrsg.). (2007). *Richtlinien zur Manuskriptgestaltung.* Göttingen: Hogrefe Verlag.
- Díaz-Loving, R., Díaz-Guerrero, R., Helmreich, R. L. & Spence, J. T. (1981). Comparación transcultural y análisis psicométrico de una medida de rasgos masculinos (instrumentales) y femeninos (expresivos). *Revista de la Asociación Latinoamericana de Psicología Social, 1* (1), 3–37.
- Eder, A. (Americas Society/ Council of the Americas, Hrsg.). (2014). *Alejandro Eder - Dir. of the Colombian Agency for Reintegration.* Zugriff am 14.04.2014. Verfügbar unter <http://www.youtube.com/watch?v=fbVbn2Ldvxc>
- Elbert, T., Hermenau, K., Hecker, T., Weierstall, R. & Schauer, M. (2012). FORNET: Behandlung von traumatisierten und nicht-traumatisierten Gewalttätern mittels Narrativer

- Expositionstherapie. In J. Endrass, A. Rossegger & B. Borchard (Hrsg.), *Interventionen bei Gewalt- und Sexualstraftätern. Risk-Management, Methoden und Konzepte der forensischen Therapie* (S. 255–276). Berlin: MWV Medizinisch-Wissenschaftliche Verlagsgesellschaft.
- Ertl, V. (2013). *Beurteilungsbogen für NET-Sitzungen*. Unveröffentlichte Skala.
- Field, A. (2009). *Discovering Statistics Using SPSS (Introducing Statistical Method). (and sex and drugs and rock 'n' roll)* (Third Edition). London: SAGE Publications Ltd.
- Foa, E. B. (1995). *Post-Traumatic Stress Diagnostic Scale (PDS)*. Minneapolis, MN: National Computer Systems.
- Goldstein, I. L. & Ford, J. K. (2002). *Training in organizations. Needs assessment development and evaluation* (4. ed). Belmont, Calif: Wadsworth.
- Grupo de Memoria Histórica. (2013). *¡Basta Ya! Colombia: Memorias de Guerra y Dignidad*. Bogotá: Imprenta Nacional.
- Haccoun, R. R. & Hamtiaux, T. (1994). Optimizing knowledge tests for inferring learning acquisition levels in single group training evaluation designs. The internal referencing strategy. *Personnel Psychology*, 47 (3), 593–604.
- Herschell, A. D., McNeil, C. B., Urquiza, A. J., McGrath, J. M., Zebell, N. M., Timmer, S. G. et al. (2009). Evaluation of a treatment manual and workshops for disseminating, parent-child interaction therapy. *Administration and policy in mental health*, 36 (1), 63–81.
- Holton, E. F. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7 (1), 5–21.
- Jonkisz, E., Moosbrugger, H. & Brandt, H. (2012). Planung und Entwicklung von Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. Aufl., S. 27–74). Berlin, Heidelberg: Springer-Verlag.
- Kanning, U. P. & Staufenbiel, T. (2012). *Organisationspsychologie* (1. Aufl.). Göttingen: Hogrefe.
- Kirkpatrick, D. L. (1959a). Techniques for evaluating training programs. *Journal of ASTD* (13), 3–9.

