



universität  
wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Wirksamkeit von Psychotherapie in freier Praxis –  
Durchführbarkeit und erste Ergebnisse einer naturalistischen Studie  
der Gesellschaft für Logotherapie und Existenzanalyse“

verfasst von / submitted by

Katharina Haubmann BA BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

UA 066 840

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Masterstudium Psychologie UG2002

Betreut von / Supervisor:

Univ.-Prof. Dr. Anton-Rupert Laireiter

*meinen Eltern*

## Inhaltsverzeichnis

<b>THEORETISCHE HINFÜHRUNG .....</b>	<b>8</b>
<b>Begriffsbestimmung und Entwicklungen im Forschungsfeld .....</b>	<b>8</b>
Psychotherapie (in Österreich) .....	8
Psychotherapieforschung.....	9
Differenzierungen in der psychotherapeutischen Wirksamkeitsforschung .....	9
Phasen der psychotherapeutischen Wirksamkeitsforschung .....	10
<b>Evidenzbasierte Praxis.....</b>	<b>12</b>
<b>Praxisbasierte Evidenz.....</b>	<b>15</b>
Naturalistische Studiendesigns .....	16
Praxisorientierte Studien im deutschsprachigen Raum .....	19
Chancen und Herausforderungen praxisorientierter Studien .....	23
<b>Psychotherapeutische Wirksamkeitsforschung bei Ausbildungstherapien.....</b>	<b>25</b>
<b>Humanistische Psychotherapie .....</b>	<b>26</b>
Logotherapie und Existenzanalyse .....	29
Existenzanalyse .....	30
Wirksamkeit existenzanalytischer Psychotherapie.....	32
<b>Zusammenfassung &amp; Forschungsinteresse .....</b>	<b>35</b>
<b>Forschungsfragen &amp; Hypothesen .....</b>	<b>37</b>
<b>EMPIRISCHER TEIL.....</b>	<b>39</b>
<b>Studienbeschreibung.....</b>	<b>39</b>
Studienablauf .....	40
Erhebungsinstrumente .....	41
Erhebung demographischer und Hintergrunddaten.....	45
Qualitative Daten.....	45
<b>Datenanalyse .....</b>	<b>46</b>
Ermittlung der Studiendurchführbarkeit und -akzeptanz .....	46
Qualitative Inhaltsanalyse nach Kuckartz .....	47
Gruppenstatistische Veränderung.....	47
Veränderung auf individueller Ebene.....	51
Zufriedenheit mit der Psychotherapie .....	52

<b>Ergebnisse .....</b>	<b>53</b>
Teilnahmebereitschaft und Studienakzeptanz der Therapeut*innen.....	53
Beschreibung der Therapeut*innenstichprobe .....	53
Qualitative Analyse zur Studienteilnahme der Therapeut*innen .....	54
Teilnahmebereitschaft und Studienakzeptanz der Klient*innen.....	56
Qualitative Analyse zur Teilnahmebereitschaft der Klient*innen .....	57
Vorläufige Abbruchsanalyse und bisheriger Fragebogenrücklauf.....	58
Klient*innenbewertungen zum Erleben der Fragebögen.....	59
Beschreibung der Klient*innenstichproben .....	61
Gesamtstichprobe der Klient*innen.....	61
Soziodemographische Daten.....	62
Skalenmittelwerte.....	63
Anteil an Personen mit klinisch (un-)auffälliger Symptomatik.....	64
Diagnosen.....	65
Prä-Post-Stichprobe .....	68
Dauer der Therapien bzw. Anzahl der Einheiten .....	68
Gruppenstatistische Veränderung zwischen Therapiebeginn und -ende (T1-T3) .....	69
Individuelle Veränderung zwischen Therapiebeginn und Therapieende .....	70
Therapiebeurteilung der Klient*innen und Therapeut*innen .....	72
Kurzfristige Therapieeffekte (T1-T2).....	75
Langfristige Therapieeffekte (T1-T4).....	76
Darstellung und Zusammenfassung der Therapieeffekte .....	78
Explorative Analysen.....	79
Completer- und ITT-Analysen der GSI-Werte .....	79
Prädiktoren des Therapieergebnisses .....	79
Zusammenhang der ESK-Gesamtskala mit den anderen Skalen.....	80
Qualitative Ergänzungen - Verstehen durch Verstanden-werden.....	80
 <b>DISKUSSION .....</b>	 <b>81</b>
 <b>Ergebnisspezifische Diskussion.....</b>	 <b>81</b>
Therapeut*innenstichprobe .....	81
Berichte der Therapeut*innen zur Studienteilnahme .....	82
Klient*innenstichprobe.....	84
Fragebogenrücklauf und Abbruch .....	85

Studienakzeptanz .....	86
Soziodemographische Daten .....	87
Skalenmittelwerte und Anteil an Personen mit klinisch auffälliger Symptomatik.....	88
Diagnosen .....	89
Prä-Post-Stichprobe .....	89
Gruppenstatistische Veränderung zwischen Therapiebeginn und Therapieende (T1-T3) .....	91
Individuelle Veränderung zwischen Therapiebeginn und Therapieenden (T1-T3) .....	92
Therapiebeurteilungen der Klient*innen und Therapeut*innen.....	93
Kurz- und langfristige Veränderungen .....	94
Explorative Analysen .....	95
<b>Studienübergreifende Diskussion .....</b>	<b>97</b>
Verschiedene Veränderungsmaße als Therapiebeurteilung .....	97
Effektstärkenmaß .....	97
Maß der reliablen Veränderung und klinische Signifikanz .....	98
Stärken und Limitationen .....	99
<b>Konklusion und Ausblick .....</b>	<b>102</b>
<b>LITERATURVERZEICHNIS .....</b>	<b>107</b>
<b>ANHANG .....</b>	<b>119</b>
<b>Abkürzungsverzeichnis.....</b>	<b>119</b>
<b>Abbildungsverzeichnis .....</b>	<b>119</b>
<b>Tabellenverzeichnis .....</b>	<b>120</b>
<b>Zusammenfassung.....</b>	<b>121</b>
<b>Abstract.....</b>	<b>122</b>

## **Wirksamkeit von Psychotherapie in freier Praxis – Durchführbarkeit und erste Ergebnisse einer naturalistischen Studie der Gesellschaft für Logotherapie und Existenzanalyse**

Psychische Störungen wie Angst und Depression zählen mittlerweile zu den am häufigsten auftretenden Erkrankungen in Österreich. Mit dieser Gesundheitsbeeinträchtigung sind nicht nur hohe finanzielle Kosten für Betroffene und Angehörige sowie für die Gesellschaft, sondern gleichfalls eine besonders große Krankheitslast verbunden (Löffler-Stastka & Hochgerner, 2021; Riffer, Sprung, Streibl & Kaiser, 2018). Bei psychischen Störungen sind Psychotherapie und / oder psychopharmakologische Behandlung indiziert. In der Versorgung von Betroffenen ist jedoch eine Schieflage zugunsten psychopharmakologischer Therapien zu beobachten (Grabenhofer-Eggerth & Sator, 2020; Rieß & Löffler-Stastka, 2022; Riffer et al., 2018). Während im Jahr 2014 für psychopharmakologische Behandlungen 256 Millionen Euro aufgewandt wurden, beliefen sich die Ausgaben der Sozialversicherung für psychotherapeutische Behandlungen im selben Zeitraum auf 76 Millionen Euro (Grabenhofer-Eggerth & Sator, 2020; Löffler-Stastka & Hochgerner, 2021). Nur 5 % der Personen, die Psychopharmaka einnehmen, erhalten einen vollfinanzierten Psychotherapieplatz (Riffer et al., 2018). Dieses Ungleichgewicht besteht trotz der vor 30 Jahren in Kraft getretenen Novellierung des *Allgemeinen Sozialversicherungsgesetzes (ASVG)*, das die Psychotherapie den anderen ärztlichen Tätigkeiten gleichstellt und in den Pflichtleistungskatalog der sozialen Krankenversicherung aufgenommen hat (Grabenhofer-Eggerth & Sator, 2020; Riffer et al., 2018). Dafür sprechen auch die Ergebnisse zahlreicher Studien und Metaanalysen, aus denen hervorgeht, dass Psychotherapie bei psychischen Störungen mindestens genauso effektiv wie eine psychopharmakologische Behandlung ist, gleichzeitig aber weniger Nebenwirkungen und längerfristige Effekte aufweist (Datler, Hochgerner, Korunka, Löffler-Stastka & Pawlowsky, 2021; Wampold & Imel, 2015).

Als Goldstandard der psychotherapeutischen Wirksamkeitsforschung gelten randomisiert-kontrollierte Studien (RCT). Dabei handelt es sich um ein experimentelles Studiendesign, das aus der medizinisch-pharmakologischen Forschung übernommen wurde und intern-valide, kausale Schlussfolgerungen über Behandlungseffekte ermöglicht (Berg, 2021; Chambless & Ollendick, 2001; Eiling, Schlipfenbacher, Hörz-Sagstetter & Jacobi, 2014; Revenstorf, 2005). Die interne Validität wird durch die Einhaltung strenger methodischer Kriterien gewährleistet, die jedoch dazu führen, dass die externe Validität sinkt und die Aussagekraft von RCT-Studienergebnisse für andere psychotherapeutische Kontexte,

aber vor allem für jene in freier Praxis kaum gegeben ist (Lutz, Castonguay, Lambert & Barkham, 2021; Seligman, 1995). Das ist insofern relevant, als in Österreich rund die Hälfte der psychotherapeutischen Behandlungen von Psychotherapeut\*innen in freier Praxis durchgeführt werden (Löffler-Stastka & Hochgerner, 2021; Tanios, Valady & Grabenhofer-Eggerth, 2020). Aus diesem Grund wird ergänzend auch eine praxisorientierte Psychotherapieforschung gefordert (Castonguay, Youn, Xiao, Muran & Barber, 2015; Herzog, Kaiser & Brakemeier, 2022; Lutz et al., 2021). Die damit verbundenen naturalistischen Studiendesigns sind in der Lage psychotherapeutische Wirksamkeit in der Routineversorgung abzubilden und ermöglichen gleichzeitig eine Evaluation und Qualitätssicherung bestehender Versorgungsleistungen (Bohart, O'Hara & Leitner, 1998; Castonguay et al., 2015; Seligman, 1995).

In der psychotherapeutischen Wirksamkeitsforschung wurden bislang hauptsächlich Behandlungen ausgebildeter Psychotherapeut\*innen evaluiert. Dagegen gibt es nur wenig veröffentlichte Untersuchungen zur Wirksamkeit von Ausbildungstherapien an ambulanten Einrichtungen und nahezu keine von jenen in freier Praxis (Victor et al., 2018). Diese Forschungslücke wird der Versorgungsleistung der, allein in Österreich, rund 4500 Psychotherapeut\*innen in Ausbildung unter Supervision nicht gerecht (Löffler-Stastka & Hochgerner, 2021). Der evaluative Charakter praxisorientierter Psychotherapiestudien kann insbesondere im Ausbildungskontext genützt und als empirische Begleitung des Supervisionsprozesses genützt werden (Rubel et al., 2017).

Zudem können sich naturalistische Studiendesigns den Besonderheiten einzelner psychotherapeutischer Verfahren anpassen. Das ist insbesondere für humanistische Verfahren, wie die Logotherapie und Existenzanalyse, von Bedeutung (Bohart et al., 1998; Leichsenring et al., 2018; Vaughan, Marshall, MacKinnon & Roose, 1994). Die Logotherapie und Existenzanalyse wird weniger über manualisierte, störungs- und symptombezogene Techniken definiert, sondern als ganzheitliches und phänomenologisches Psychotherapieverfahren beschrieben, im Zuge dessen den Klient\*innen zu einer eigenverantwortlichen und authentischen Lebensgestaltung verholfen werden soll (Längle, 2014).

Das Ziel der vorliegenden Arbeit ist es, die Behandlungsqualität von existenzanalytischen Psychotherapeut\*innen in Ausbildung unter Supervision in freier Praxis anhand einer Zwischenanalyse einer laufenden Studie der Gesellschaft für Logotherapie und Existenzanalyse zu untersuchen. Zusätzlich soll die Studienakzeptanz und Anwendbarkeit des eingesetzten naturalistischen Untersuchungsdesigns evaluiert werden.

## Theoretische Hinführung

### Begriffsbestimmung und Entwicklungen im Forschungsfeld

#### *Psychotherapie (in Österreich)*

Für die Definition von Psychotherapie wird auf das Ergebnis ausführlicher Diskussionen zurückgegriffen, welches in die Formulierung im österreichischen Berufsgesetz mündete (Slunecko, 2017). Darin wird Psychotherapie als eine Behandlung von psychosozialen und -somatischen Störungen definiert, die durch eine gesetzlich geregelte Ausbildung erlernt wird und sich am neuesten Stand wissenschaftlicher psychotherapeutischer Forschung orientiert. Ziel einer psychotherapeutischen Behandlung ist es, Symptome zu lindern, krankmachende Verhaltensweisen zu verändern und Gesundheit, Reifung und persönliche Entwicklung der Betroffenen zu fördern (Psychotherapiegesetz, BGBl. 361/1990).

Im Gegensatz zu vielen anderen nationalen Gesundheitssystemen werden in Österreich 23 Verfahren als eigenständige Psychotherapieschulen gesetzlich anerkannt (Fäh, 2006; Sagerschnig & Valady, 2021). Diese können wiederum in vier große Grundorientierungen, sogenannte Cluster, zusammengefasst werden. Die Psychodynamische Orientierung vereint psychoanalytische sowie tiefenpsychologisch fundierte Verfahren, die Verhaltenstherapeutischen Verfahren ohne weitere Unterteilung, die Systemische Orientierung mit der Systemischen Familientherapie und der Neuro-Linguistischen Psychotherapie sowie die Humanistische Orientierung mit Verfahren, wie die Personzentrierte Psychotherapie, die Integrative Therapie und die Existenzanalyse und Logotherapie, um nur einige zu nennen (Sagerschnig & Valady, 2021).

Im internationalen Vergleich gilt die zweigeteilte Berufsausbildung von Psychotherapeut\*innen in Österreich als umfassend und qualitativ hochwertig (Datler et al., 2021). Der erste Teil, das *Propädeutikum*, soll eine allgemeine Einführung in die psychotherapeutische Theorie und Praxis geben. Beim *Fachspezifikum* handelt es sich um den zweiten Teil der Ausbildung. Neben 300 theoretischen Stunden wird mit 1600 Stunden Praxis der Fokus auf praktische Erfahrung gesetzt (Löffler-Stastka et al., 2018). Nach ungefähr der Hälfte der praktischen und theoretischen Ausbildung im Fachspezifikum erfolgt eine Zulassung zur eigenständigen Psychotherapie unter Supervision, die mit dem Status *Psychotherapeut\*in in Ausbildung unter Supervision* verbunden ist und Therapien auch in eigener Praxis ermöglicht. Diese Behandlungen müssen schriftlich aufgezeichnet und von entsprechend ausgebildeten Psychotherapeut\*innen supervidiert werden.

## **Psychotherapieforschung**

Thematisch kann psychotherapeutische Forschung in die zwei weit gefassten Kategorien, *Ergebnis-* und *Prozessforschung*, unterteilt werden (Caspar & Jacobi, 2004; Lutz et al., 2021). Während in *Prozessstudien (process-studies)* Wirkfaktoren und Behandlungsverläufe untersucht werden, und hauptsächlich der Forschungsfrage nachgegangen wird, *wie* Psychotherapie wirkt, werden in *Wirksamkeits-* oder *Ergebnisstudien (outcome-studies)* vorwiegend die Resultate einer Behandlung evaluiert, indem untersucht wird, *ob* und *wie sehr* Psychotherapie wirkt. (Lutz et al., 2021). Psychotherapiestudien können auch beiden Formen zugeordnet werden, zumal Prozessforschung ohne Ergebniserhebung nur wenig Aufschluss über die tatsächliche Wirkung von bestimmten Prozessen zulässt. Umgekehrt bedarf es neben der Erkenntnis, dass ein Verfahren wirkt, bestenfalls auch das Wissen, auf welche Weise diese Wirkung erzielt wird (Orlinsky, Grawe & Parks, 1994). Aus dieser Verwobenheit beider Forschungszugänge hat sich die *Prozess-Ergebnis-Forschung (process-outcome-research)* entwickelt, welche den Zusammenhang von Prozessfaktoren mit den Therapieergebnissen untersucht (Caspar & Jacobi, 2004).

Weil für die vorliegende Arbeit bedeutender wird vorwiegend auf die psychotherapeutischen *Wirksamkeitsforschung (outcome studies)* Bezug genommen.