- Kirkpatrick, D. L. (1959b). Techniques for evaluating training programs: Part 2 - Learning. *Journal of ASTD* (13), 21–26.
- Kirkpatrick, D. L. (1960a). Techniques for evaluating training programs: Part 3 - Behavior. *Journal of ASTD* (14), 13–18.
- Kirkpatrick, D. L. (1960b). Techniques for evaluating training programs: Part 4 - Results. *Journal of ASTD* (14), 28–32.
- Kirkpatrick, D. L. (1996). Great Ideas Revisited. Techniques for Evaluating Training Programs. Revisiting Kirkpatrick's Four-Level Model. *Training and Development*, 50, 54–59.
- Kirkpatrick, D. L. (1998). *Evaluating Training Programs. The Four Levels* (2nd). San Francisco: Berrett-Koehler Publishers.
- Kraiger, K. (2002). Decision-based evaluation. In K. Kraiger (Hrsg.), *Creating, implementing, and managing effective training and development. State-of-the-art lessons for practice* (S. 331–375). San Francisco, CA: Jossey-Bass.
- Krohne, H. W. & Hock, M. (2007). *Psychologische Diagnostik. Grundlagen und Anwendungsfelder*. Stuttgart: W. Kohlhammer.
- Landy, F. J. & Conte, J. M. (2010). *Work in the 21st century. An introduction to industrial and organizational psychology* (3. ed). Hoboken, NJ: Wiley.
- Langenscheidt Diccionario Moderno Alemán. Español-alemán; alemán-español.* (2009) (Nueva ed). Berlin, München: Langenscheidt.
- Maedl, A., Schauer, E., Odenwald, M. & Elbert, T. (2010). Psychological Rehabilitation of Ex-combatants in Non-Western, Post-conflict Settings. In E. Martz (Hrsg.), *Trauma Rehabilitation After War and Conflict* (S. 177–213). New York, NY: Springer New York.
- Martínez-Gómez, J. A., Guerrero-Rodríguez, S. I. & Rey-Anacona, C. A. (2012). Evaluación de la validez de constructo y la contabilidad del inventario de masculinidad y femineidad en adolescentes y adultos jóvenes colombianos. *Avances en Psicología Latinoamericana*, 30, 170–181.
- Martino, S., Ball, S., Nich, C., Frankforter, T. L. & Carroll, K. M. (2009). Correspondence of motivational enhancement treatment integrity ratings among therapists, supervisors, and observers. *Psychotherapy Research*, 19 (2), 181–193.

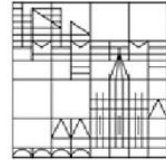
- Mathieu, J. E., Tannenbaum, S. I. & Salas, E. (1992). Influences of individual and situational characteristics on measures of training effectiveness. *Academy of Management Journal*, 35 (828-847).
- Mayring, P. (2003). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (11. Aufl.). Weinheim: Beltz.
- Metschke, R. & Wellbrock, R. (2002). *Datenschutz in Wissenschaft und Forschung* (3. Aufl.) (Berliner Beauftragter für Datenschutz und Informationsfreiheit, Hrsg.), Berlin.
- Moor, M. (2007). A new start, an open end. The reintegration of individual demobilized combatants in Colombia. *Intervention*, 5 (3), 191–202.
- Morgenstein, J. (2008). Consolidating disarmament. Lessons from Colombia's reintegration program for demobilized paramilitaries.
- Morgenstern, J., Morgan, T. J., McCrady, B. S., Keller, D. S. & Carroll, K. M. (2001). Manual-guided cognitive-behavioral therapy training: A promising method for disseminating empirically supported substance abuse treatments to the practice community. *Psychology of Addictive Behaviors*, 15 (2), 83–88.
- Noe, R. A. & Colquitt, J. A. (2002). Planning for training impact. Principles of training effectiveness. In K. Kraiger (Hrsg.), *Creating, implementing, and managing effective training and development. State-of-the-art lessons for practice* (S. 53–79). San Francisco, CA: Jossey-Bass.
- Odenwald, M., Hinkel, H. & Schauer, E. (2007). Challenges for a future reintegration programme in Somalia:.. outcomes of an assessment on drug abuse, psychological distress and preferences for reintegration assistance. *Intervention*, 5 (2), 109–123.
- Pfeiffer, A. (2011). Treating trauma patients with Narrative Exposure Therapy. Effects on the therapist and their evaluation of the therapy.
- Pichot, P., López-Ibor Aliño, Juan J. & Valdés Miyar, M. (1995). *DSM-IV. Manual diagnóstico y estadístico de los trastornos mentales*. Barcelona: Masson, S.A.
- Reintegración ACR (Autor). (2010). *Carlos Cano*. Verfügbar unter http://www.youtube.com/watch?v=ZbSD_4XmOts

- Ruf, M., Schauer, M., Neuner, F., Catani, C., Schauer, E. & Elbert, T. (2010). Narrative exposure therapy for 7- to 16-year-olds: A randomized controlled trial with traumatized refugee children. *Journal of Traumatic Stress, 23* (4), 437–445.
- Salas, E. & Cannon-Bowers, J. A. (2001). The Science of Training. A Decade of Progress. *Annual Review of Psychology, 52*, 471–499.
- Schaal, S., Elbert, T. & Neuner, F. (2009). Narrative Exposure Therapy versus Interpersonal Psychotherapy. A Pilot Randomized Controlled Trial with Rwandan Genocide Orphans. *Psychotherapy and Psychosomatics, 78*, 298–306.
- Schauer, M., Neuner, F. & Elbert, T. (2011). *Narrative Exposure Therapy. A short-term treatment for traumatic stress disorder* (2nd rev. and expanded ed.). Göttingen: Hogrefe.
- Schmitt, M. & Eid, M. (2007). Richtlinien für die Übersetzung fremdsprachlicher Messinstrumente. *Diagnostica, 53* (1), 1–2.
- Scriven, M. (1991). *Evaluation thesaurus* (4. ed). Newbury Park: Sage.
- Seng, A. C., Prinz, R. J. & Sanders, M. R. (2006). The role of training variables in effective dissemination of evidence-based parenting interventions. *International Journal of Mental Health Promotion, 8* (4), 20–28.
- Stieglitz, R.-D. (2008). *Diagnostik und Klassifikation in der Psychiatrie*. Stuttgart: Kohlhammer.
- Stockmann, R. (2000). Evaluation in Deutschland. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 11–40). Opladen: Leske & Budrich.
- Sugrue, B. & Rivera, R. (2005). ASTD 2005 state of the industry report.
- Tannenbaum, S. I., Cannon-Bowers, J. A., Salas, E. & Mathieu, J. E. (1993). *Factors that influence training effectiveness A conceptual model and longitudinal analysis*. Technical Rep. No. 93-011. Orlando: FL: Naval Training Systems Center.
- Teo, T. (2005). *The Critique of Psychology. From Kant to Postcolonial Theory*. Boston, MA: Springer Science+Business Media.

6 Anhang

A1 Fragebogen pre

Universität
Konstanz



Evaluación del entrenamiento en Terapia de Exposición Narrativa

Procedimiento y objetivo de la evaluación

Este cuestionario hace parte de la evaluación del entrenamiento en Terapia de Exposición Narrativa (TEN) en Colombia. Tiene como objetivo realizar declaraciones sobre la efectividad y la eficiencia del entrenamiento. Para ello le solicitamos llenar muy cuidadosamente este cuestionario antes y después del entrenamiento. No existen respuestas correctas ni incorrectas, lo que nos interesa es su opinión y experiencia personal a lo largo del entrenamiento.

Análisis de los datos

Los datos que usted provea serán tratados de forma anónima y en ningún caso proveerán información que pueda referirse a usted directamente. Sus datos serán utilizados exclusivamente para fines investigativos y bajo ninguna circunstancia serán entregados a terceras personas.

Derechos

Su participación en este cuestionario es voluntaria. Usted no está obligado a tomar parte de ella. Si usted decidiera no participar, no se generará ningún inconveniente ni represalia en su contra. Usted puede decidir interrumpir su participación sin necesidad de alegar razones puntuales.

Datos de contacto

Si tiene alguna pregunta, duda o sugerencia puede ponerse en contacto durante el entrenamiento con la **Dr. Claudia Bueno** o a **Christian Schart (B. Sc.)** para inquietudes durante y fuera del entrenamiento.

Christian Schart
Steinstraße 17a
78467 Konstanz
ALEMANIA
E-Mail: christian.schart@uni-konstanz.de

Con gusto le informaremos acerca de los principales resultados de la evaluación. ¡Gracias de antemano por su colaboración!

Código de identificación:

Para poder comparar sus datos antes y después del entrenamiento, necesitamos una forma de poder codificar la información. Con el fin de realizarlo de manera anónima, por favor rellene el siguiente código según lo requerido.