**Differenzierungen in der psychotherapeutischen Wirksamkeitsforschung.** Im Kontext der psychotherapeutischen Wirksamkeitsforschung werden die Begriffe *potenzielle Effektivität (efficacy)*, *praktische Effektivität (effectiveness)* und seltener auch *Kosteneffektivität (efficiency)* unterschieden (Herzog et al., 2022; Jacobi & Reinecker, 2018). *Efficacy-Studien*, auch Laborstudien genannt, prüfen die Wirksamkeit von Psychotherapie unter methodisch möglichst optimalen Bedingungen mittels randomisiert-kontrollierter Untersuchungsdesigns (randomised controlled trial, RCT), welche zu den experimentellen Forschungsmethoden gezählt werden. RCT-Studien zeichnet eine hohe interne Validität aus, also eine hohe Aussagekraft der Ergebnisse für die Bedingungen innerhalb eines Studiendesigns. *Effectiveness-Studien* (Feldstudien) untersuchen die Wirksamkeit von Psychotherapie unter klinischen Praxisbedingungen und werden mit naturalistischer Psychotherapieforschung assoziiert (Heekerens, 2005; Leichsenring & Rüger, 2004; Lutz, 2003; Tschuschke & Freyberger, 2015). Feldstudien zeichnet eine hohe externe Validität aus, also eine hohe Aussagekraft der Ergebnisse für Bedingungen außerhalb des Studiendesigns, weswegen Befunde aus naturalistischen Studien meist eine größere Generalisierbarkeit und klinische Aussagekraft aufweisen (Heekerens, 2005; Shadish, Cook & Campbell, 2002). Während Laborstudien die potenzielle und Feldstudien die praktische Wirksamkeit einer

Intervention ermitteln sollen, ist bei *Efficiency-Studien* das Kosten bzw. Aufwand-Nutzen-Verhältnis von entscheidender Bedeutung (Jacobi & Reinecker, 2018; Seligman, 1995).

Die psychotherapeutische Wirksamkeitsforschung kann in Anlehnung an die pharmakologische Medikamentenprüfung zudem in vier Phasen eingeteilt werden (Jacobi, 2020). Während *Phase I* die Konzeptentwicklung und *Phase II* die weiterführende Exploration zum Inhalt haben, beschreiben *Phase III* die kontrollierte Wirksamkeitsprüfung durch RCT-Studien und *Phase IV* Untersuchungen von Behandlungen unter Routinebedingungen. Phase IV-Studien überprüfen die klinische Nützlichkeit in der Versorgungsrealität mittels naturalistischer Studiendesigns und entsprechen einer praxisorientierten Wirksamkeitsforschung (Herzog et al., 2022; Hiller, Schindler, Andor & Rist, 2011; Jacobi, 2020; Laireiter, Schaireiter, Schirl-Russeger, Baumgartner, Längle & Sauer, 2013).

Zu betonen ist, dass sich die methodische Diskussion dieser Arbeit auf das quantitative Spektrum der (Psychotherapie-) Wissenschaft beschränkt und qualitative Erhebungs- und Analyseformen nahezu unerwähnt bleiben. Diese stellen aber eine bedeutsame Methode, gerade auch in der Erforschung von Psychotherapie, dar (Gelo, Lagetto, Dinoi, Belfiore, Lombi, Blasi, Aria & Ciavolino, 2020).

**Phasen der psychotherapeutischen Wirksamkeitsforschung.** Der Phasenbegriff lässt sich auch in einigen Überblicksarbeiten zur (historischen) Entwicklung der psychotherapeutischen Wirksamkeitsforschung finden (Ardito & Rabellino, 2011; Eckert, 2013; Grawe, 1992; Löffler-Stastka, 2013; Orlinsky, Rønnestad & Willutzki, 2004). Phasen sind an dieser Stelle als sich gegenseitig beeinflussende, überlappende und historisch gewachsene Forschungstraditionen zu verstehen, welche bis heute andauern können und noch von mehr oder weniger großer Bedeutung sind (Ardito & Rabellino, 2011; Jacobi, 2020; Lutz et al., 2021).

Vor den 1950er Jahren bestand die Forschung zur Wirksamkeit psychotherapeutischer Behandlungen überwiegend aus unsystematischen Einzelfallstudien (Lutz et al., 2021). In Folge kam es zu einer weitreichenden Kritik von Hans Eysenck (1952), in der behauptet wurde, dass durch die bisherigen empirischen Wirksamkeitsuntersuchungen spontane Remissionen nicht ausgeschlossen bzw. Wirkungen darüber hinaus nicht nachgewiesen werden könnten, da die unsystematischen Fallanalysen keine validen Aussagen über die Wirksamkeit von Psychotherapie im Allgemeinen zulassen würden (Caspar & Jacobi, 2004; Eysenck, 1952; Heekerens, 2005; Jacobi, 2020; Löffler-Stastka, 2013). Daraufhin wurden vermehrt experimentellen Studiendesigns in Form von Kontrollgruppenuntersuchungen

eingesetzt, deren Ergebnisse die provokante Annahme Eysencks widerlegen sollten (Caspar & Jacobi, 2004). Diese Entwicklung ist heute als *Legitimations- bzw. Rechtfertigungsphase* bekannt (Jacobi, 2020). Eine Annäherung an naturwissenschaftlich fundierte Methoden der Medizin hielt Einzug und randomisiert-kontrollierte Studien gewannen auch in der Beurteilung der Wirksamkeit psychotherapeutischer Interventionen an Bedeutung (Grawe, 1997; Löffler-Stastka, 2013). Eysencks (1952) These konnte im Zuge zahlreicher Untersuchungen widerlegt werden (Eckert, 2013; Eysenck, 1952, Fäh, 2006).

Das vermehrte Forschungsaufkommen zur Wirksamkeit von Psychotherapie durch standardisierte Studiendesigns führte zu einer *Wettbewerbsphase*, in der versucht wurde die Überlegenheit einer Psychotherapiemethode gegenüber einer anderen nachzuweisen (Jacobi, 2020; Löffler-Stastka, 2013). In den 1960er-Jahren entstanden umfangreiche Übersichtsarbeiten in Form von Metaanalysen, die die Ergebnisse einzelner Studien, auch verfahrensübergreifend, miteinander vergleichbar machten (Eckert, 2013). Einige Autor\*innen dieser Metaanalysen kamen zu dem Schluss, dass sich die Wirksamkeit der verschiedenen Psychotherapieverfahren im Großen und Ganzen nicht signifikant voneinander unterschied. Diese Annahme wurde in der Literatur als das *Äquivalenzparadoxon* oder auch als das *Dodo-Vogel-Verdikt* bekannt (Caspar & Jacobi, 2004; Löffler-Stastka, 2013). Kritik am Forschungsaufkommen der Wettbewerbsphase bezog sich insbesondere auf die unspezifischen Untersuchungen der allgemeinen Wirksamkeit von Psychotherapieverfahren. Dagegen blieben Fragen nach der Effektivität konkreter psychotherapeutischer Methoden bzw. Techniken unbeantwortet, und auch mögliche Einflüsse von Therapiedauer oder Klient\*innen- bzw. Psychotherapeut\*innenvariablen auf das Therapieergebnis fanden wenig Berücksichtigung (Jacobi, 2020; Löffler-Stastka, 2013).

Das zunehmende Aufkommen von störungsspezifischen Fragestellungen und Wirksamkeitsuntersuchungen einzelner Interventionstechniken wird als *Verschreibungs- oder Indikationsphase* bezeichnet, in der die Theorien der Therapieverfahren zugunsten einer störungs- und technikspezifischen Perspektive in den Hintergrund zu treten begannen (Grawe, 1997; Jacobi, 2020). Neben der Effektivität einzelner Verfahren entwickelte sich in der sogenannten *Prozessforschungsphase* zunehmend das Bemühen den Einfluss psychotherapeutischer (Prozess-) Faktoren auf das Behandlungsergebnis zu untersuchen (Grawe, 1997; Jacobi, 2020; Jacobi & Reinecker, 2018).

Zwei weitere Phasen können einerseits mit der Etablierung von (Gold-) Standards in der psychotherapeutischen Praxis und Forschung und andererseits mit der wachsenden Bedeutung von praxisorientierten Studien in Zusammenhang gebracht werden (Castonguay,

Barkham, Youn & Page, 2021; Chambless & Ollendick, 2001; Orlinsky et al., 2004; Rudolf, 1998). Beide Forschungsschwerpunkte werden im Anschluss ausführlicher beschrieben.

### **Evidenzbasierte Praxis**

Von namhaften Forschungseinrichtungen - wie der *American Psychological Association (APA)* oder dem *Cochrane Institut* - werden Richtlinien zur Prüfung und Berichterstattung der Wirksamkeit von Psychotherapie vorgegeben (Chambless & Ollendick, 2001; Kriz, 2019; Leichsenring & Rieger, 2004; Rabenstein, 2020; Wampold, 2021). In den Leitlinien wird betont, dass insbesondere experimentelle Forschungsdesigns, also randomisiert-kontrollierte Studien (RCT-Studien), bzw. Metaanalysen selbiger eine methodisch valide Einschätzung zur Wirksamkeit einer Intervention zulassen würden und einen höheren Evidenzgrad als beispielsweise quasi-experimentelle Studiendesigns aufweisen (Chambless & Ollendick, 2001; Fydrich & Schneider, 2007; Persons & Silberschatz, 1998; Rabenstein, 2020). Diese Entwicklungen werden innerhalb der Phase der *Konsolidierung, Standardisierung und Elaboration* beschrieben (Orlinsky et al., 2004).

In RCT-Studien soll der Einfluss einer unabhängigen Variable (Intervention) auf eine abhängige Variable (Symptomausprägung) untersucht werden. Um ein möglichst intern valides Ergebnis zu erhalten, um so eine kausale Schlussfolgerung zu ermöglichen, müssen methodische Anforderungen erfüllt sein (Chambless & Ollendick, 2001). Wichtige Prinzipien sind die *Randomisierung*, also die zufällige Zuweisung der Studienteilnehmer\*innen zu den Untersuchungsbedingungen und der *Vergleich* zwischen einer Interventionsgruppe und einer weiteren Behandlungs- oder Wartelistengruppe (Kontrollgruppe) (Leichsenring & Rieger, 2004). Zudem sollen mögliche andere Einflussgrößen (Kontroll- bzw. Störvariablen) durch Ausschluss oder durch kontrollierter Gleichhaltung in allen Vergleichsgruppen für das Ergebnis irrelevant gemacht werden (Chambless & Ollendick, 2001). Weitere methodische Anforderungen sind die Behandlung nach *Manual*, welche bei Personen mit *einer Diagnose* über eine *fixierte Zeitspanne* durchgeführt werden sollen, um den Effekt einer zeitlich und inhaltlich genau definierten Intervention (Dosis) für ein konkretes Störungsbild untersuchen zu können (Seligman, 1995).

Die Etablierung experimenteller Methoden als (Gold-) Standard in der psychotherapeutischen Wirksamkeitsforschung führte zu einer Annäherung an das medizinische Modell, insbesondere an die inhaltsnahe Psychopharmakologie (Castonguay et al., 2021; Kriz, 2019). Kritiker\*innen weisen darauf hin, dass die Übernahme medizinischer (Forschungs-) Prämissen voraussetze, dass psychotherapeutische Behandlungen mit bio-

medizinischen Interventionen, wie der Verabreichung eines Medikamentes, gleichzusetzen seien. Das wird jedoch in Frage gestellt (Revenstorf, 2005; Wampold & Imel, 2015).

Diese Kritik wird unter anderem an der Relevanz des *Settings*, also den Rahmenbedingungen, in denen eine Intervention durchgeführt wird, deutlich. In der Medizin wird angenommen, dass unabhängig vom Setting (z.B. Medikamentengabe in einer Klinik oder in einer Praxis) die Dosis eines Pharmakons weitgehend gleich wirkt (Dick & Kringler, 2007; Leichsenring & Rüger, 2004; Wampold, 2021). Der größte Prädiktor in der pharmakologischen Wirksamkeit wird der *spezifischen Behandlungsform*, dem Medikament, zugeschrieben. Dazu passt die methodische Stärke von RCT-Studien, den potenziellen Behandlungseffekt spezifischer Interventionen in einem kontrollierten Umfeld zu untersuchen (Deaton & Cartwright, 2018; Strauss, 2019; Tschuschke & Freyberger, 2015; Wampold, 2021). Unspezifischen Wirkfaktoren, wie der Umgang der Ärzt\*innen mit Patient\*innen, das gegenseitige Vertrauen in eine Intervention etc., wird ein wesentlich geringerer Einfluss auf die (pharmakologische) Wirkvarianz zugesprochen, weswegen diese in der medizinischen Forschung eher peripher behandelt werden (Revenstorf, 2005).

Ergebnisse aus großen Metanalysen der Psychotherapieforschung legen jedoch gegenteilige Schlüsse nahe. Der wesentliche Teil der Wirksamkeit einer psychotherapeutischen Behandlung ist auf unspezifische bzw. kontextuelle Wirkfaktoren zurückzuführen, während spezifische Wirkfaktoren eine eher untergeordnete Rolle einnehmen (vgl. Tabelle 1) (Strauss, 2019; Wampold, 2021).

**Tabelle 1**

*Wirkung von Psychotherapie und Einfluss von kontextuellen bzw. spezifischen Faktoren (Strauss, 2019, S. 6)*

Faktor(-gruppe)	Variable	Studien	Patienten	ES
Wirkung von Psychotherapie	Psychotherapie	> 500	> 80.000	0,80
Kontextuelle Faktoren	Allianz	190	> 14.000	0,57
	Empathie	59	3.599	0,63
	Zielkonsens/Zusammenarbeit	15	1.302	0,72
	Wertschätzung	18	1.067	0,56
	Kongruenz/Echtheit	16	863	0,49
	Kulturelle Adaptation der Therapie	21	950	0,32
	Person des Psychotherapeuten	29+17	14.519	0,35–0,55
Spezifische Wirkfaktoren	Unterschiede zwischen Treatments	295	> 5.900	< 0,20
	Spezifische Techniken	30	871	0,01
	Adhärenz	28	1.334	0,04
	Kompetenz für störungsspezifische Therapie	633	633	0,14

Wie aus Tabelle 1 hervorgeht, weisen spezifische Wirkfaktoren, wie die konkrete Technik oder die Manualtreue (Adhärenz), einen deutlich geringeren Einfluss auf das psychotherapeutische Behandlungsergebnis (*Cohen's d* < 0.20) auf als allgemeine bzw. kontextuelle Faktoren, wie die Arbeitsbeziehung oder das Einfühlungsvermögen der Therapeut\*innen mit Effektstärken zwischen *Cohens' d* = 0.35 bis 0.57 (Pfammatter & Tschacher, 2012; Revenstorf, 2005; Strauss, 2019; Wampold, 2021). Daneben werden an den RCT-Studienbedingungen auch die fixierten Zeiteinheiten, die strengen Einschlusskriterien wie die Monosymptomatik, die Störungsspezifität und Symptomfokussierung und besonders auch die Behandlung nach Manual als für den psychotherapeutischen Prozess (in freier Praxis) untypisch kritisiert. Dieser wird eher als *selbstkorrigierend*, *ressourcenbezogen* und *zeitlich* bzw. *individuell flexibel* beschrieben wird (Castonguay, Barkham, Lutz & McAleavey, 2013; Legewie, 2015; Persons & Silberschatz, 1998; Seligman, 1995). Zudem sind gewisse Psychotherapieverfahren, wie jene des humanistischen Clusters, aufgrund ihrer theoretischen und praktischen Ausrichtung schwieriger mit konkreten Handlungsanweisungen im Sinne einer Technisierung der Behandlung zu vereinbaren (Bohart et al., 1998).

Diese Überlegungen und Befunde legen für die psychotherapeutische Wirksamkeitsforschung den paradoxen Schluss nahe, dass je methodisch korrekter und somit intern valider randomisiert-kontrollierte Studien durchgeführt werden, desto geringer ist die externe Validität dieser Ergebnisse. Das bedeutet, dass weniger Aussagen über psychotherapeutische Behandlungen in anderen Kontexten, insbesondere in freier Praxis, gemacht werden können (Barkham & Margison, 2007; Deaton & Cartwright, 2018; Desmet, 2013; Eiling et al., 2014; Hiller et al., 2011; Legewie, 2015; Seligman, 1995; Tschuschke & Freyberger, 2015). Ein wesentliches Ziel von Psychotherapieforschung soll immer auch die Qualitätssicherung und -verbesserung des psychotherapeutischen Handelns im klinischen Alltag sein (Hiller et al., 2011; Laireiter et al., 2013; Legewie, 2015; Thir & Batthyány, 2016). Aus diesem Grund wird für die Beurteilung der (praktischen) Wirksamkeit von Psychotherapieverfahren eine Zusammenschau von Forschungsbefunden des gesamten Methodenspektrums mit dessen jeweiligen Vor- und Nachteilen eingefordert, anstatt einen *Goldstandard* zu definieren (Frieden, 2017; Jensen & Corralejo, 2017; Legewie, 2015; Seligman, 1995; Tschuschke et al., 2009; Wampold, 2021).

In der psychotherapeutischen Wirksamkeitsforschung werden als Ergänzung zu den RCT-Studien zunehmend naturalistische Studiendesigns hervorgehoben, die sich den Bedingungen der Psychotherapie im tatsächlichen Versorgungssetting anpassen und dadurch extern valider sind (Bohart et al., 1998; Buchholz, 2007; Castonguay et al., 2015;

Leichsenring et al., 2019). Statt der Konzeption der *evidenzbasierten Praxis* soll die Idee der *praxisbasierten Evidenz* die psychotherapeutische Wirksamkeitsforschung anleiten (Barkham & Margison, 2007; Dick & Kringler, 2007; Löffler-Stastka & Datz, 2016; Lutz et al., 2021).

### **Praxisbasierte Evidenz**

Als praxisbasierte Evidenz gelten in der Psychotherapieforschung Ergebnisse aus naturalistischen Psychotherapiestudien (Barkham & Margison, 2007). Ein bedeutsamer Unterschied zwischen naturalistischen und experimentellen Untersuchungen liegt in der Art des Studiendesigns, das sich in naturalistischen Studien dem zu untersuchenden Phänomen – der Psychotherapie (in freier Praxis) - anpasst und nicht umgekehrt das Phänomen der Methode angepasst wird (Seligman, 1995; Strauss et al., 2015). Daten werden in jenem Kontext erhoben, über den sie Aufschluss geben sollen. Aus diesem Grund weisen Ergebnisse aus praxisorientierten Studien eine höhere Repräsentativität, also externe Validität, für das betreffende Feld auf als solche aus experimentellen Studien (Deaton & Cartwright, 2018; Hiller et al., 2011; Laireiter & Kopetzky, 2015; Leichsenring & Rüger, 2004; Löffler-Stastka & Datz, 2016; Seligman, 1995; Shadish, Matt, Navarro, & Phillips 2000).

In den letzten Jahren ist eine Zunahme an naturalistischen Praxisstudien in der psychotherapeutischen Wirksamkeitsforschung feststellbar (Hiller et al., 2011). Einige Autor\*innen erkennen in diesen Entwicklungen eine neue Forschungsphase, nämlich die der *Praxisorientierten Psychotherapieforschung* (Castonguay et al., 2013; Rudolf, 1998). Im Vordergrund steht die gegenseitige Ergänzung der unterschiedlichen methodischen Ansätze mit deren jeweiligen Vor- und Nachteilen (Barkham & Margison, 2007; Castonguay et al., 2021; Heckerens, 2005; Hiller et al., 2011; Tschuschke & Freyberger, 2015).

Im Zuge von praxisorientierten Studien wird überdies eine Gleichrangigkeit zwischen den Forschungsverantwortlichen und den psychotherapeutischen Praktiker\*innen als Studienbeteiligte angestrebt. Letztere gelten als wichtige Verbündete (*stakeholder*) bei der Durchführung naturalistischer Studien (Castonguay et al., 2015; Eiling et al., 2014; Figl & Laireiter, 2022; Strauss et al., 2015). Durch eine intensivere Einbindung von Psychotherapeut\*innen in die Studienbelange wird ein Austausch zwischen Forschung und Praxis gefördert und eine *Bottom-Up-Wissenschaftspraxis* betrieben. Dieses Vorgehen kann blinde Flecken und Unachtsamkeiten vonseiten der Forscher\*innen verhindern und deren Kompetenzen mit den Erfahrungen von Kliniker\*innen ergänzen. Gleichzeitig kann es auch für Praktiker\*innen von Vorteil sein sich aktiv mit der Forschung ihres Berufsfeldes auseinanderzusetzen. Untersuchungsergebnisse deuten darauf hin, dass eine Forschungsteilnahme der Psychotherapeut\*innen mit einer intensiveren Beschäftigung mit

den Forschungsinhalten ihres Faches einhergeht (Castonguay et al., 2013; Castonguay et al., 2015; Löffler-Stastka & Datz, 2016; Strauss, 2019). Aufgrund der Verwobenheit von Praxis und Forschung in naturalistischen Studiendesigns können Belange von Praktizierenden und Klient\*innen direkter erfasst und auch so die praktische Relevanz der Forschung erhöht werden (Deaton & Cartwright, 2018). Ein solches Vorgehen schafft auch die Möglichkeit die häufig diskutierte *Praxis-Forschungslücke (practice-science-gap)* zu verringern, die durch eine experimentelle *Top-Down-Forschungspraxis* eher vergrößert wird, da dabei methodologische Überlegungen klinischen Belangen vorgezogen werden (Bohart et al., 1998; Castonguay et al., 2015; Kriz, 2019; Legewie, 2015; Lutz et al., 2021; Seligman, 1995; Strauss et al., 2015).

### ***Naturalistische Studiendesigns***

Bislang gibt es noch keinen einheitlichen Kriterien- bzw. Methoden katalog, welcher als Orientierung bei der Konzeption naturalistischer Studien genutzt werden kann und anhand dessen wissenschaftliche Qualität gewahrt und Ergebnisse methodisch eingeordnet werden können. Allerdings gibt es mittlerweile durch Forschungsprojekte einige Beispiele und Empfehlungen wie Studiendesigns im naturalistischen Setting ausgestaltet sein können und dabei ist insgesamt eine größere Methodenvielfalt als bei experimentellen Untersuchungen feststellbar (Hiller et al., 2011; Leichsenring & Rieger, 2004; Seligman, 1995; Tschuschke et al., 2009).

Gerade zu Beginn der aufkeimenden Kritik an Laborstudien in der Psychotherapieforschung wurden retrospektive Studiendesigns als mögliche Ergänzung vorgeschlagen. Anhand des Klient\*innenbefindens und ihrer Zufriedenheit mit dem Therapieergebnis können Rückschlüsse über die Effektivität von Psychotherapie in der Praxis gezogen werden (Laireiter et al., 2013; Lutz et al., 2021; Seligman, 1995). Mittlerweile haben sich aber prospektive Verlaufsstudien für die Untersuchung der Wirksamkeit von Psychotherapie in freier Praxis durchgesetzt. Psychometrische Fragebögen bleiben dabei ein häufig eingesetztes Instrument, um Veränderungen abzubilden (Altmann, Thielemann & Strauss, 2016; Castonguay et al., 2015; Legewie, 2015; Strauss et al., 2015).

Die nachstehenden methodischen Empfehlungen, beziehen sich, wenn nicht anders gekennzeichnet, auf die Ausführungen von Hiller und Kolleg\*innen (2011) bzw. Leichsenring und Rieger (2004). Um methodische Qualität im naturalistischen Setting zu gewährleisten, fordern die Autor\*innen ähnliche, aber angepasste methodische Richtlinien wie es sie für Laborstudien gibt. Während Leichsenring und Rieger (2004) eher allgemein und wissenschaftstheoretisch bleiben, beschreiben Hiller und Kolleg\*innen (2011)

anwendungsorientierte Erhebungs- und Auswertungsmethoden. Letztere forschen jedoch im Kontext von Universitätsambulanzen, ein Bereich, der wiederum andere kontextuelle Rahmenbedingungen als Psychotherapie in freier Praxis aufweist.

Eine naturalistische Studie mit hoher methodischer Qualität charakterisiert sich durch ein prospektives Design, das eine möglichst offene Stichprobengenerierung und zeitlich bzw. inhaltlich freie Therapieverläufe zulässt, um auf diese Weise klinisch repräsentative Daten erheben zu können (Legewie, 2015). Den Rahmenbedingungen von Psychotherapie in freier Praxis entsprechend, ist keine Randomisierung zu unterschiedlichen Untersuchungsbedingungen möglich, aber auch nicht nötig, da naturalistische Studien den Versorgungsalltag abbilden sollen.

Eine natürliche Kontrollgruppe wäre im Zuge von Wartelisten an Ambulanzen umsetzbar. In freier Praxis ist diese Option jedoch weniger gegeben. Durch das Fehlen einer randomisierten Kontrollgruppe können - streng genommen - etwaige Veränderungen nicht direkt auf eine Intervention zurückgeführt werden, da viele unkontrollierte Variablen systematische Verzerrungen der Ergebnisse hervorrufen können. Methodisch können aber sowohl das Fehlen von Kontrollgruppen als auch die fehlende Randomisierung ausgeglichen werden. So können etwa Ergebnisse von abgeschlossenen Therapien mit jenen von abgebrochenen Therapien verglichen werden, ohne dass dabei eine gesonderte Kontrollgruppe erhoben werden muss. Durch statistische bzw. theoretische Balancierungsmaßnahmen, wie die *Stratifizierung* (Randomisierung nach der Datenerhebung), die *Matching-Pairs-Technik* (Aufnahme von Personen mit möglichst ähnlichen Ausprägungen in den Einflussvariablen) oder die manuelle Kontrolle bekannter Einflussgrößen, kann eine fehlende Randomisierung nachträglich einigermaßen ausgeglichen und eine Homogenisierung von Vergleichsgruppen erwirkt werden (Heckerens, 2005; Leichsenring & Rüger, 2004; Shadish et al., 2000).

Da Psychotherapie in freier Praxis häufig ohne Manuale durchgeführt wird, kann stattdessen von den Therapeut\*innen eine Beschreibung der durchgeführten Behandlungsinhalte abgefragt werden (Bohart et al., 1998; Chambless & Ollendick, 2001). Alternativ können die durchgeführten, therapeutischen Techniken mittels Ton- oder Videoaufzeichnung ermittelt werden. Es gibt jedoch Hinweise, dass diese Form der Überprüfung der Techniktreue (*adherence*) ergebnisverzerrend wirkt und den Therapieverlauf beeinflussen kann. Zudem äußern Psychotherapeut\*innen häufig Vorbehalte gegenüber Ton- und Videoaufzeichnung von Therapieeinheiten (Castonguay et al., 2015; Figl & Laireiter, 2022; Seligman, 1995).

Anhand erheblich weniger Einschlusskriterien als bei Laboruntersuchungen und der größeren Stichprobenvariabilität in naturalistischen Studien sind eine transparente Berichterstattung der Datenerhebung und eine genaue Stichprobenbeschreibung von großer Bedeutung. Trotz heterogener Gegebenheiten soll auf diese Weise ein nachvollziehbarer Überblick der Stichprobenzusammensetzung gewährleistet werden (Legewie, 2015). Diagnosen sollten möglichst objektiv durch Verwendung von klinischen Interviews entsprechend dem aktuellen Klassifikationssystem der *International Classification of Diseases* (ICD) oder des *Diagnostic and Statistical Manual of Mental Disorders* (DSM) zumindest zu Therapiebeginn erhoben werden. Eine klinische Einordnung der Stichprobe kann auch durch den Einsatz von Fragebögen erfolgen. Valide und reliable Erhebungsinstrumente können zuverlässige Daten erheben, deren Ergebnisse auch eine Vergleichbarkeit mit jenen aus RCT-Studien zulassen. Die Annahme, dass Ergebnisse aus naturalistischen Studien zu systematischen Überschätzungen von Therapieeffekten, z.B. durch den *Selbstselektionsbias* (self-selection-bias), führen, konnte durch Vergleiche mit Ergebnissen aus Labor- und Feldstudien entkräftet werden (Heekerens, 2005; Leichsenring & Rieger, 2004; Minami, Wampold, Serlin, Hamilton, Brown & Kircher, 2008; Shadish et al., 2000).

Einen deutlichen Unterschied zwischen experimentellen und naturalistischen Psychotherapiestudien können die Erhebungszeiträume ausmachen. Während in RCT-Studien die Aufnahme von Teilnehmer\*innen und die Durchführung der untersuchten Therapien meist synchron ablaufen, erfolgt in naturalistischen Studiendesigns die Rekrutierung von Teilnehmer\*innen über mehrere Monate, sogar Jahre hinweg. Die untersuchten Therapien haben, je nach Erscheinen der Klient\*innen in der Praxis, unterschiedliche Beginn- und Endzeiten und weisen je nach Studiendesign in den meisten Fällen auch verschiedene Zeitumfänge auf. Naturalistische Studien können somit beträchtliche Erhebungszeiträume umfassen und bieten ähnlich wie eine (offene) Psychotherapie wenig zeitliche Orientierung. Hohe Abbruchquoten sind über die lange Zeitperiode einer Studie ein häufiges Problem naturalistischer Forschungsprojekte (Frieden, 2017).

Dokumentation und Berichterstattung über Abbruchs- bzw. Rücklaufquoten sollten immer angeführt und mit jenen anderer Studien (sofern vorhanden) verglichen werden. Abbrüche, also nicht reguläre Therapieabschlüsse, können in *qualitätsrelevante* bzw. *nicht-qualitätsrelevante* Abbrüche unterteilt werden. Von Hiller et al. (2011) werden als Gründe für qualitätsrelevante Abbrüche z.B. Unzufriedenheit der Klient\*innen mit dem Setting, der Arbeitsweise oder auch mit den Therapeut\*innen angeführt (Hiller et al., 2011). Diese sollten

bei der Evaluation der Behandlungsqualität mitbedacht werden, da sie mit hoher Wahrscheinlichkeit mit dem therapeutischen Prozess in Zusammenhang stehen (ebda.). Nichtqualitätsrelevante Abbrüche sind aus therapeutischer Sicht nicht. Gründe können neu auftretende Indikationen wie eine stationäre Behandlung, der Beginn eines Rehabilitationsaufenthaltes, eine Bewilligung eines vollfinanzierten Therapieplatzes, ein Wohnortwechsel, der Tod des\*r Klient\*in, Schwangerschaft, organisatorische Probleme etc. sein (Altmann et al., 2016; Hiller et al., 2011; Strauss et al., 2015).

Vor dem Hintergrund abgebrochener Fälle (Dropouts) hat sich das Berichten und Bearbeiten zweier Stichproben durchgesetzt: einerseits die Stichprobe jener Personen, die die Studie *und* die Therapie zumindest bis zu Therapieende durchgeführt haben (*Completer-Analyse, C*) und andererseits eine Gesamtstichprobe mit allen Personen, die die Studie begonnen haben (*Intention-to-Treat-Analyse, ITT*). Eine Methode der Imputation für die fehlenden Testzeitpunkte bei Abbruchsfällen ist die *Last-Observation-Carried-Forward* - Technik (*LOCF*) (Hiller et al., 2011; Steffanowski et al., 2011; Strauss et al., 2015). Dabei werden die Daten des zuletzt vorhandenen Testzeitpunktes einer Person für deren fehlende Daten zu Therapieende eingesetzt. Personen, die nur Daten zu Therapiebeginn haben, weisen somit keinen Behandlungseffekt auf (Steffanowski et al., 2011).

Eine weitere Analyseform, die in klinischen Studien häufig angewendet wird, ist die *Complete-Case-Analyse (CC)*, bei der nur jene Personen in die Analyse eingeschlossen werden, die vollständige Datensätze zu den jeweiligen Messzeitpunkten aufweisen (Jakobsen, Gluud, Wetterslev & Winkel, 2017). Eine ähnliche Auswertungsform ist die *Available-Case-Analyse (AC)*, bei der alle vorhandenen (Skalen-) Werte aufgenommen werden, und dadurch unterschiedliche Stichprobengrößen pro Skala vorhanden sein können (Little & Rubin, 2020). Diese Analyse-Stichproben können, je nach Studie, auch Personen mit abgebrochenen Therapien beinhalten, die dennoch bereit waren, an der Erhebung zu Therapieende und gegebenenfalls zum Katamnesezeitpunkt teilzunehmen (Cramer, Koemeda, Tschuschke, Schulthess & von Wyl, 2014; Gallas, Kächele, Kraft, Kordy & Puschner 2008).

### ***Praxisorientierte Psychotherapieforschung im deutschsprachigen Raum***

Während in angloamerikanischen Versorgungssystemen psychotherapeutischen Behandlungen eine durchschnittliche Dauer von weniger als 10 Einheiten zugestanden bekommen, können ambulante Psychotherapien im deutschsprachigen Raum einen weit höheren zeitlichen Umfang aufweisen (Fäh, 2006; Gallas, Puschner, Kühn & Kordy, 2010). Aus diesem Grund werden im Anschluss groß angelegte, naturalistische Studienprojekte ausschließlich aus dem deutschsprachigen Raum vorgestellt, deren Ergebnisse als

Vergleichsrahmen und Orientierung zur Interpretation der Ergebnisse dieser Arbeit dienen (*benchmarking*) (Hiller et al., 2011; Minami et al., 2008; Seligman, 1995; Shadish et al., 2000).

**QS-PSY-BAY-Studie – Qualitätssicherung ambulanter Psychotherapie (Steffanowski, Kramer, Fembacher, Glahn, Bruckmayer, von Heymann, Kriz, Pfaffinger & Wittmann, 2011).** Diese praxisorientierte Studie wurde im Raum Bayern in Deutschland als prospektive Ein-Gruppen-Längsschnittstudie konzipiert (Altmann et al., 2016). In diesem Projekt wurde ein Feedbacksystem in der Psychotherapie in freier Praxis untersucht, und es kam durch eine Kooperation zwischen Krankenversicherungsanstalten und Psychotherapeut\*innen zustande. Mehrere unabhängige Institute waren jeweils mit der Planung, der Erhebung und der Datenaufbereitung bzw. -analyse betraut. Die Datenerhebung dauerte knapp über zwei Jahre (2007-2009). Eine Studienteilnahme haben 217 Psychotherapeut\*innen und 1696 Klient\*innen begonnen. Die angegebenen Abbruchquoten lagen bei 34 % zu Therapieende und bei 52 % zum Katamnesezeitpunkt 12 Monate nach Therapieende. In einer (vorläufigen) Prä-Post-Completer-Stichprobe lagen Daten von 18 % der Anfangsstichprobe vor. Erhoben wurde zu Therapiebeginn, zu einer Verlaufsmessung nach ca. zehn Einheiten, zu Therapieende und zwölf Monate nach Therapieende (Steffanowski et al., 2011). Es konnten Prä-Post-Effektstärken einer Completer- und einer Intention-To-Treat-Stichprobe zwischen Cohens'  $d_C = 1.10$  ( $n = 309$ ) und Cohens'  $d_{ITT} = 0.74$  ( $n = 1696$ ) ermittelt werden. Bei der Katamneseerhebung ein Jahr nach Therapieende konnten Effektstärken von Cohens'  $d_K = 1.10$  ( $n = 309$ ) gefunden werden, was für ein stabiles Therapieergebnis spricht (Steffanowski et al., 2011; Strauss et al., 2015).