- Primeras dos letras del nombre de su padre: (Ejemplo: José → JO)
- Primeras dos letras del nombre de su madre: (Ejemplo: María → MA)
- Su día de nacimiento: (Ejemplo: 08.12.1972 → 08)

Actitudes

En una escala de 1 a 5 siendo 1 = muy poco confiado, 2 = poco confiado, 3 = moderadamente confiado, 4 = altamente confiado, 5 = muy altamente confiado responda la siguiente pregunta.

¿Qué tan confiado se siente usted de poder implementar la TEN de forma efectiva?	1	2	3	4	5
--	---	---	---	---	---

Las siguientes expresiones tienen que ver con su actitud personal frente a terapias manualizadas que fueron comprobadas por medio de ensayos clínicos.

Por favor indique su posición en una escala de 1 a 5 siendo 1 = En absoluto, 2 = En cierta medida, 3 = Moderadamente, 4 = En gran medida, 5 = En muy gran medida

Me agrada usar nuevos tipos de terapia y/o intervenciones para ayudar a mis clientes.	1	2	3	4	5
Estoy dispuesto a probar nuevos tipos de terapia y/o intervenciones incluso si tengo que seguir un manual de tratamiento.	1	2	3	4	5
Sé mejor que los investigadores académicos cómo cuidar a mis clientes	1	2	3	4	5
Estoy dispuesto a probar nuevos tipos de terapia y/o intervenciones desarrollados por investigadores.	1	2	3	4	5
Los tratamientos y/o intervenciones basados en investigaciones no son útiles clínicamente.	1	2	3	4	5
La experiencia clínica es más importante que utilizar terapias y/o intervenciones manualizadas.	1	2	3	4	5
No usaría terapias y/o intervenciones manualizadas.	1	2	3	4	5
Probaría una nueva terapia y/o intervención incluso si ésta es muy diferente de lo que estoy acostumbrado a hacer.	1	2	3	4	5

Si usted recibiera entrenamiento en una terapia y/o intervención que fuera nueva para usted, ¿qué tan probable sería que usted la adoptara si...:

... fuese intuitivamente atractiva?	1	2	3	4	5
... "tuviera sentido" para usted?	1	2	3	4	5
... fuera requerida por su supervisor?	1	2	3	4	5
... fuera requerida por su agencia?	1	2	3	4	5
... fuera requerida por su Estado?	1	2	3	4	5
... fuera usada por colegas que estuviesen contentos con ella?	1	2	3	4	5
... usted sintiera que ha tenido suficiente entrenamiento para utilizarla correctamente?	1	2	3	4	5

1 = En absoluto, 2 = En cierta medida, 3 = Moderadamente, 4 = En gran medida, 5 = En muy gran medida

Preguntas de conocimiento:

A continuación encontrará algunas preguntas de conocimiento sobre el **Trastorno por Estrés Postraumático (TEPT)**. Por favor marque todas las respuestas que usted considere correctas. Puede haber ninguna, una o más respuestas correctas para cada pregunta.

Diagnóstico

1. ¿Con cuáles instrumentos puede usted evaluar si una persona cumple con los criterios para el diagnóstico de un TEPT?
 - a. Escala para el Trastorno por Estrés Postraumático administrada por el Clínico (CAPS)
 - b. Escala diagnóstica de Estrés Postraumático (PSD)
 - c. Figura humana de Machover
 - d. Escala de Hamilton para la depresión (HAMD)

2. ¿Qué eventos pueden aumentar el riesgo de desarrollar un TEPT?
 - a. Ser violado sexualmente
 - b. Ser torturado
 - c. Estar involucrado en un accidente automovilístico
 - d. Ser testigo de un asesinato

3. ¿Cuáles son características de los eventos que se asocian más frecuentemente con el desarrollo de un TEPT?
 - a. Posibilidad latente de muerte
 - b. Decepción
 - c. Impotencia
 - d. Alegría intensiva

4. ¿Por cuánto tiempo debe una persona presentar síntomas de un TEPT antes de que pueda ser emitido un poder emitir un el diagnóstico?
 - a. 2 semanas
 - b. 1 mes
 - c. 3 meses
 - d. 1 año