Ohne Vergleichsgruppe konnten keine Aussagen darüber getroffen werden, ob der Erhalt von Rückmeldungen einen Einfluss auf das Therapieergebnis hatte. Fragen zu Therapieende bezüglich der Studienakzeptanz weisen auf eine gute Annahme des Studiendesigns vonseiten der Klient\*innen hin. Auch das Rückmeldesystem wurde als sinnvoll und hilfreich wahrgenommen und die Mehrheit der Befragten würde sich Rückmeldungen bei allen Psychotherapien wünschen. Als nachteilig wurde der Einfluss der psychometrischen Rückmeldungen zu den Fragebogenergebnisse beschrieben, die darüber entschieden, ob eine Therapie indiziert war. Die teilnehmenden Psychotherapeut\*innen empfanden dieses Vorgehen als Eingriff in ihre professionelle Autonomie. Ergebnisse psychometrischer Testverfahren sollten als Ergänzung des therapeutischen Eindruckes genützt werden, diesen aber nicht ersetzen (Altmann et al., 2016; Steffanowski et al., 2011; Strauss et al., 2015).

**TRANS-OP-Studie – Mit Transparenz und Ergebnisorientierung zur Optimierung der psychotherapeutischen Versorgung: Eine Studie zur Evaluation ambulanter Psychotherapie (Gallas, Kächele, Kraft, Kordy & Puschner 2008).** In der TRANS-OP-Studie wurde der Fokus auf den zeitlichen Umfang von ambulanter Psychotherapie gelegt. Um diese Thematik näher zu beleuchten, wurde eine Stichprobe von 714 privat versicherten Psychotherapieklient\*innen anhand eines prospektiven Ein-Gruppen-Längsschnittdesigns erhoben (Altmann et al., 2016). Dafür wurden alle Klient\*innen, die in den Jahren von 1998 bis 2000 um eine Kostenübernahme von Psychotherapiebehandlungen bei einer deutschen Krankenkasse angesucht hatten, gebeten bei der Studie mitzumachen. Daten wurden zum Zeitpunkt der Antragstellung und eineinhalb bzw. zwei Jahre nach Antragstellung erhoben.

Im Unterschied zu den anderen Studien wurden bei der TRANS-OP-Studie zuerst die Klient\*innen rekrutiert und erst in späterer Folge auch deren Therapeut\*innen eine Teilnahme an der Studie angeboten. Die Gesamtstichprobe der Klient\*innen umfasste zu einem großen Teil männliche Teilnehmer mit höheren Bildungsabschlüssen. Die für Psychotherapieklient\*innen ungewöhnliche Verteilung wird durch das Klientel der projektinitiierenden Privatversicherung erklärt, welche vorrangig aus bildungsnahen, männlichen Personen bestand (Altmann et al., 2016). Entgegen der gängigen Praxis wurden keine Gruppenmittelwertsunterschiede und Effektstärken berechnet. Stattdessen wurden die Dauer und der Umfang von ambulanter Psychotherapie und deren Ergebnisqualität auf individueller Ebene durch das Maß der *reliablen* bzw. der *klinisch relevanten* Veränderung ermittelt.

Dabei erreichten 62 % aller Klient\*innen einer Available bzw. Complete-Case-Stichprobe ( $n = 542$ ) (klinisch unauffällig und klinisch auffällig) eine reliable Verbesserung ihrer psychischen Beschwerden. Eine klinisch signifikante Veränderung konnte bei 36 % aller Personen mit klinisch auffälliger Symptomatik zu Therapiebeginn nachgewiesen werden (Altmann et al., 2016; Gallas et al., 2008).

**PAP-S-Studie – Praxisstudie ambulante Psychotherapie Schweiz (Tschuschke, Cramer, Koemeda, Schulthess & von Wyl, 2015).** Bei der PAP-S-Studie handelt es sich ebenfalls um eine prospektive Ein-Gruppen-Längsschnittstudie. Es wurden alle 22 zugelassenen psychotherapeutischen Verfahrenseinrichtungen in der Schweiz gebeten, ihre Psychotherapeut\*innen in freier Praxis zur Studie einzuladen. Im Zeitraum von 2007 bis 2012 nahmen 86 Psychotherapeut\*innen von neun verschiedenen Psychotherapieinstituten an der Studie teil. Die Therapeut\*innen wurden ersucht, alle neuen Klient\*innen in freier Praxis über

die Möglichkeit einer freiwilligen Studienteilnahme aufzuklären. Rund 23 % der angesprochenen Personen ( $n = 1660$ ) haben einer Studienteilnahme zugesagt ( $n = 362$ ). Daten wurden jeweils zu Therapiebeginn, nach jeder fünften Sitzung, zu Therapieende und ein Jahr nach Therapieende mittels Fragebögen erhoben. Alle Datenerhebungen, außer die regelmäßigen Verlaufsmessungen nach jeweils fünf Einheiten, wurden in einem externen Assessmentzentrum durchgeführt. Zu Therapieende erklärten sich 73 % aller Studienanfänger dazu bereit, eine Testung durchzuführen. Ein Jahr nach Therapieende waren 40 % aller Studienanfänger\*innen bereit, in das Assessment-Zentrum für die Katamnesemessung zu kommen. Zusätzlich wurden die Therapeut\*innen gebeten – sofern sich die Klient\*innen einverstanden erklärt hatten – von jeder Sitzung eine Tonaufzeichnung zu machen und einen Interventionsbogen auszufüllen, um die Methodentreue (*adherence*) zu erheben.

Bei einer Completer-Stichprobe ( $n = 264$ ) konnten Prä-Post Effektstärken zwischen Cohens'  $d_C = 0.78$  und  $0.99$  ermittelt werden. Ein Jahr nach Behandlungsende betrugen die Effektstärken bei den verbliebenen Klient\*innen ( $n = 146$ ) zwischen Cohens'  $d_K = 1.00$  und  $1.12$  und hatten sich somit erhöht. Einen signifikanten Einfluss auf das Therapieergebnis hatten die psychische Belastungsintensität der Klient\*innen zu Therapiebeginn und das therapeutische Arbeitsbündnis. Die Konzepttreue (*adherence*) hatte dagegen keinen signifikanten Einfluss auf das Behandlungsergebnis, wobei Konzepttreue bei weniger erfahrenen Therapeut\*innen ausgeprägter vorhanden war. Auch die Therapiedauer hatte keinen wesentlichen Einfluss auf das Ergebnis (Altmann et al., 2016; Cramer et al., 2014; Tschuschke et al., 2015).

**TK-Studie – Modellprojekt der Technikerkasse zum Qualitätsmonitoring in der ambulanten Psychotherapie (Wittmann, Lutz, Steffanowski, Kriz, Glahn, Völkle, et al., 2011).** Das Modellprojekt der Technikerkasse zum Qualitätsmonitoring in der ambulanten Psychotherapie umfasst ein clusterrandomisiertes Zwei-Gruppen-Längsschnittdesign. Die TK-Studie unterscheidet sich damit von den anderen vorgestellten Untersuchungen hinsichtlich der Zufallszuteilung der Therapeut\*innen zu den Studienbedingungen. Es sollte untersucht werden, ob Rückmeldungen zu den psychometrischen Fragebogenergebnissen der Klient\*innen an deren Psychotherapeut\*innen die Ergebnisqualität der Psychotherapien verbessern. Dafür wurden die Therapeut\*innen ( $n = 403$ ) den beiden Untersuchungsgruppen (Erhalten von Rückmeldungen bzw. keine Rückmeldung) randomisiert zugeteilt. Die Studie wurde als Kooperation aus Forschenden, Therapeut\*innen und Beteiligten der Techniker Krankenkasse realisiert. Der Erhebungszeitraum lag zwischen den Jahren 2005 und 2010.

Es konnten keine signifikanten Unterschiede zwischen den beiden Untersuchungsbedingungen hinsichtlich der Therapieeffekte festgestellt werden. Zudem wurde keine Verringerung der Therapiedauer durch die Bereitstellung von Rückmeldungen erzielt, was besonders für die Krankenkasse von Interesse gewesen wäre. Die gemittelte Prä-Post-Effektstärke einer *Complete-case*-Stichprobe über beide Untersuchungsbedingungen hinweg belief sich auf Cohens'  $d_{CC} = 1.20$ . Die *bereinigte* Effektstärke ergab einen Wert von Cohens'  $d = 1.00$ , dafür wurde die Effektstärke eines ermittelten signifikanten Selektionseffektes abgezogen. Zu Therapieende lagen verwertbare Daten von 594 Personen (37 % aller Studienanfänger\*innen) vor. Es konnte eine Abbruchquote von 58 % zu Therapieende und von 71 % zum Katamnesezeitpunkt ermittelt werden. Aufgrund dieser hohen Abbruchquote wurden keine weiterführenden Analysen durchgeführt. Die Studienautor\*innen beschrieben eine gute Annahme und Durchführbarkeit des Rückmeldungssystems in der psychotherapeutischen Routineversorgung. Jedoch wurde nicht erhoben, ob und wie die Therapeut\*innen die Rückmeldungen verwendet hatten und wie sehr der therapeutische Eindruck mit den Ergebnissen übereinstimmte. Daneben wurden die Klient\*innen auch zu ihrer Therapiezufriedenheit befragt. Drei Viertel der Klient\*innen gaben zu Therapieende an, mit den Therapieergebnissen (sehr) zufrieden zu sein (Altmann et al., 2016; Strauss et al., 2015; Wittmann & Steffanowski, 2011).

Die vorgestellten Studien wiesen somit bei insgesamt 2438 Studienteilnehmer\*innen eine durchschnittliche Effektstärke um Cohens'  $d_{Prü-Post} \approx 1.00$  auf (Altmann et al., 2016). Alle Stichproben wurden störungsheterogen erhoben und entgegen den Empfehlungen von Hiller und Kolleg\*innen (2011) wurden alle - und nicht nur pathologische Personen - in die Analysen eingeschlossen (Hiller et al., 2011).

### ***Chancen und Herausforderungen praxisorientierter Studien***

Die vorgestellten Studien zeigen neben Chancen auch Herausforderungen, die mit naturalistischen Studiendesigns verbunden sein können. Naturalistische Studien sind sehr zeit- und ressourcenintensiv. Durchführende Forscher\*innen berichten von ungefähr zehn Jahren Planungs-, Durchführungs- und Auswertungs- bzw. Veröffentlichungszeit (Altmann et al., 2016; Steffanowski et al., 2011; Strauss et al., 2015). Die einzelnen Psychotherapien können sich über einen mehrjährigen Zeitraum erstrecken, und die untersuchten Behandlungen haben alle unterschiedliche Beginn- und Endzeitpunkte. Bei der Möglichkeit einer Fragebogenrückmeldung bedarf es außerdem einer kontinuierlichen Datenaufbereitung und -auswertung. Dadurch entsteht ein evaluativer Studiencharakter bereits während der Datenerhebung und nicht erst bei der finalen Analyse. Aufgrund der komplexen

Datenerhebung, -aufbereitung und -auswertung sind genaue Planung und Erprobung, z.B. im Zuge von Pilotstudien, wichtig für das Gelingen naturalistischer Studien (Cope, 2015; Donald, 2018; Gadke, Kratochwill & Gettinger, 2021; Tickle-Degnen, 2013; Whitehead, Sully & Campbell, 2014).

Die Durchführbarkeit eines naturalistischen Studiendesigns hängt neben einer guten Planung auch von teilnahmebereiten Therapeut\*innen ab. Ein derartig zeitlich breitangelegtes Vorgehen ist nur durch eine enge und langfristige Zusammenarbeit zwischen Forschungsteam und Psychotherapeut\*innen realisierbar (Laireiter & Kopetzky, 2015). Therapeut\*innen sind dabei gleichzeitig Studienteilnehmer\*innen und -durchführende. In naturalistischen Psychotherapiestudien kommt ihnen häufig die Aufgabe zu, den Klient\*innen eine Studienteilnahme anzubieten, gegebenenfalls klinische Interviews durchzuführen, Testbögen auszuteilen und die ausgefüllten Fragebögen an die Forschungsdurchführenden zu übermitteln. Geringe Schnittmengen und Berührungspunkte zwischen Praxis und Forschung sowie Vorbehalte von Psychotherapeut\*innen gegenüber einer Studienteilnahme (*practice-science-gap*) können das Durchführen einer Psychotherapiestudie erheblich erschweren (Castonguay et al., 2013; Strauss et al., 2015).

In der Literatur wurden als Gründe für eine geringe Studienakzeptanz seitens der Therapeut\*innen Befürchtungen einer Kontrolle der therapeutischen Arbeit, eines möglichen Autonomieverlusts oder einer Störung des psychotherapeutischen Prozesses durch die psychometrischen Erhebungen ermittelt (Strauss et al., 2015). Zusätzlich wurden auch fehlende Zeitressourcen, ein gewisses Misstrauen gegenüber psychometrischen Erhebungs- und Auswertungsformen und damit einhergehende datenschutzrechtliche Bedenken angegeben (Altmann et al., 2016; Figl & Laireiter, 2022; Steffanowski et al., 2011; Strauss, 2019; Taubner, Klasen & Munder, 2016).

Mit einem von Beginn an offenen Kommunikationsstil kann ein Vertrauensverhältnis zwischen Forschungsdurchführenden und -teilnehmenden geschaffen werden. Idealerweise soll der Informationsfluss wechselseitig und auf gleicher Ebene erfolgen. Transparenz, Kontinuität und eine fortwährende Kontaktmöglichkeit werden dabei als essentiell bezeichnet (Castonguay et al., 2015; Strauss et al., 2015). Die Studienanforderungen sollen für Therapeut\*innen und Klient\*innen zumutbar sein. Eine Studienteilnahme soll zeitlich und inhaltlich gut mit dem therapeutischen Alltag vereinbar sein. Für eine rege Studienbeteiligung förderlich sind außerdem Anreize im Zuge einer Teilnahme, wie beispielsweise die Anerkennung von Aus- und Weiterbildung. Für ein aussichtsreiches, naturalistisches

Forschungsvorhaben ist Motivationsarbeit bei den teilnehmenden Psychotherapeut\*innen unerlässlich (Figl & Laireiter, 2022; Strauss et al., 2015; Taubner et al., 2016).

Trotz der Schwierigkeiten, die naturalistische Studiendesigns mit sich bringen können, gibt es auch zahlreiche Vorteile. Infolge der Verbindung von Elementen aus Prozess- bzw. Ergebnisforschung und des inhaltlich und methodisch offeneren Studiendesigns können vielfältige Themen untersucht werden. Dosis-/Wirkungszusammenhänge, die mögliche Kostenreduktion im Gesundheitssystem durch Psychotherapie, Prädiktoren für Therapieeffekte, wie z.B. der Einfluss der Arbeitsbeziehung zwischen Therapeut\*innen und Klient\*innen, oder der Erhalt von Rückmeldungen sind nur einige Beispiele, die durch praxisorientierte Studien untersucht werden können (Altmann et al., 2016; Castonguay et al., 2015; Hiller et al., 2011; Löffler-Stastka & Datz, 2016; Strauss et al., 2015).

Ein großer Vorteil ist zudem der klinische Nutzen von Daten aus naturalistischen Studien. Durch eine Datenerhebung im Kontext des psychotherapeutischen Versorgungsalltags weisen Befunde aus praxisorientierten Untersuchungen eine größere Repräsentativität für die tägliche Praxis auf. Belange der Therapeut\*innen und der Klient\*innen können direkter erfasst und untersucht werden. Gleichzeitig werden Qualitätssicherung und Evaluierung der täglichen Praxis ermöglicht, was eine wichtige Funktion der Psychotherapieforschung darstellt (Deaton & Cartwright, 2018; Laireiter et al., 2013; Seligman, 1995; Thir & Batthyány, 2016). Durch das unmittelbare Rückmelden der Fragebogenergebnisse entsteht eine zusätzliche Informationsquelle zum Zustand der Klient\*innen, was insbesondere bei Ausbildungstherapien hilfreich sein kann (Rubel et al., 2017).

### **Psychotherapeutische Wirksamkeitsforschung bei Ausbildungstherapien**

Die Qualitätssicherung von Ausbildungstherapien wird in der psychotherapeutischen Forschungsgemeinschaft als ein zunehmend wichtiges Untersuchungsfeld angesehen (Löffler-Stastka et al., 2019). Bisher wurden Studien zur Wirksamkeit von Ausbildungsbehandlungen hauptsächlich an psychotherapeutischen Ambulanzen – und vorwiegend bei verhaltenstherapeutischen Therapien – durchgeführt (Victor et al., 2018). Daten zur Behandlungsqualität von Ausbildungstherapien anderer Verfahren und in freier Praxis wurden dagegen kaum bis gar nicht zur Veröffentlichung gebracht. Die Ergebnisse der wenigen Studien aus Deutschland zu diesem Thema deuten auf hohe Therapieeffekte bei Behandlungen von Ausbildungskandidat\*innen in ambulanten Einrichtungen hin, welche mit jenen fertig ausgebildeter Therapeut\*innen vergleichbar sind (Paine, Bell, Sandage, Rupert, Bronstein, O'Rourke, Stavros, Moon & Kehoe, 2019; Peikert, Baum, Barchmann, Schröder &

Kropp, 2014; Victor et al., 2018). So konnten von Peikert und Kolleg\*innen (2014) keine wesentlichen Unterschiede im wahrgenommenen Behandlungserfolg zwischen Klient\*innen einer verhaltenstherapeutischen Ausbildungsambulanz und Klient\*innen mit regulären Behandlungen bei ausgebildeten Verhaltenstherapeut\*innen in freier Praxis finden (Peikert et al., 2014). Des Weiteren ermittelten Victor und Kolleg\*innen (2018) bei Behandlungen an mehreren verhaltenstherapeutischen Ausbildungsambulanzen große Therapieeffekte mit Stärken zwischen Cohen's  $d = 1.24$  und  $2.02$  für eine Completer-Stichprobe. Zudem gaben sie Response-Raten von 75% bzw. 85% und Remission-Raten von 56% bzw. 73% an (Victor et al., 2018).

Vor dem Hintergrund teils prekärer Arbeitsverhältnisse von Psychotherapeut\*innen während der Ausbildung aufgrund fehlender Verträge mit Sozialversicherungen und geringeren Stundenhonoraren sollten weitere Studien zur Wirksamkeit von Ausbildungstherapien an vorherige Befunde anknüpfen und die bisherige Evidenz vergrößern. Ein vermehrtes Forschungsaufkommen in diesem Bereich könnte die Versorgungsrelevanz, die Psychotherapeut\*innen i.A.u.S. in der öffentlichen Gesundheitsversorgung tatsächlich leisten, adäquat abbilden. Dadurch könnte der Stellenwert der Versorgungsleistung von Psychotherapeut\*innen i.A.u.S. bei gesundheitspolitischen Verantwortungsträger\*innen, aber auch bei Klient\*innen, erhöht werden (Löffler-Stastka et al., 2018; Peikert et al., 2014).

Das betrifft in Österreich immerhin 4.380 Personen (Stand 01.06.2020), die sich in fachspezifischer Ausbildung befinden (Löffler-Stastka & Hochgerner, 2021). Von den über 40 Ausbildungseinrichtungen in Österreich weist die Gesellschaft für Logotherapie und Existenzanalyse (GLE-Ö) mit rund 424 Ausbildungskandidat\*innen (10%) die größte Anzahl an Auszubildenden auf (Sagerschnig & Valady, 2021). Die Logotherapie und Existenzanalyse wird zu den humanistischen Psychotherapieverfahren gezählt.

### **Humanistische Psychotherapie**

Der humanistische Psychotherapiecluster begann sich in den 1960er Jahren systematisch zu formieren und grenzte sich als dritte große Psychotherapiegruppierung einerseits von dem deterministisch-analytischem Menschenbild der (damaligen) Psychoanalyse und andererseits von der kausalen-reiz-reaktions-Sichtweise der (damaligen) Verhaltenstherapie ab (Kriz, 2014). Zum Cluster der humanistischen Psychotherapie werden in Österreich die *Existenzanalyse*, die *Existenzanalyse und Logotherapie*, die *Gestalttheoretische Psychotherapie*, die *Integrative Gestalttherapie*, die *Integrative Therapie*, die *Klientenzentrierte Psychotherapie*, die *Personzentrierte Psychotherapie* und das *Psychodrama* gezählt (Kierein & Sagl, 2020).

Die Verfahren im humanistischen Cluster weisen eine große methodische Vielfalt auf, die sich von nahezu technikfreien Verfahren (Personzentrierte Psychotherapie) bis hin zu Verfahren mit klar strukturierteren Interventionen (Gestalttherapie) erstrecken können (Angus, Watson, Elliott, Schneider & Timulak, 2015; Gaab, Ziem & Flückinger, 2020; Stumm, 2017). Trotz dieser Vielfalt können Übereinstimmungen in den theoretischen Grundgerüsten der humanistischen Verfahren ausgemacht werden (Bohart et al., 1998; Gaab et al., 2020; Stumm, 2017). Humanistische Psychotherapietheorien sind in der *Existenzphilosophie* verwurzelt und beschäftigen sich mit Fragen nach dem Menschsein in der Welt und den Bedingungen für ein erfülltes Leben (Gaab et al., 2020; Kriz, 2014; Stumm, 2017). Der Mensch wird als aktives, holistisches und sinnorientiertes Wesen verstanden, das sich im Spannungsfeld zwischen Autonomie, Freiheit und Selbstbestimmtheit bei gleichzeitig sozialer Interdependenz einordnen muss (Kriz, 2014; Stumm 2017).

Humanistische Psychotherapiebehandlungen folgen einigen Grundprinzipien, von denen in den folgenden Abschnitten vier genauer beschrieben werden (Angus et al., 2015):

Der Fokus in humanistischen Psychotherapien liegt weniger auf den Diagnosen oder der Reduzierung von Symptomen einer Störung, sondern in erster Linie auf *dem persönlichen Wachstum* der Klient\*innen (Gaab et al., 2020). Es wird davon ausgegangen, dass jedem Menschen ein Potenzial zur *Selbstverwirklichung* innewohnt, also dem Bedürfnis und die Fähigkeit die eigenen Begabungen und Wünsche zu verwirklichen und das eigene Leben konstruktiv sich und der (Um-) Welt gegenüber zu gestalten (Helle, 2019). Innerhalb humanistischer Therapien sollen die Klient\*innen zu einem kongruenten Selbstverständnis und -ausdruck verholfen werden, um ein für sie bedeutsames, also sinnvolles, Leben führen zu können (Cain, Keenan & Rubin, 2016). Auf diese Weise soll indirekt auch auf bestehende Symptomausprägungen eingewirkt werden. Störungsbilder stellen zwar wichtige Orientierungspunkte im psychotherapeutischen Prozess dar, müssen aber nicht vordergründig (be-) handlungsanleitend sein (Shean, 2014).

Als wesentliche Komponente in der therapeutischen Arbeit wird das subjektive *Erleben* der Klient\*innen beschrieben, das deren Lebensrealität formt und gestaltet (Gaab et al., 2020). Der Mensch wird in der humanistischen Psychotherapie als ganzheitliche Entität begriffen, welche von zahlreichen Faktoren wechselseitig und unterschiedlich stark beeinflusst wird. Bedingungen wie die körperliche und die psychische Konstitution, die sozialen Ressourcen, die bisherigen Erfahrungen und der kulturelle Kontext, um nur einige zu nennen, stehen in fortwährender Wechselwirkung zueinander und bestimmen das Erleben eines Menschen (Cain et al., 2016). Dabei wird die Einzigartigkeit jeder einzelnen Person

deutlich, die innerhalb der Therapie bei jedem\*r Klient\*in neu erschlossen werden muss, wobei die Klient\*innen als Expert\*innen ihrer selbst betrachtet werden (Gaab et al., 2020; Stumm, 2017).

Der Behandlungserfolg wird in humanistisch fundierten Psychotherapien in hohem Maße an einer vertrauensvollen und tragfähigen *Therapeut\*in-Klient\*in-Beziehung* festgemacht (Angus et al., 2015; Helle, 2019). Humanistische Psychotherapeut\*innen verstehen sich weniger als Durchführende von (standardisierten) Techniken und Manualen, denn als Beziehungsanbieter\*innen und Dialogpartner\*innen, die flexibel auf die individuelle Lebenssituation der Klient\*innen eingehen (Bohart et al., 1998; Cain et al., 2016). Die therapeutische Beziehungsgestaltung richtet sich dabei an das Erleben und den Bedürfnissen der Klient\*innen und ist geprägt von einer professionellen, empathischen, wertschätzenden und aufrichtigen Haltung der Therapeut\*innen. Auf dieser Weise soll eine stützende Beziehungserfahrung geschaffen werden, die es den Klient\*innen durch wohlwollender Rückmeldung ermöglicht sich selbst besser zu verstehen, anzunehmen und gegebenenfalls neu auszurichten (Helle, 2019).

Eine gelungene, therapeutische Beziehung wird durch ein phänomenologisches Vorgehen geschaffen. *Phänomenologie* kann als Erkenntnisform definiert und im therapeutischen Kontext auch als *Wesensschau* bezeichnet werden (Längle, 2014). Dabei werden einerseits die Gesetzmäßigkeiten der Logik von den Geschehnissen eines denkenden und fühlenden Bewusstseins und andererseits eine *beschreibende* von einer *verstehenden* Haltung unterschieden. So kann beispielsweise der Zustand und die Situation von Klient\*innen über Symptomen, Störungsbildern, Theorien und Vorannahmen beschrieben jedoch nicht darauf reduziert oder verstanden werden (Angus et al., 2015). *Verstehen* im Sinne der Phänomenologie beschreibt ein möglichst unvoreingenommenes, hypothesen- und theoriefreies Erfassen, das sich unmittelbar am Phänomen, dem Erleben einer Person, orientiert (Cain et al., 2016).

Diese Prinzipien prägen humanistische Psychotherapien nicht nur in den Behandlungen, sondern auch in ihrer Forschungsausrichtung (Angus et al., 2015; Bohart et al., 1998; Horvath, Del Re, Flückiger & Symonds, 2011; Längle, 2007; Längle & Kriz, 2012). Daraus ergibt sich ein deutlicher Unterschied zur quantitativen Forschungsdominanz in der psychotherapeutischen Wirksamkeitsforschung, die sich hauptsächlich auf quantifizierbare und operationalisierbare Variablen wie Veränderungen in der Symptomausprägungen als Abbildung des Therapiefortschrittes verständigt (Revenstorf, 2005; Schneider & Längle, 2012). Trotz dieser Diskrepanz zwischen der humanistischer

Epistemologie und quantitativer Forschungsmethodik verzeichnet der humanistische Cluster ein steigendes nomothetisches Forschungsaufkommen, dessen Befunde für die Bedeutung humanistischer Konzepte im psychotherapeutischen Kontext sprechen (Angus et al., 2015; Cain et al., 2016; Helle, 2019; Längle & Kriz, 2012; Shean, 2014).

So deuten Studienergebnisse darauf hin, dass humanistische Psychotherapieverfahren große Behandlungseffekte aufweisen können, welche auch langfristig bei Katamnesemessungen nachgewiesen werden konnten (Angus et al., 2015; Helle, 2019). Als möglichen Grund für die nachhaltigen Therapieeffekte führen Forscher\*innen den Fokus humanistischer Verfahren auf den Aufbau von Selbstverständnis und -bestimmung anstelle von primärer Symptomreduzierung durch konkrete Verhaltensmodifikation an (Angus et al., 2015; Längle et al., 2000). Zudem finden originär humanistische Konzepte immer häufiger Eingang in die klinisch-psychologische Forschung. Befunde von Studien aus der Positiven Psychologie oder der sinnzentrierten und emotionsfokussierten Forschung stützen indirekt auch das Theoriegerüst humanistischer Verfahren (Cain et al., 2016; Thir & Batthyány, 2016; Vos, Cooper, Correia & Craig, 2015; Vos & Vitali, 2018; Wong, 2017).

### ***Logotherapie und Existenzanalyse***

Charakteristisch für das von Viktor Frankl entwickelte psychotherapeutische Verfahren ist die Bedeutung der Sinnmotivation für ein erfülltes Leben (Kriz, 2014; Längle, 2015). Frankl erkannte in den 1930er-Jahren die Notwendigkeit einer *Rehumanisierung* der Psychotherapie, weswegen die Frage nach dem (Lebens-) Sinn als zentrale anthropologische Komponente in die psychotherapeutischen Arbeitstheorien mitaufgenommen wurde (Längle, 2015). Das damalige medizinisch-psychiatrische Weltbild der Psychoanalyse und der Individualpsychologie sollte um das menschliche Bedürfnis nach Sinn erweitert werden (Kriz, 2014; Stumm, 2017).

Aus logotherapeutischer bzw. existenzanalytischer Sicht stellt *Sinn* eine wichtige Motivation des Menschen dar, welche zu einer authentischen und verantwortungsbewussten Existenz beitragen kann (Thir & Batthyány, 2016). Konzepte wie Freiheit, Verantwortung und Selbstbestimmung werden dabei als wichtige Dimensionen des menschlichen Daseins verstanden, die im logotherapeutischen Psychotherapieprozess berücksichtigt werden (Längle, 2021). Besonders die *Selbst-Distanzierung*, die Fähigkeit, sich selbst von einer Außenperspektive zu betrachten sowie die *Selbst-Transzendenz*, die Fähigkeit, sich ohne Selbstaufgabe mit etwas Äußerem, wie z.B. mit einer Aufgabe oder mit anderen Menschen zu verbinden, stellen zwei wichtige Pole in der Verortung der eigenen Person in der Welt dar (Stumm, 2017).

Ein Mangel an Bedeutsamkeit oder ein fehlender Wille zur Sinnsuche in der eigenen Lebensgestaltung kann zu einem *existenziellen Vakuum* führen, dessen Folgen ein gestörter Dialog mit sich und der Umwelt sowie das Fehlen von Ausdruck und Austausch sein können – Umstände, die wiederum Frustrationserleben mit sich bringen können (Längle, 2021). Ein anhaltender Zustand von (existenzieller) Frustration kann in klinisch relevante Symptome und Störungsbilder münden (Längle, 2015; Kriz, 2014; Stumm, 2017; Thir & Batthyány, 2016). Die ursprüngliche therapeutische Theorie und Methode Frankls prägt eine gewisse Rationalität und eine kognitive Betrachtungsweise (Längle & Klaassen, 2019; Reitinger & Bauer, 2019; Riemeyer, 2007). In diesem Sinne können logotherapeutische Methoden wie z.B. die *paradoxe Intention* und *Dereflexion* heute auch im systemischen und verhaltenstherapeutischen Ansatz gefunden werden (Kriz, 2014). Aufbauend auf die Theorien und Methoden der Logotherapie wurden von Alfred Längle mitunter Konzepte zur subjektiven Emotionalität und eine phänomenologische Herangehensweise integriert (Längle, 2015; Stumm, 2017; Kriz, 2014). Sie wurde so zur Existenzanalyse erweitert.

### ***Existenzanalyse***

Während die Logotherapie primär die Motivation des *Willens zum Sinn* hervorhebt und Klient\*innen zur Suche danach befähigen möchte, ist die Personale Existenzanalyse ein phänomenologisches Psychotherapieverfahren, das einer Person zu einem eigenverantwortlichen Umgang mit sich und der Welt verhelfen möchte (Längle, 2021). *Existenz* wird dabei als ein im Idealfall sinnerfülltes, in Freiheit und Selbstverantwortung gestaltetes Leben definiert, das von einer Person selbstbestimmt geführt werden kann. Der Wille zum Sinn stellt eine wesentliche, aber nicht ausschließliche, Motivation der menschlichen Existenz dar (Längle, Van Deurzen, Craig, Schneider, Tantam, & Plock 2019; Längle, 2014). Die kognitiv verstandene Sinn-Motivation des Menschen wird ergänzt um das subjektive und emotionale Erleben eines Individuums. Die Methode der Phänomenologie soll es den Therapeut\*innen ermöglichen, sich dem Erleben der Klient\*innen mit einer verstehenden Haltung anzunähern (Längle & Klaassen, 2019; Reitinger & Bauer, 2019). Trotz der Hervorhebung der Selbstbestimmung und Freiheit einer Person kann eine sinnerfüllte Existenz nicht (nur) aus sich selbst heraus entstehen. Erst die Zuwendung zur Umwelt über Begegnungen mit Personen oder anhand von Aufgaben jedweder Art, komplementiert das (Sinn-) Erleben eines Menschen (Längle, 2015).

Im *Strukturmodell* der Personalen Existenzanalyse werden vier Grundmotivationen des Menschen beschrieben, die als Dimensionen einer erfüllten Existenz gelten und je nach Person und Lebenslage vielfältige Ausprägungsformen aufweisen können (Längle, 2015).

Die erste Grundmotivation, *Sein-Können* oder *Weltbezug* (Längle, 2015, 2019), wird von den Lebensumständen einer Person und deren Fähigkeit, diesen zu begegnen, bestimmt. Damit sind einerseits Annehmen und Aushalten und andererseits aktive Begegnung mit der eigenen Befindlichkeit und den Lebensumständen gemeint, die (noch) nicht veränderbar sind. Bei Erfüllung dieser Motivation kann ein Gefühl von Vertrauen und Aufgehoben-Sein entstehen. Defizite in dieser Grundmotivation werden mit Symptomen von Angststörungen bzw. psychotischen Zuständen in Verbindung gebracht (Längle, 2021).

Die zweite Grundmotivation wird als *Leben-Mögen* oder *Lebensbezug* (Längle, 2015, 2019) bezeichnet. Dabei ist insbesondere das subjektive (Wohl-) Befinden gemeint. Es umschreibt ein kongruentes Verbunden-Sein mit sich und der Welt, welches durch Offenheit und Resonanz geprägt ist. Ein Mangel in diesem Bereich kann zu Isolation, Selbstablehnung und Entfremdungsgefühlen führen und steht mit affektiven Störungen in Zusammenhang (Längle, 2015).

Die dritte Grundmotivation, *Selbstsein-Dürfen* oder *Selbstbezug* (Längle, 2015, 2019), ähnelt dem psychologischen Konzept des Selbstwertes einer Person. Einerseits ist das Bewusstsein über sich selbst durch Selbstannahme und -akzeptanz bedeutend für eine Erfüllung dieser Grundmotivation, andererseits bedarf es auch wertschätzender und respektvoller Rückmeldung der (sozialen) Umwelt. Ein Mangel in dieser Dimension kann mit Symptomen von (histrionischen) Persönlichkeitsstörungen in Verbindung gebracht werden (Längle, 2015).