5. ¿Quién puede desarrollar un TEPT?
 - a. Sólo hombres (Gender Difference)
 - b. Cualquiera que haya vivido suficientes experiencias traumáticas (Building-Block Effect)
 - c. Sólo los adultos (Youth-Protection Factor)

6. ¿Cuáles son los tres criterios diagnósticos principales para diagnosticar un TEPT?
 - a. Intrusiones
 - b. Estrés
 - c. Evitación
 - d. Hiperactivación

7. ¿Cuál es la característica que diferencia al TEPT de otros trastornos de ansiedad?
 - a. Falta del reconocimiento del miedo
 - b. La presencia del criterio A
 - c. El estrés crónico
 - d. Evitación de situaciones sociales

8. Según el DSM-V, ¿cómo se clasifica el criterio A?
 - a. Respuesta objetiva
 - b. Respuesta natural
 - c. Respuesta subjetiva
 - d. Respuesta directa

9. ¿Cuántos síntomas dentro del criterio B son necesarios para decretarlo como presente?
 - a. Uno
 - b. Dos
 - c. Tres
 - d. Cinco

Síntomas

10. ¿Cuáles características y/o síntomas son particularmente típicos en personas con TEPT?

- a. Actuar o sentir como si el acontecimiento traumático volviese a ocurrir
- b. Aumento de peso
- c. Pérdida de autoestima
- d. Problemas de sueño
- e. Dificultades en la concentración
- f. Disminución de la inteligencia
- g. Recuerdos recurrentes y angustiosos
- h. Euforia

Terapia

11. ¿Cuáles factores pueden influenciar positivamente el resultado de la terapia en un TEPT?

- a. Contacto físico cercano entre terapeuta y paciente
- b. Creación de una atmósfera cercana
- c. Presionar al paciente desde el principio de la terapia a describir detalladamente el evento traumático
- d. Que el/la paciente se introduce nuevamente en la experiencia traumática y siente como si la estuviera viviendo nuevamente
- e. Que el/la paciente recuerde la situación traumática (cuándo y dónde fue, al igual que cómo se sintió, siendo siempre consciente de que la situación real ya acabó

Datos sociodemográficos:

Para el análisis estadístico requerimos algunos datos sobre usted y sus experiencias.

Edad (en años):

Género: masculino femenino no especificado

En caso de haber recibido capacitación y/o formación en otras de formas de terapia, por favor especifique cuáles:

¿Cuántos años lleva ejerciendo la profesión de psicólogo?

Si usted ha trabajado antes de este entrenamiento en el programa de la ACR, por favor indique por cuántos meses:

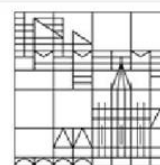
Si usted ya ha recopilado experiencia en el trabajo con las víctimas y/o victimarios del conflicto armado colombiano, por favor indique durante cuántos meses y describa brevemente sus funciones:

En caso de que usted o alguna persona cercana (familiar, amigo, ...) haya tenido contacto directo con las guerrillas o paramilitares, por favor describa brevemente el tipo de contacto establecido con estos grupos.

¡Gracias por su participación!

A2 Fragebogen post

Universität
Konstanz



Evaluación del entrenamiento en Terapia de Exposición Narrativa

Procedimiento y objetivo de la evaluación

Este cuestionario hace parte de la evaluación del entrenamiento en Terapia de Exposición Narrativa (TEN) en Colombia. Tiene como objetivo realizar declaraciones sobre la efectividad y la eficiencia del entrenamiento. Para ello le solicitamos llenar muy cuidadosamente este cuestionario antes y después del entrenamiento. No existen respuestas correctas ni incorrectas, lo que nos interesa es su opinión y experiencia personal a lo largo del entrenamiento.

Análisis de los datos

Los datos que usted provea serán tratados de forma anónima y en ningún caso proveerán información que pueda referirse a usted directamente. Sus datos serán utilizados exclusivamente para fines investigativos y bajo ninguna circunstancia serán entregados a terceras personas.

Derechos

Su participación en este cuestionario es voluntaria. Usted no está obligado a tomar parte de ella. Si usted decidiera no participar, no se generará ningún inconveniente ni represalia en su contra. Usted puede decidir interrumpir su participación sin necesidad de alegar razones puntuales.