Die vierte Dimension entspricht dem Sinn-Konzept nach Frankl und wird als *Handeln-Sollen* oder *Sinnbezug* bezeichnet (Längle, 2015, 2019). Die eigene Existenz, das eigene Sein und Tun werden dabei in einen größeren Kontext gestellt. Individueller Lebenssinn kann Antworten auf die Frage nach dem *Wofür* geben. Existenzielles Bedeutungserleben ermöglicht Orientierung in der Gegenwart und darüber hinaus. Ein Mangel in der vierten Grundmotivation kann zu Lebensfrustration und innerer Leere führen. Defizite in diesem Bereich können mit übermäßigem Idealismus und aggressivem, fanatischem Verhalten sowie mit Suizidalität und Süchten in Zusammenhang gebracht werden (Längle, 2015).

Die vier Grundmotivationen können individuell sehr unterschiedlich ausgeprägt sein. Um dieser Individualität gerecht zu werden, wird der therapeutische Prozess klient\*innen-zentriert und dialogisch gestaltet (Kriz, 2014; Längle & Klaassen, 2019; Vos & Vitali, 2018). Klient\*innen werden als holistische Entitäten begriffen, deren einzigartige, emotional-konstruktivistische Lebensrealität nur in ihrer kontextuellen Gesamtheit verstanden und bewegt werden kann (Längle et al., 2019; Reitinger & Bauer, 2019).

Ein zentrales Ziel in der existenzanalytischen Therapie ist es, den Klient\*innen ein Leben mit innerer Zustimmung zu ermöglichen, indem eine offene, dialogische und eigenverantwortliche Haltung in Bezug auf sich selbst und der Welt angeregt wird (Längle, 2014). Dies geschieht weniger durch die aktive Suche nach (Lebens-) Sinn, sondern in Form eines Anregens der Klient\*innen zu einem Erspüren, Begreifen und Ausleben der eigenen Umstände, Wünsche, Bedürfnisse und der eigenen Emotionalität (Görtz, 2020). Dadurch soll eine authentische Entscheidungsfähigkeit unterstützt werden, die zu einem verantwortungsvollen Handeln und freien Empfinden verhelfen soll. Spezifische Verhaltensmodifikationen und direkte Symptombehandlung spielen dabei eine eher untergeordnete Rolle (Görtz, 2020; Kriz, 2014).

Anstelle von spezifischen Behandlungstechniken wird die therapeutische Beziehung auf Basis der Phänomenologie als für den Behandlungsprozess wesentlich beschrieben (Längle, 2014). Existenzanalytische Psychotherapeut\*innen bieten mit dem phänomenologischen Zugang eine offene Haltung gegenüber dem Erleben der Klient\*innen an. Therapeutische Offenheit bedeutet, die personeneigenen Phänomene der Klient\*innen als Maß der Dinge zu betrachten und diese Phänomene im holistischen Gefüge der Person zu verstehen und zu verorten (Görtz, 2020). Im dialogischen Therapieprozess soll das professionelle Verstehen der Therapeut\*innen den Klient\*innen zu einem besseren Verständnis ihrer eigenen Lebenssituation und unter anderem zur Integration schwieriger Lebensumstände verhelfen (Längle, 2014; Launeanu, Klaassen & Muir, 2019).

### **Wirksamkeit existenzanalytischer Psychotherapie.**

Trotz der Positionierung der Existenzanalyse in der idiographischen Methodik gibt es auch nomothetisch-quantitative Forschungsergebnisse, die die Wirksamkeit sinnzentrierter Interventionen stützen und die Relevanz von Sinn und Bedeutung für die Psychopathologie und das Wohlbefinden hervorheben (Thir & Batthyány, 2016; Vos, Cooper, Correia & Craig, 2015; Vos & Vitali, 2018).

Studienergebnisse geben Hinweise, dass Sinnerleben positiv mit Lebensqualität, Hoffnung, Emotionsregulation, Selbstwirksamkeit und Resilienz korreliert (Thir & Batthyány, 2016; Vos & Vitali, 2018). Dagegen kann geringes Sinnerleben mit Angst und Hoffnungslosigkeitserleben, z.B. bei Krebspatient\*innen, einhergehen und mit Psychopathologien wie Depression oder Angst in Zusammenhang gebracht werden (Vos & Vitali, 2018). Ergebnisse aus Katamnesestudien deuten zusätzlich darauf hin, dass sinnzentrierte Psychotherapie nachhaltige Therapieeffekte aufweisen können (Thir & Batthyány, 2016; Vos et al., 2015). Auch die Gesellschaft für Logotherapie und

Existenzanalyse in Österreich (GLE-Ö) hat bereits einige naturalistische Studien durchgeführt, deren Ergebnisse für eine gute praktische Wirksamkeit dieser therapeutischen Ausrichtung sprechen und in den folgenden Absätzen beschrieben werden.

**Wirksamkeit im stationären Setting (Längle, Görtz, Rauch, Jarosik & Haller, 2000).** In den Jahren 1994 bis 2000 wurde auf der *Station Carina* im Vorarlberger *Krankenhaus Maria Ebene* eine naturalistische Wirksamkeitsstudie bei Suchtpatient\*innen im stationären Setting durchgeführt. Bei der prospektiven Studie mit mehreren Messzeitpunkten handelte es sich um eine der ersten empirischen Überprüfungen dieser Art der Personalen Existenzanalyse nach Alfred Längle. Als Vergleichsgruppe wurden Daten von Patient\*innen der Station erhoben, die Einzeltherapien einer anderen Verfahrenskonzeption (Tiefenpsychologie, Verhaltenstherapie oder Systemische Familientherapie) erhielten. Es konnten Daten von 337 Personen erhoben werden. Davon erhielten 280 Personen eine existenzanalytische Einzeltherapie und 57 Personen eine Einzeltherapie nach einem anderen Verfahren (Längle et al., 2000). Das Durchschnittsalter der Stichprobe lag bei 41 Jahren. 60 % davon waren Männer und 74 % der Stichprobe litt unter einer Alkoholabhängigkeit.

Die Existenzanalyse wies in den Studienergebnissen eine mit den anderen psychotherapeutischen Schulen vergleichbare Effektivität bei Klient\*innen mit Suchproblematiken auf. Die Prä-Post-Effektstärke existenzanalytischer Behandlungen lag auf der Gesamtskala des verfahrenseigenen Erhebungsinstruments, der Existenzskala, bei Cohen's  $d = 0.46$ , was für einen kleinen bis mittleren Effekt spricht. Die Prä-Katamnese-Effektstärke auf derselben Skala lag bei Cohen's  $d = 0.74$  und kann als mittlerer bis großer Effekt eingeschätzt werden. Die erhöhten Effektstärken sechs Monate nach Therapieende deuten auf einen langfristigen Therapieeffekt der Existenzanalyse hin. Für die Studienautor\*innen ist dies ein Hinweis auf den verfahrenstypischen Fokus auf den Aufbau von Autonomie und Eigenverantwortung statt einer vorwiegenden Symptombearbeitung, was zu längerfristigen Effekten führen könnte (Längle et al. 2000). Es wurden keine Effektstärkenangaben für die Behandlungen der anderen Verfahren gemacht. Jedoch wurden die erzielten Rohwertsdifferenzen angegeben. So wies die existenzanalytische Behandlungsgruppe ( $n = 174$ ) bei der ESK-Gesamtskala eine Prä-Post-Rohwertsdifferenz von 14.2 Testpunkten auf und die Behandlungsgruppe der anderen Verfahren ( $n = 40$ ) veränderte sich in ihrer Testausprägung um 13.4 Testpunkte (ebda.).

**Wirksamkeit in freier Praxis (Längle, Görtz, Probst, Probst, Lopatka, Kubin & Steinert, 2005).** In den Jahren 1993 bis 2004 wurde von der Gesellschaft für Logotherapie und Existenzanalyse eine naturalistische Studie mit dem Ziel durchgeführt, die Wirksamkeit der existenzanalytischen Therapie in freier Praxis zu erheben (Längle, Görtz, Probst, Probst, Lopatka, Kubin & Steinert, 2005). Es handelte sich auch hier um eine prospektive Studie mit mehreren Testzeitpunkten. Es konnten verwertbare Daten von zwölf Therapeut\*innen aus dem Raum Graz und Wien und von 248 Klient\*innen erhoben werden. Die Klient\*innenstichprobe bestand zu 71 % aus Frauen. Das Durchschnittsalter betrug 32 Jahre. Es wurden signifikante Verbesserungen der Symptomatik zwischen Therapiebeginn und -ende mit Effektstärken von Cohend's  $d = 0.99$  für die Gesamtskalen der Existenzskala (ESK) bzw. durchschnittlich Cohend's  $d = 0.98$  für die Skalen der Kieler Änderungssensitiven Symptomliste (KASSL) ermittelt, was für einen großen Therapieeffekt spricht. Zudem wurden nach allen zehn Einheiten Zwischenerhebungen durchgeführt. Aus den Verlaufsdaten wird ersichtlich, dass sich die Effektstärken vom ersten bis zum letzten Messzeitpunkt mit jedem weiteren Testzeitpunkt erhöhten (Längle, Görtz, Probst, Probst, Lopatka, Kubin & Steinert, 2005).

**Retrospektives Studiendesign in freier Praxis (Laireiter, Schaireiter, Russegger, Baumgartner, Längle & Sauer, 2013).** Im Zuge einer retrospektiven Studie wurden Therapeut\*innen und deren Klient\*innen gebeten nach Therapieabschluss den Erfolg ihrer Psychotherapien einzuschätzen (Laireiter, Schaireiter, Russegger, Baumgartner, Längle & Sauer, 2013). Dabei sollte die praktische Wirksamkeit hinsichtlich einer Verbesserung der körperlichen, psychischen und sozialen Beeinträchtigung und des subjektiv erlebten Therapieerfolgs ermittelt werden. Personen mit abgeschlossener bzw. abgebrochener Therapie wurden miteinander verglichen. Der Erhebungszeitraum lag zwischen den Jahren 1996 und 1998. Es konnten 26 Therapeut\*innen angeworben werden, die wiederum 106 ehemalige Klient\*innen rekrutieren konnten. Die Klient\*innenstichprobe bestand zu 69 % aus Frauen und die Teilnehmer\*innen wiesen ein Durchschnittsalter von 37 Jahren auf. Knapp über 20 % der Klient\*innenstichprobe beendete die Therapie vorzeitig. Die Klient\*innen gaben zu Therapieende eine hohe bis sehr hoch erlebte Verbesserung ihrer Symptombelastung an. Durch die retrospektive Beurteilung zum Zustand zu Therapiebeginn und den Angaben zu Therapieende konnten Effektstärken von Cohen's  $d > 0.80$  ermittelt werden. Rund 90 % der Klient\*innen gaben an, eine positive Veränderung erlebt zu haben. Weitere 10 % gaben an keine Veränderung wahrgenommen zu haben (Laireiter, Schaireiter, Russegger, Baumgartner, Längle & Sauer, 2013).

***Zufriedenheit mit der Existenzanalyse (Laireiter, Schirl, Kimeswenger, Längle & Sauer, 2000).*** 26 teilnahmebereite Psychotherapeut\*innen wurden ersucht, alle Klient\*innen zwischen 20 und 60 Jahren, die im Jahr 1997 bei ihnen eine Psychotherapie abgeschlossen bzw. abgebrochen hatten, um eine Studienteilnahme zu bitten. 108 Klient\*innen waren bereit, Fragebögen zur Beurteilung der Psychotherapie auszufüllen und zurückzusenden. Die Stichprobe bestand aus 69 % Frauen und das Durchschnittsalter der Klient\*innen lag bei 37 Jahren. Von den 108 Klient\*innen wiesen 23 Personen einen vorzeitigen Therapieabbruch auf. Die Abbruchsgründe konnten mehrheitlich als nicht-qualitätsrelevant (z.B. Umzug, finanzielle Probleme etc.) eingestuft werden. Der Therapieerfolg wurde von der Gesamtstichprobe als überwiegend positiv beschrieben. Rund 56 % der Befragten beurteilten die Therapie als (sehr) erfolgreich und 40 % beschrieben sie als gering bzw. etwas erfolgreich. Besonders die Psychotherapeut\*innen wurden rückblickend als sehr positiv erlebt. Weniger gut wurden hingegen das therapeutische Vorgehen und die therapeutischen Effekte beurteilt. Insgesamt kann hier von einer durchschnittlichen Zufriedenheit gesprochen werden. Die Gruppe der Abbrecher\*innen bewertete den therapeutischen Prozess und dessen Ergebnisse signifikant schlechter als jene Gruppe, die die Therapie regulär beendet hatten. Die Studienautor\*innen fassen die Ergebnisse auf einer Schulnotenskala zusammen, wobei die Qualität der Behandlung durch die Therapeut\*innen und deren Gestaltung der therapeutischen Beziehung mit *gut* bis *sehr gut* beurteilt wurden. Die Durchführung der Therapie erhielt von den Studienautor\*innen eine *gute* bis *zufriedenstellende* Note, und der Therapieerfolg wurde als *zufriedenstellend* bzw. *genügend* beurteilt (Laireiter, Schirl, Kimeswenger, Längle & Sauer, 2000).

Die Ergebnisse der von der GLE-Ö durchgeführten Studien deuten auf eine insgesamt gute Wirksamkeit der Existenzanalyse in der Versorgungspraxis hin. Eine häufig erwähnte Limitation von Studien zur Wirksamkeit existenzanalytischer Psychotherapien ist der geringe Stichprobenumfang sowohl bei den Klient\*innen als auch bei den Therapeut\*innen (Laireiter et al., 2013; Vos et al., 2015). Zudem fehlt bislang eine Überprüfung der praktischen Wirksamkeit von Psychotherapie bei existenzanalytischen Psychotherapeut\*innen in Ausbildung unter Supervision (i.A.u.S.) in freier Praxis, jedoch nicht nur im Verfahren der Existenzanalyse.

### **Zusammenfassung & Forschungsinteresse**

Die Wirksamkeit von Psychotherapie wurde in zahlreichen RCT- Studien ermittelt und gilt als evident (Wampold & Imel, 2015). Die Stärke von experimentellen Studiendesigns ist es, den potenziellen Behandlungseffekt spezifischer Interventionen für konkrete Störungen

unter bestimmten Voraussetzungen sowie in einem kontrollierten Umfeld zu untersuchen. Ergebnisse daraus haben somit eine hohe interne Validität (Chambless & Ollendick, 2001; Persons & Silberschatz, 1998). Die Generalisierbarkeit von Ergebnisse aus RCT-Untersuchungen ist für andere Behandlungskontexte wie der psychotherapeutischen Routineversorgung jedoch weniger gegeben, ihre externe Validität wird als gering eingestuft (Lutz et al., 2021; Seligman, 1995). Zudem deuten Studienergebnisse auf die Bedeutsamkeit von allgemeinen und kontextuellen Wirkfaktoren und einen vergleichbar eher geringen Einfluss von spezifischen Techniken auf die Wirkvarianz von Psychotherapie hin (Deaton & Cartwright, 2018; Strauss, 2019; Tschuschke & Freyberger, 2015). Psychotherapieverfahren, deren Behandlungskonzept primär auf kontextuellen Wirkfaktoren und Ressourcenaktivierung beruht, können adäquater mit methodisch und thematisch offeneren Studiendesigns untersucht werden (Castonguay et al., 2013; Vaughan et al., 1994). Das betrifft unter anderem Verfahren, die dem humanistischen Psychotherapiecluster zugeordnet werden (Bohart et al., 1998). Dazu zählt auch die Existenzanalyse. Das Ziel einer existenzanalytischen Behandlung ist es, die Klient\*innen zu einer erfüllten Existenz zu befähigen (Görtz, 2020; Längle, 2014). Der durch die Phänomenologie gestützte Therapieprozess der Existenzanalyse ist mit den methodischen Anforderungen von RCT-Studien weniger kompatibel. In der praxisorientierten Psychotherapieforschung wird Psychotherapie im Versorgungssetting untersucht (Castonguay et al., 2015; Herzog et al., 2022; Hiller et al., 2011). Dafür werden naturalistische Studiendesigns eingesetzt, die sich den Bedingungen der Praxis anpassen können (Tschuschke & Freyberger, 2015).

Im Jahre 2020 wies der Ausbildungsverein der Gesellschaft für Logotherapie und Existenzanalyse die größte Gruppe an fachspezifischen Ausbildungskandidat\*innen in ganz Österreich auf (Sagerschnig & Valady, 2021). Psychotherapien, die von Auszubildenden durchgeführt werden, nehmen zwar eine wichtige Rolle in der Versorgungslandschaft ein, wurden bislang jedoch wenig untersucht (Löffler-Stastka et al., 2018; Victor et al., 2018). Während es zumindest einige Studien zur Ergebnisqualität von Ausbildungstherapien in ambulanten Einrichtungen gibt, fehlen Studien bei Psychotherapeut\*innen i.A.u.S. in freier Praxis nahezu gänzlich (Paine et al., 2019; Peikert et al., 2014; Victor et al., 2018). Mit der naturalistischen Studie vorliegender Arbeit soll sich dieser Forschungslücke angenommen werden.