Datos de contacto

Si tiene alguna pregunta, duda o sugerencia puede ponerse en contacto durante el entrenamiento con la **Dr. Claudia Bueno** o a **Christian Schart (B. Sc.)** para inquietudes durante y fuera del entrenamiento.

Christian Schart
Steinstraße 17a
78467 Konstanz
ALEMANIA
E-Mail: christian.schart@uni-konstanz.de

Con gusto le informaremos acerca de los principales resultados de la evaluación. ¡Gracias de antemano por su colaboración!

Código de identificación:

Para poder comparar sus datos antes y después del entrenamiento, necesitamos una forma de poder codificar la información. Con el fin de realizarlo de manera anónima, por favor rellene el siguiente código según lo requerido.

- Primeras dos letras del nombre de su padre: (Ejemplo: José → JO)
- Primeras dos letras del nombre de su madre: (Ejemplo: María → MA)
- Su día de nacimiento: (Ejemplo: 08.12.1972 → 08)

Reacciones al entrenamiento:

A continuación queremos conocer su opinión sobre el entrenamiento.

¿Recomendaría este entrenamiento a un compañero de trabajo? Sí No

En una escala de 1 a 5 siendo 1 = muy insatisfecho, 2 = insatisfecho, 3 = Ni satisfecho ni insatisfecho, 4 = satisfecho, 5 = muy satisfecho, responda las siguientes preguntas:

¿Cómo evaluaría el contenido del curso?	1	2	3	4	5
De forma general, ¿cuán satisfecho se sintió usted con el curso?	1	2	3	4	5

Actitudes

En una escala de 1 a 5 siendo 1 = muy poco confiado, 2 = poco confiado, 3 = moderadamente confiado, 4 = altamente confiado, 5 = muy altamente confiado responda la siguiente pregunta.

¿Qué tan confiado se siente usted de poder implementar la TEN de forma efectiva?	1	2	3	4	5
--	---	---	---	---	---

Las siguientes expresiones tienen que ver con su actitud personal frente a terapias manualizadas que fueron comprobadas por medio de ensayos clínicos.

Por favor indique su posición en una escala de 1 a 5 siendo 1 = En absoluto, 2 = En cierta medida, 3 = Moderadamente, 4 = En gran medida, 5 = En muy gran medida

Me agrada usar nuevos tipos de terapia y/o intervenciones para ayudar a mis clientes.	1	2	3	4	5
Estoy dispuesto a probar nuevos tipos de terapia y/o intervenciones incluso si tengo que seguir un manual de tratamiento.	1	2	3	4	5
Sé mejor que los investigadores académicos cómo cuidar a mis clientes	1	2	3	4	5
Estoy dispuesto a probar nuevos tipos de terapia y/o intervenciones desarrollados por investigadores.	1	2	3	4	5
Los tratamientos y/o intervenciones basados en investigaciones no son útiles clínicamente.	1	2	3	4	5
La experiencia clínica es más importante que utilizar terapias y/o intervenciones manualizadas.	1	2	3	4	5
No usaría terapias y/o intervenciones manualizadas.	1	2	3	4	5
Probaría una nueva terapia y/o intervención incluso si ésta es muy diferente de lo que estoy acostumbrado a hacer.	1	2	3	4	5

Si usted recibiera entrenamiento en una terapia y/o intervención que fuera nueva para usted, ¿qué tan probable sería que usted la adoptara si...:

... fuese intuitivamente atractiva?	1	2	3	4	5
... "tuviera sentido" para usted?	1	2	3	4	5
... fuera requerida por su supervisor?	1	2	3	4	5
... fuera requerida por su agencia?	1	2	3	4	5
... fuera requerida por su Estado?	1	2	3	4	5
... fuera usada por colegas que estuviesen contentos con ella?	1	2	3	4	5
... usted sintiera que ha tenido suficiente entrenamiento para utilizarla correctamente?	1	2	3	4	5

1 = En absoluto, 2 = En cierta medida, 3 = Moderadamente, 4 = En gran medida, 5 = En muy gran medida

Preguntas de conocimiento:

A continuación encontrará algunas preguntas de conocimiento sobre el **Trastorno por Estrés Postraumático (TEPT)**. Por favor marque todas las respuestas que usted considere correctas. Puede haber ninguna, una o más respuestas correctas para cada pregunta.

Diagnóstico

1. ¿Con cuáles instrumentos puede usted evaluar si una persona cumple con los criterios para el diagnóstico de un TEPT?
 - a. Escala para el Trastorno por Estrés Postraumático administrada por el Clínico (CAPS)
 - b. Escala diagnóstica de Estrés Postraumático (PSD)
 - c. Figura humana de Machover
 - d. Escala de Hamilton para la depresión (HAMD)

2. ¿Qué eventos pueden aumentar el riesgo de desarrollar un TEPT?
 - a. Ser violado sexualmente
 - b. Ser torturado
 - c. Estar involucrado en un accidente automovilístico
 - d. Ser testigo de un asesinato

3. ¿Cuáles son características de los eventos que se asocian más frecuentemente con el desarrollo de un TEPT?
 - a. Posibilidad latente de muerte
 - b. Decepción
 - c. Impotencia
 - d. Alegría intensiva

4. ¿Por cuánto tiempo debe una persona presentar síntomas de un TEPT antes de que pueda ser emitido un poder emitir un el diagnóstico?
 - a. 2 semanas
 - b. 1 mes
 - c. 3 meses
 - d. 1 año

5. ¿Quién puede desarrollar un TEPT?
 - a. Sólo hombres (Gender Difference)
 - b. Cualquiera que haya vivido suficientes experiencias traumáticas (Building-Block Effect)
 - c. Sólo los adultos (Youth-Protection Factor)

6. ¿Cuáles son los tres criterios diagnósticos principales para diagnosticar un TEPT?
 - a. Intrusiones
 - b. Estrés
 - c. Evitación
 - d. Hiperactivación

7. ¿Cuál es la característica que diferencia al TEPT de otros trastornos de ansiedad?
 - a. Falta del reconocimiento del miedo
 - b. La presencia del criterio A
 - c. El estrés crónico
 - d. Evitación de situaciones sociales

8. Según el DSM-V, ¿cómo se clasifica el criterio A?
 - a. Respuesta objetiva
 - b. Respuesta natural
 - c. Respuesta subjetiva
 - d. Respuesta directa

9. ¿Cuántos síntomas dentro del criterio B son necesarios para decretarlo como presente?
 - a. Uno
 - b. Dos
 - c. Tres
 - d. Cinco

Síntomas

10. ¿Cuáles características y/o síntomas son particularmente típicos en personas con TEPT?
 - a. Actuar o sentir como si el acontecimiento traumático volviese a ocurrir
 - b. Aumento de peso
 - c. Pérdida de autoestima
 - d. Problemas de sueño
 - e. Dificultades en la concentración
 - f. Disminución de la inteligencia
 - g. Recuerdos recurrentes y angustiosos
 - h. Euforia

Terapia

11. ¿Cuáles factores pueden influenciar positivamente el resultado de la terapia en un TEPT?
- a. Contacto físico cercano entre terapeuta y paciente
 - b. Creación de una atmósfera cercana
 - c. Presionar al paciente desde el principio de la terapia a describir detalladamente el evento traumático
 - d. Que el/la paciente se introduce nuevamente en la experiencia traumática y siente como si la estuviera viviendo nuevamente
 - e. Que el/la paciente recuerde la experiencia traumática (cuándo y dónde fue, al igual que cómo se sintió, siendo siempre consciente de que la situación real ya acabó

¡Gracias por su participación!

A3 Beurteilungsbogen DLE-NET

Cuestionario DLE para evaluar las sesiones de NET

Este cuestionario contiene cuatro áreas: **Diagnóstico**, **Línea de vida**, **Exposición** y **habilidades terapéuticas**. Primero elija la fase (D, L ó E) e identifique los items correspondientes incluyendo los items del área T. Por favor evalúe, si los elementos correspondientes fueron aplicados en la sesión de la siguiente manera: "excelente" (2), de "bien" (1) o "mal" (0). Si no puede valorar un item elija la opción "sin dato" y escriba una nota breve.