## **Forschungsfragen & Hypothesen**

Das primäre Ziel der Studie in vorliegender Arbeit ist die Evaluierung der praktischen Wirksamkeit der Existenzanalyse bei Psychotherapeut\*innen in Ausbildung unter Supervision in freier Praxis. In Anlehnung an die Herausforderungen und Probleme, welche bei naturalistischen Praxisstudien in der Psychotherapieforschung entstehen können, soll zuvor die Durchführbarkeit des naturalistischen Studiendesigns geklärt werden. Die nachstehenden Fragen und Hypothesen sollen im Zuge dieser Arbeit untersucht werden:

Fragestellung I: Ist das naturalistische Studiendesign durchführbar?

### **1.1 Wird die Studie von den Therapeut\*innen akzeptiert und angenommen?**

- 1.1.1 Wie hoch ist die Rücklaufquote bei den Therapeut\*innen?
- 1.1.2 Ist die erlangte Stichprobe der Therapeut\*innen repräsentativ für die Zielgruppe?
- 1.1.3 Wie gut lässt sich die Studie in den therapeutischen Alltag integrieren?
- 1.1.4 Welche Gründe für eine Teilnahme werden angeführt? Gibt es wahrgenommene Vorteile?
- 1.1.5 Wird die Möglichkeit der Rückmeldungen angenommen? Wie werden diese verwendet?

### **1.2 Wird die Studie von den Klient\*innen akzeptiert und angenommen?**

- 1.2.1 Wie hoch ist die Rücklaufquote bei den Klient\*innen?
- 1.2.2 Gibt es Unterschiede zwischen den Personen, die an der Studie teilnehmen und jenen, die eine Studienteilnahme ablehnen?
- 1.2.3 Wie sind die Meinungen der Klient\*innen zu den Fragebögen bezüglich Länge, Dauer, Verständnis und Ausfüllerleben?
- 1.2.4 Welche Gründe für eine Teilnahme werden angeführt? Gibt es wahrgenommene Vorteile?
- 1.2.5 Wie hoch ist die Abbruchsquote bzw. der Fragebogenrücklauf? Welche Abbruchsgründe gibt es?
- 1.2.6 Ist die erlangte Stichprobe der Klient\*innen repräsentativ für die Zielgruppe? Welche Personen suchen eine existenzanalytische Psychotherapie bei Auszubildenden in freier Praxis auf?
- 1.2.7 Welche Charakteristika weist die Gesamtstichprobe der Klient\*innen zu Therapiebeginn auf?

Fragestellung II: Ist die existenzanalytische Ausbildungstherapie in freier Praxis wirksam?

## **2.1. Gibt es gruppenstatistisch signifikante Veränderungen zwischen Therapiebeginn und –ende?**

2.1.1. Können Veränderungen in den Skalenwerten der Prä-Post-Stichprobe zwischen Therapiebeginn und -ende gefunden werden?

***Prä-Post-Hypothese:*** Es gibt einen Unterschied in den Gruppenmittelwerten der Skalen, Global Severity Index (GSI), Depression und Angst der Symptom-Checkliste-90, Gesamtskala des Inventars zur Erfassung Interpersonaler Probleme (IIP-D), Gesamtskala der Existenzskala (ESK), Ryff-Skala und Global Assessment of Functioning (GAF), zwischen Therapiebeginn und -ende.

## **2.2 Gibt es reliable und klinisch signifikante Veränderung?**

2.2.1 Wie viele Personen der Prä-Post Stichprobe weisen eine reliable Veränderung auf?

2.2.2 Wie viele Personen der Prä-Post Stichprobe weisen zwischen Therapiebeginn und -ende eine klinisch signifikante Verbesserung auf?

## **2.3 Wie beurteilen die Klient\*innen und ihre Therapeut\*innen den psychotherapeutischen Prozess zu Therapieende?**

2.3.1 Wie schätzen die Klient\*innen und ihre Therapeut\*innen den Therapieerfolg zu Therapieende ein?

2.3.2 Wie zufrieden sind die Klient\*innen und ihre Therapeut\*innen mit der durchgeführten Behandlung zu Therapieende?

2.3.3 Wurden aus Sicht der Klient\*innen und ihren Therapeut\*innen die Therapieziele erreicht?

## **2.4 Gibt es kurz- bzw. langfristige, gruppenstatistisch signifikante Veränderungen zwischen Therapiebeginn und der Verlaufsmessung bzw. der Katamnese-messung?**

2.4.1. Können Veränderungen in den Skalenwerten der Verlauf-Stichprobe im Zeitraum zwischen Therapiebeginn und der Verlaufsmessung gefunden werden?

***Verlauf-Hypothese:*** Es gibt einen Unterschied in den Gruppenmittelwerten der Skalen, Global Severity Index (GSI), Depression und Angst der Symptom-Checkliste-90, Gesamtskala des Inventars zur Erfassung Interpersonaler Probleme (IIP-D), Gesamtskala der Existenzskala (ESK), Ryff-Skala und Global Assessment of Functioning (GAF), zwischen Therapiebeginn und der Verlaufsmessung danach.

2.4.2. Können Veränderungen in den Skalenwerten der Katamnese-Stichprobe im Zeitraum zwischen Therapiebeginn und der Katamnese-messung gefunden werden?

***Katamnese-Hypothese:*** Es gibt einen Unterschied in den Gruppenmittelwerten der Skalen, Global Severity Index (GSI), Depression und Angst der Symptom-Checkliste-90, Gesamtskala des Inventars zur Erfassung Interpersonaler Probleme (IIP-D), Gesamtskala der Existenzskala (ESK) und Ryff-Skala, zwischen Therapiebeginn und der Katamnese-messung, sechs Monate nach dem Therapieende.

## Empirischer Teil

### Studienbeschreibung

Bei der hier dargestellten und evaluierten *Pilotstudie zur Evaluation der existenzanalytischen Psychotherapie bei Psychotherapeut\*innen in Ausbildung unter Supervision in freier Praxis* handelt es sich um ein prospektives Ein-Gruppen-Studiendesign mit vier Testzeitpunkten. Das laufende Studienprojekt wird von den Forschungsverantwortlichen der Gesellschaft für Logotherapie und Existenzanalyse Österreich (GLE-Ö) mit dem vorrangigen Ziel durchgeführt, die praktische Wirksamkeit existenzanalytischer Ausbildungstherapien zu evaluieren.

Die Studie wurde nach den *Helsinki-Ethikrichtlinien* konzipiert und wird darauf ausgerichtet durchgeführt (World Medical Association, 2013). Es wurde keine Präregistrierung vorgenommen. Durch eine Pilotphase soll abgeklärt werden, wie und mit welchen Mitteln eine zukünftige, großangelegte Praxisstudie durchgeführt werden kann. Weitere Informationen dazu sind auf der Webseite der Einrichtung und bei den Forschungsverantwortlichen zu finden (<https://existenzanalyse.org/service/forschung/>).

Eine Studienteilnahme ist für Auszubildende im psychotherapeutischen Fachspezifikum ab erfolgreich abgeschlossener Prüfung zum klinischen Teil der Ausbildung (therapeutisches Supervisionsstadium) möglich. Nach Studienanmeldung der Psychotherapeut\*innen werden diese gebeten, alle Klient\*innen im Einzelsetting in freier Praxis, nach Erfüllung der Einschlusskriterien, eine Teilnahme an der Studie anzubieten. Ausgeschlossen werden Personen ohne Volljährigkeit und mit intellektueller Beeinträchtigung. Die Studienanwerbung und -teilnahme der Klient\*innen erfolgte ausschließlich über die Psychotherapeut\*innen. Die Therapeut\*innen werden ersucht, in einer Liste zu notieren, wie viele Personen eine Studienteilnahme zugesagt und abgelehnt bzw. die Voraussetzungen nicht erfüllt haben, und sollen dabei Angaben zu deren Geschlecht und Alter machen.

Durch die Schaffung von Anreizen sollte möglichen Vorbehalten von Psychotherapeut\*innen an dem Forschungsprojekten teilzunehmen entgegengewirkt werden. Einerseits gibt es für Therapeut\*innen die Möglichkeit Rückmeldungen in Form von statistischen Auswertungen der individuellen Fragebogenergebnisse der Klient\*innen zu den jeweiligen Testzeitpunkten zu erhalten und andererseits konnte die Abschlussarbeit des psychotherapeutischen Fachspezifikums im Zuge der Studienteilnahme verfasst werden. Dafür sollte auch ein Fall zu Therapiebeginn und -ende interviewt und anschließend nach existenzanalytisch, phänomenologischen Prinzipien analysiert werden.

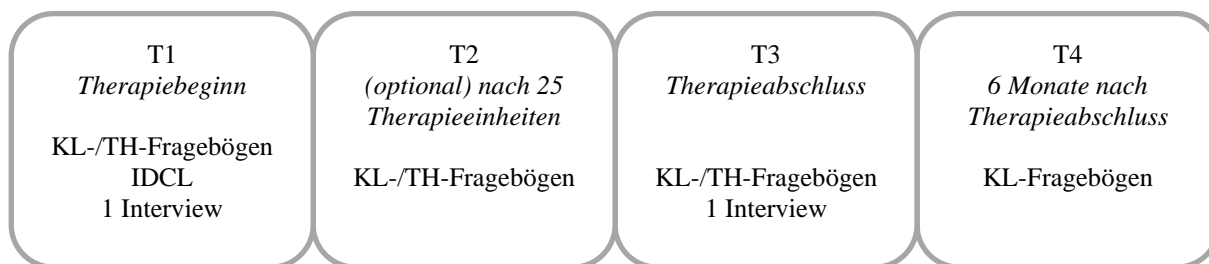
## Studienablauf

Der erste Testzeitpunkt (t1) ist mit Therapiebeginn festgelegt. Klient\*innen, die sich bereit erklären, an der Studie teilzunehmen, bekommen von den Psychotherapeut\*innen den Testbogen mit einer Aufklärung zu den Studienbedingungen und einer Einverständniserklärung (*Informed Consent*) mitgegeben. Nach dem Ausfüllen werden die Testbögen ohne Einverständniserklärung und ohne Realnamen und Unterschrift der Klient\*innen an das Forschungsteam gesandt. Dadurch soll die Anonymität der Teilnehmer\*innen gegenüber Dritten gewahrt bleiben. Die Therapeut\*innen bestätigen mit Unterschrift im Therapeut\*innen-Fragebogen zu Therapiebeginn, dass eine Einverständniserklärung der Klient\*innen vorliegt und diese bei ihren Patient\*innen-Akten verwahrt wird. Mittels pseudonymisierter Codes, die aus Teilen der Namen der Eltern und deren Geburtsjahren bestehen, kann eine Zuordnung zwischen den einzelnen Fragebögen innerhalb der Klient\*innen sowie zu den jeweiligen Psychotherapeut\*innen erfolgen.

Zusätzlich zu den Fragebögen sollen zu Therapiebeginn vorhandene Störungsbilder mittels einer validierten Diagnose-Checkliste (IDCL, Hiller et al., 1995) erhoben werden. Die Therapeut\*innen werden gebeten diese Diagnostik in einer gesonderten Einheit (keine Therapieeinheit) durchzuführen, um keinen finanziellen Nachteil für die Teilnehmer\*innen zu erwirken. Dieses Prozedere sollte innerhalb der ersten fünf Therapieeinheiten stattfinden (vgl. Abbildung 1).

### Abbildung 1

Studiendesign vorliegender Arbeit



*Anmerkung.* Messzeitpunkte und eingesetzte Methoden; KL = Klient\*innen; TH = Therapeut\*innen; IDCL = Internationale Diagnosecheckliste nach ICD-10 (Hiller et al., 1995).

Nach 25 Einheiten erfolgt eine Zwischenerhebung (t2) mittels Fragebögen bei den Klient\*innen, die sich zu diesem Zeitpunkt noch in Psychotherapie befinden. Andernfalls wird zu Therapieende (t3) der dafür vorgesehene Fragebogen mitgegeben (vgl. Abbildung 1). Eine Therapie gilt als abgeschlossen, wenn die krankheitswertige Störung erfolgreich behandelt worden ist. Es besteht die Möglichkeit, die Therapie mit anderen Zielen weiterzuführen. Nach sechs Monaten wird den Therapeut\*innen der Katamnese-Fragebogen

(t4) von den Forschungszuständigen zugesendet. Die Therapeut\*innen werden ersucht, diesen an die Klient\*innen postalisch weiterzuleiten. Ein vorfrankiertes und an das Forschungsteam adressiertes Kuvert ist dem Fragebogen beigelegt, um abermals finanzielle Nachteile für teilnehmende Klient\*innen zu vermeiden.

Für die Abschlussarbeit sollten quantitative Daten von zumindest drei Klient\*innen zu Therapiebeginn und -ende und ein Interview jeweils zu Therapiebeginn und -ende eines Falles vorliegen. Zudem werden die teilnehmenden Psychotherapeut\*innen gebeten, im Zuge ihrer Abschlussarbeit einen Erfahrungsbericht zur Studienteilnahme zu verschriftlichen und diesen auch dem Forschungsteam zukommen zu lassen.

### ***Erhebungsinstrumente***

Um das psychische Befinden der Klient\*innen zu ermitteln, werden Fremdeinschätzungsverfahren bei den Therapeut\*innen und Selbsteinschätzungsfragebögen bei den Klient\*innen eingesetzt. Diese werden im Anschluss kurz beschrieben.

**Internationale Diagnosen Checklisten für ICD -10 (IDCL).** Zur Diagnostik wird von den Psychotherapeut\*innen bis zur fünften Einheit die *Internationale Diagnosen Checkliste für ICD-10*, in Form eines semistrukturierten Interviews, durchgeführt. Die Checkliste dient der reliablen und validen Erhebung von Störungskriterien im Sinne der ICD-10-Klassifikation und wurde von Hiller, Zaudig und Mombour (1995) entwickelt. Die Dauer des diagnostischen Interviews variiert je nach Anzahl an vorliegenden Störungen und dem Frage- bzw. Erzählstil. Pro Fall sollte mit mindestens 20 bis 30 Minuten gerechnet werden. Bei Vorliegen der entsprechenden Diagnosekriterien kann eine Störung diagnostiziert werden. Das Interview ermöglicht eine valide und gleichzeitig ökonomische Diagnoseeinschätzung. Die Checklisten für affektive Störungen der Diagnosegruppe *F3* und Angst- bzw. Anpassungsstörungen der Diagnosegruppe *F4* weisen mit einem Cohen's  $\kappa > .70$  eine hohe Interraterreliabilität auf (Hiller et al., 1995). Detailliertere Informationen zur Anwendung wird den Therapeut\*innen bei der Studieneinschulung und im Zuge von Vernetzungstreffen durch das Forschungsteam vermittelt.

**Global Assessment of Functioning (GAF).** Die GAF-Skala ist ein international weit verbreitetes Instrument um das globale Funktionsniveau einer Person zu erfassen (Aas, 2011; Saß, 2003). Damit soll das psychologische, soziale und berufliche Funktionsniveau der Klient\*innen aus Sicht der Psychotherapeut\*innen eingeschätzt werden. In einem Wertebereich von 0 - 100 kann, zwischen den Polen *Gesundheit* und *Krankheit*, die Ausprägung des Belastungs- bzw. Funktionsniveau beurteilt werden, wobei ein Wert von 100 die bestmögliche Zustandsbeschreibung entspricht. Die Reliabilitäten variieren je nach

Schulung der einschätzenden Personen zwischen Cohen's  $\alpha = .54$  und  $\alpha = .90$ . Die Skala weist eine gute Veränderungssensitivität für Behandlungseffekte auf (Heuft & Lange, 2015).

**Working Alliance Inventory - Short Form revised (WAI-SR).** Das Inventar zur Erfassung der therapeutischen Arbeitsbeziehung zählt zu den am häufigsten eingesetzten Fragebögen, um die therapeutische Arbeitsbeziehung zu ermitteln (Munder, Wilmers, Leonhart, Linster & Barth, 2010). Beim verwendeten Verfahren handelt es sich um eine Kurzform mit zwölf Items. Auf einer fünf-stufigen Likert-Skala von *selten* bis *immer* sollen Aussagen wie „*Mein\*e Therapeut\*in und ich stimmen überein, woran es für mich wichtig ist zu arbeiten*“ eingeschätzt werden. Die deutsche Kurzversion weist eine gute interne Konsistenz mit Cronbach's  $\alpha = .90$  für die Gesamtskala auf (Wilmers et al., 2008). In der Studie wird der Fragebogen bei der Verlaufsmessung (t2) und zu Therapieende (t3) eingesetzt. Als Kennwert der Arbeitsbeziehung wurde der gemittelte WAI-SR-Gesamtwert errechnet, welcher sich aus der Summe aller Items geteilt durch deren Anzahl (12) ergibt. Es wurden nur vollständige Tests verrechnet.

**Ryff-Skala zur Erfassung des psychologischen Wohlbefindens.** Der Fragebogen zur Erhebung des psychologischen Wohlbefindens wird, nach der Entwicklerin, die Ryff-Skala genannt (Ryff, 1989). In einer Weiterentwicklung entstand eine Kurzform mit 18-Items, welche in dieser Studie zum Einsatz kam (Bernsteiner & Boggatz, 2016).