Fecha:

Fase: D L E

Nombre del/la terapeuta:

Fase	Item	0 mal	1 bien	2 excelente	X sin dato
D	1. Conocimiento de los instrumentos diagnósticos Visión global de los Items y sus constructos; no solamente leído	0	1	2	
	2. Entrevista diagnóstica Estilo de la entrevista: una conversación, no un interrogatorio	0	1	2	
L	3. Elementos de la línea de vida Todos los elementos presentes de forma suficiente y con variedad (tamaño, color, ...); explicados e implantados correctamente; al final de la sesión recogidos por el cliente. Foto con consentimiento del cliente.	0	1	2	
	4. Fases de la línea de vida Todos las fases representadas adecuada y correctamente (nacimiento, vida, futuro, final)	0	1	2	
	5. Nivel Ninguna exposición pero suficiente información de trasfondo sobre los los eventos narrados	0	1	2	
	6. Visión global Terapeuta tiene una visión global sobre la vida, eventos traumáticos y positivos del cliente (incluido apuntes)	0	1	2	
E	7. Narración Lectura de la narración o resumen de las últimas sesiones al comienzo de la sesión y en caso necesario completar la información adicional. Continuación de la narración a lo largo de la línea de la vida „Lifeline“, sin ningún desvío arbitrario de la cronología de vida	0	1	2	
	8. Piedra / Exposición Es un evento traumático (Criterio-A) y se encuentra en el foco central (cámara lenta); tiempo adecuado antes y después del evento; realmente tiene lugar una exposición al miedo	0	1	2	
	9. Integración contextual („memoria fría") Preguntar de manera suficiente la información de contexto (cuándo, dónde, quién, qué), de tal manera que sea posible la integración cronológica de las experiencias relatadas en la historia de vida del paciente y así se construya un conocimiento declarativo autobiográfico.	0	1	2	
	10. Activación de la "memoria caliente" Desarrollar todos los componentes de la "memoria caliente" hasta enmarcarlos en una exposición clara (pensamientos, sentimientos, fisiología, percepciones sensoriales, comportamiento)	0	1	2	
	11. La integración del punto 9 y 10 es clara	0	1	2	
	12. Relación con el presente Mantener la relación con el presente: comparar el „antes" y "ahora". Evitación de disociaciones y flashbacks y/o reconocimiento rápido de éstas posibles situaciones y la aplicación de las técnicas de "reorientación". Continuando con la narración después de consolidar la reorientación completa.	0	1	2	

	13. Prevención de la evitación Prevenir las „estrategias de evitación" durante la exposición (q. indicando antes de comenzar a usar el baño, a fumar, evitando "saltos" esporádicos en la narración, aplazando las apreciaciones, valoraciones y reconociones de los eventos pasados hasta el final de la sesión)	0	1	2	
	14. Final de la sesión Finalizar la sesión en un punto „relativamente" seguro. Se ha presentado una reducción clara del miedo y de la activación del paciente.	0	1	2	

T	15. Relación terapéutica Hay una relación terapéutica, marcada por la confianza, empatía, etc.	0	1	2	
	16. Comprensión Comprensión del mundo del paciente (pensamientos, sentimientos, sensaciones corporales, percepciones, comportamiento) (Ej. Por medio de preguntas concretas) y transmitir esto al paciente de forma empática, precisa y clara (Ej. Parafraseando, resumiendo, verbalizando pensamientos, comportamientos, sentimientos, sensaciones corporales y percepciones).	0	1	2	
	17. Eficacia en la interacción Aumentar la eficacia de la terapia mostrando atención, interés, cuidado, autenticidad y competencia profesional.	0	1	2	
	18. Interés Mantener un equilibrio entre estar interesado, calidez emocional y permisividad; en lugar de una actitud distante, no centrada en el paciente.	0	1	2	

<i>Suma de los puntos correspondientes</i>				
--	--	--	--	--