Das *psychologische* Wohlbefinden wird vom *subjektiven* Wohlbefinden abgegrenzt. Das subjektive Wohlbefinden entspricht einer *hedonistischen* Betrachtungsweise und bedeutet weitestgehend das Vorhandensein von positiven Gefühlen wie (allgemeine oder spezielle) Freude und Zufriedenheit bzw. die Abwesenheit von negativem Erleben wie Leidenszustände (Hausler et al., 2017). Das psychologische Wohlbefinden wird dagegen breiter definiert und weist eine *eudaimonische* Betrachtungsweise auf (Bosch, 2019; Hausler et al., 2017). Neben positiven Erlebenszustände werden dazu auch persönliches Wachstum, Einstellung zum Leben, Haltung gegenüber sich und anderen gezählt und weist somit Ähnlichkeiten zu den Konzepten der Existenzskala auf (Längle, Orgler & Kundi, 2000; Ryff & Keyes, 1995).

Mit der Ryff-Skala können die sechs Subskalen mit je drei Items nämlich *Selbstakzeptanz, Beherrschung der Umgebung, positive Beziehungen zu anderen, Lebensziele, persönliches Wachstum und Autonomie* gebildet werden. Jeder Aussage kann auf einer sechs-stufigen Skala zwischen *ganz falsch* bis *ganz richtig* zugestimmt werden. Eine Aussage lautet z.B.: „*Ich habe normalerweise das Gefühl, dass ich mit meiner Situation gut zurechtkomme*“. Die interne Konsistenz ist mit Cronbach's  $\alpha$  von .33 bis .56 für die einzelnen Skalen als inakzeptabel bis schlecht einzustufen. Für die Gesamtskala der Kurzform wurden keine

Angaben gefunden. Jedoch korreliert die Kurzform mit der ursprünglichen Langform (120 Items) mit einem  $r$  von .70 bis .80 sehr hoch, was wiederum für eine gute Konstruktvalidität spricht (Ryff & Keyes, 1995). In vorliegender Arbeit wird die Gesamtsumme der Ryff-Skala als Maß des psychologischen Wohlbefindens der Klient\*innen herangezogen. Es wurden nur vollständige Tests verrechnet.

**Existenz-Skala (ESK).** Die ESK erhebt die erlebte existenzielle Erfüllung bzw. die Ausprägung an wahrgenommener Bedeutsamkeit im Leben einer Person. Darunter ist die Fähigkeit zu verstehen, mit sich und der Welt umgehen zu können (Längle, Orgler & Kundi, 2000). Theoretisch wurde die ESK nach dem Sinnkonzept von Frankl entwickelt (Längle, 2014). Auf einer sechsstufigen Skala sollen Äußerungen wie „*Es gibt eigentlich nichts in meinem Leben, wofür ich mich richtig einsetzen möchte.*“ oder „*Die Frage, ob ich etwas tun will, stellt sich selten, weil ich meistens etwas tun muss.*“ zwischen den Polen *stimmt* und *stimmt nicht* eingeschätzt werden. Der Fragebogen umfasst 46 Items und ermöglicht es die vier Subskalen Selbstdistanzierung, Selbsttranszendenz, Freiheit und Verantwortung zu bilden. Aus diesen Skalen können drei Summenskalen gebildet werden: Die *Personalität* (Cronbach's  $\alpha = .87$ ), die sich aus den Skalen Selbstdistanzierung und -transzendenz ergibt und die *Existenzialität* (Cronbach's  $\alpha = .90$ ), die sich aus den Skalen Freiheit und Verantwortung ergibt. Aus beiden Summenskalen kann die Gesamtskala gebildet werden. Die Gesamtskala gibt die wahrgenommene Fähigkeit einer Person an, eine für sich bedeutsame und kongruente Lebensgestaltung mit den Angeboten und Forderungen der persönlichen Lebensrealität führen zu können. Je höher dieser Wert ist, desto erfüllter erlebt sich eine Person. Eine interne Konsistenz von Cronbach's  $\alpha = .93$  für die Gesamtskala kann als exzellent eingestuft werden (Längle, Orgler & Kundi, 2000). Für die Berechnung der Gesamtskala dürfen höchstens zwei fehlende Werte (5 %) vorliegen.

**Symptom-Checkliste-90-Standard (SCL-90).** Die Symptomcheckliste von Franke (2014) erfasst anhand 90 Items die subjektiv empfundene psychische und physische Beeinträchtigung der letzten sieben Tage. Der Fragebogen gilt als transdiagnostisches Breitbandverfahren, da mehrere Störungsbilder gescreent werden können. Es können neun Subskalen, nämlich *Somatisierung, Zwanghaftigkeit, Unsicherheit im Sozialkontakt, Depressivität, Ängstlichkeit, Aggressivität, Phobische Angst, Paranoides Denken und Psychotizismus*, gebildet werden. Für vorliegende Arbeit werden die Subskalen Ängstlichkeit (ANGS) und Depressivität (DEPR) als Kennwerte für die häufigsten Symptomcluster in der psychotherapeutischen Zielpopulation verwendet (Cramer et al., 2014; Laireiter et al., 2013; Steffanowski et al., 2011; Wittmann & Steffanowski, 2011). Die Subskala *Ängstlichkeit*

besteht aus zehn Items und erhebt Verhaltensweisen und Symptome, die mit Angst verbunden sind. Die interne Konsistenz der Skala ist mit Cronbach's  $\alpha = .88$  als gut zu bezeichnen. Die Subskala *Depression* ergibt sich aus dreizehn Items, die die Bandbreite klinisch depressiver Symptome erheben. Die Skala weist mit Cronbach's  $\alpha = .89$  ebenso eine gute Reliabilität auf.

Neben den beiden globalen Kennwerten, *Positive Symptom Distress Index* (PSDI) und *Positive Symptom Total* (PST), kann ein globaler Index, der *Global Severity Index* (GSI), gebildet werden. Dieser errechnet sich aus dem gemittelten Summenscore aller 90 Items und misst die global erlebte Belastung einer Person über alle Subskalen hinweg. Ein Cronbach's  $\alpha$  von .98 spricht für eine exzellente interne Konsistenz (Franke, 2014). In vorliegender Arbeit wird als Kennwert für die allgemeine Symptombelastung der *GSI* verwendet. Laut Manual können Skalen mit maximal drei fehlenden Werten auf der Ängstlichkeitsskala, maximal vier fehlenden Werten auf der Depressionsskala und maximal 18 fehlenden Werten auf der GSI-Skala verrechnet werden.

**Inventar zur Erfassung interpersonaler Probleme (IIP-D).** Das Inventar zur Erfassung interpersonaler Probleme erhebt Schwierigkeiten im Umgang mit anderen Menschen bzw. interpersonale Verhaltensweisen, die schwer fallen oder im Übermaß vorhanden sind (Becker & Mohr, 2005; Horowitz, Strauss & Kordy, 2000). Das Verfahren besteht aus 64 Items, dessen Aussagen auf einer fünfstufigen Likert-Skala von *gar nicht* bis *sehr* zugestimmt werden kann. Es können acht Subskalen (*zu autokratisch, zu streitsüchtig, zu abweisend, zu introvertiert, zu unterwürfig, zu ausnutzbar, zu fürsorglich* und *zu expressiv*) errechnet werden. Ein gemittelter Summenwert über alle Items hinweg bildet den IIP-Gesamtwert, der als Kennwert für das generelle Ausmaß an interpersonalem Problemen beschrieben wird (Horowitz et al., 2000). In vorliegender Arbeit wird der Gesamtwert als Maß für interpersonale Probleme herangezogen.

Für die Subskalen liegen interne Konsistenzen zwischen Cronbach's  $\alpha = .71$  bis  $.82$  vor und sind somit akzeptabel bis gut. Für das Gesamtmaß wurden keine Informationen gefunden. Laut Manual wird der Grenzwert für fehlende Items bei fünf Prozent aller Items angesetzt. Es werden daher nur Tests verwendet, die maximal drei fehlende Werte aufweisen (Horowitz et al., 2000).

**Assessment of DSM-IV Personality Disorders (ADP-IV).** Beim ADP-IV handelt es sich um ein Selbstbeurteilungsinstrument zur Erfassung möglicher Persönlichkeitsstörungen. Für eine seriöse Diagnosestellung von Persönlichkeitsstörungen werden klinische Interviews empfohlen (Renn, Höfer, Schüßler, Rumpold, Smrekar, Janecke & Doering, 2008). Bei vorliegender Studie kommen 35 Items zum Einsatz, die Cluster-B-Persönlichkeitsstörungen

screenen (*antisoziale, Borderline, histrionische und narzisstische Persönlichkeitsstörung*). Die internen Konsistenzen können mit Werten zwischen *Cronbach's  $\alpha = .72$*  (histrionisch) bis *Cronbach's  $\alpha = .85$*  (antisozial) bzw. *Cronbach's  $\alpha = .91$*  für den gesamten Cluster-B als akzeptabel bis exzellent eingestuft werden. Die Fragen werden auf einer 7-Stufigen Skala beantwortet. Wird dabei ein Wert ab fünf angekreuzt, wird ersucht ein zusätzliches dreistufiges Distress-Rating zu dieser Frage auszufüllen (Doering, Renn, Höfer, Rumpold, Smrekar, Janecke, Schatz & Schotte, 2007).

### ***Erhebung demographischer und Hintergrunddaten***

Den teilnehmenden Klient\*innen wird neben den psychometrischen Testverfahren zu jedem Testzeitpunkt ein allgemeiner Fragebogen vorgegeben. Darin werden zum einen soziodemographische Daten und zum anderen Informationen zu Medikamenteneinnahmen, dem Vorliegen von Krankheiten, früherer Inanspruchnahme psychosozialer Angebote und relevante Fragen zur Studienakzeptanz, wie der Meinung zur Fragebogenlänge, zur Konzentrationsfähigkeit und die Gründe der Studienteilnahme erhoben. Zusätzlich werden die Klient\*innen auch nach möglichen negativen Veränderungen durch die Therapie, wichtigen Ereignissen außerhalb der Therapie und auch nach Besonderheiten, die sie für den Verlauf der Therapie für wichtig erachten, gefragt. Diese Fragen können offen beantwortet werden. Zu Beginn und am Ende der Fragebögen wird außerdem die momentane Stimmung auf einer sieben-stufigen Skala zwischen den Polen *schlechteste mögliche* und *bestmögliche* Stimmung abgefragt. Damit sollen unmittelbare Effekte auf die Stimmung durch das Ausfüllen des Fragebogens ermittelt werden.

Auch die Therapeut\*innen werden gebeten zu jedem Fall und Testzeitpunkt (außer bei der Katamnesemessung) einen Fragebogen auszufüllen. Darin wird gebeten die GAF-Skala einzuschätzen und offene Fragen nach möglichen Besonderheiten für den Psychotherapieverlauf zu beantworten. Bei der Verlaufsmessung und zu Therapieende wird zusätzlich eine Einschätzung zum bisherigen Behandlungserfolg, zur Zufriedenheit mit, und zur Zielerreichung im Zuge der Psychotherapie anhand einer Ordinalskala erbeten. Zu Therapieende wird den Psychotherapeut\*innen Fragen zum Abschluss der Therapie gestellt. Es wird erhoben, ob es sich um ein reguläres oder vorzeitiges Therapieende gehandelt hat. Bei Abbrüchen wird der Grund erfragt. Zudem wird die Anzahl der Einheiten erhoben.

***Qualitative Daten.*** Neben den offenen Fragen in den allgemeinen Klient\*innen- und Therapeut\*innenfragebögen wurden den teilnehmenden Therapeut\*innen sechs Fragen zur Studiendurchführung gestellt, die sie am Ende Ihrer Teilnahme schriftlich beantworten und dem Forschungsteam zukommen lassen sollten. Die Fragen sind im Folgenden angeführt:

- Wie gut ließ sich die Studie in den therapeutischen Alltag integrieren?
- Wie waren die Reaktionen der Klient\*innen auf das Angebot, bei der Studie mitmachen zu können?
- Wie waren (falls geäußert) die Meinungen der Klient\*innen zu den Fragebögen?
- Haben Sie die Möglichkeit in Anspruch genommen, eine Rückmeldung zur Diagnostik zu erhalten? Wenn ja, wie haben Sie die Rückmeldung verwertet und war diese hilfreich im therapeutischen Prozess?
- Wo hatten Sie Schwierigkeiten? Hatten Sie stets eine Ansprechperson, die Ihnen bei Ihren Schwierigkeiten helfen konnte?
- Was kann optimiert werden?

### **Datenanalyse**

Sämtliche statistische Berechnungen wurden mit *IBM SPSS 25* durchgeführt. Das Signifikanzniveau wurde bei allen statistischen Tests dieser Arbeit auf  $\alpha = .05$  festgelegt. Bei mehrfachen (t-) Tests wurde der p-Wert mit der als konservativ geltenden *Bonferroni-Korrektur* adjustiert (Armstrong, 2014; Bortz & Schuster, 2010).

### ***Ermittlung der Studiendurchführbarkeit und -akzeptanz***

Die Betrachtung der Rücklaufquote und der Vergleich der Gesamtstichprobe der Therapeut\*innen und Klient\*innen mit Stichproben aus vergleichbaren Studien sollten eine Einschätzung über die erreichte Personengruppe und deren Repräsentativität für das Interessensfeld ermöglichen. Zudem wurde eine vorläufige Abbruchsanalyse durchgeführt. Informationen über Abbruchsfälle konnten dem Therapeut\*innenfragebogen zu Therapieende und vereinzelt den Listen der Therapeut\*innen zum Klient\*innenrücklauf entnommen werden. Es wurde eine Unterteilung in *qualitätsrelevante* bzw. *nicht-qualitätsrelevante* Abbrüche durchgeführt. Qualitätsrelevante Abbrüche stehen in Zusammenhang mit der Qualität der Therapie. Das können interaktionelle Probleme, Unzufriedenheit mit dem Fortschritt, plötzliches Fernbleiben etc. sein. Dagegen haben nicht-qualitätsrelevante Abbrüche einen unvermeidbaren Charakter, wie ein Wohnortwechsel, eine Schwangerschaft bzw. Geburt, ein stationärer Aufenthalt, eine Kur oder ein Rehabilitationsaufenthalt usw. (Hiller et al., 2011). Neben der Abbruchsanalyse wird der vorläufige Stand an Datensätzen zu den jeweiligen Testzeitpunkten berichtet.

Die Meinungen der Klient\*innen zu den Fragebögen bezüglich Länge, Dauer und Verständlichkeit sowie ihre Angaben zum Teilnahmegrund werden deskriptiv beschrieben. Um einen möglichen statistisch signifikanten Unterschied zwischen der Stimmung unmittelbar vor und nach dem Ausfüllen der Fragebögen zu untersuchen, wurde der non-

parametrischer *Wilcoxon-Rang-Test* durchgeführt. Ergänzend fließen zur Beantwortung der Frage nach Akzeptanz und Durchführbarkeit die Ergebnisse einer qualitativen Analyse der Erfahrungsberichte der Therapeut\*innen ein, die auch nach den geäußerten Meinungen der Klient\*innen zu den Fragebögen gefragt wurden. Auf das Vorgehen, der dafür verwendeten qualitativen Inhaltsanalyse nach Kuckartz (2018), wird im Folgenden kurz eingegangen.

**Qualitative Inhaltsanalyse nach Kuckartz.** Zur Analyse der schriftlichen Erfahrungsberichte der Therapeut\*innen kam die Methode der *inhaltlich-strukturierenden Analyse* nach Kuckartz (2018) zum Einsatz. Dabei sollen aus dem Text Inhalte identifiziert und konzeptualisiert und zu Kategorien zusammengefasst werden. Im Gegensatz zu Mayring (2015), der ein theoriengeleitetes (deduktives) Vorgehen beschreibt, tritt Kuckartz (2018) für eine induktive und deduktive, also gemischte, Form der Entwicklung von Kategorien ein und betont stärker das Arbeiten direkt am Material (Kuckartz, 2018; Mayring, 2015). Diese Forderung wird im Phasenmodell des Analyseprozesses verdeutlicht. *Phase I* beschreibt eine initiierte Textarbeit mit dem Markieren wichtiger Textstellen und dem Schreiben von Memos. In *Phase II* werden thematischen Hauptkategorien entwickelt, um in *Phase III* das gesamte Material anhand der Hauptkategorien zu codieren. In *Phase IV* werden die codierten Textstellen mit gleichen Kategorien zusammengefasst, um daraus in *Phase V* Subkategorien zu bestimmen. Ein erneuter Kodier-Prozess in *Phase VI* soll ein weiteres Ausdifferenzieren von Kategorien prüfen. In *Phase VII* erfolgt die endgültige Auswertung des bearbeiteten Materials. Die Befunde der inhaltlich-strukturierenden Analyse sollen als qualitative Ergänzung der quantitativen Daten dienen. Dabei werden die qualitativen Ergebnisse nicht gesammelt, sondern in den jeweiligen thematischen Kapiteln gesondert berichtet und in Bezug mit den quantitativen Daten gesetzt. So werden beispielsweise die von den Therapeut\*innen angeführten Themen zum Erleben der Klient\*innen beim Ausfüllen des Fragebogens im Abschnitt der Klient\*innenstichprobe angeführt.

Bei der qualitativen Analyse der therapeutischen Erfahrungsberichte handelt es sich um eine ergänzende Nebenanalyse zur Klärung der Akzeptanz des Studiendesigns. Im Gegensatz zu Mayring (2015) ermöglicht die Analyse nach Kuckartz (2018) eine freiere Auswertung direkt am Material, weswegen auf dessen Analysemethode zurückgegriffen wurde.

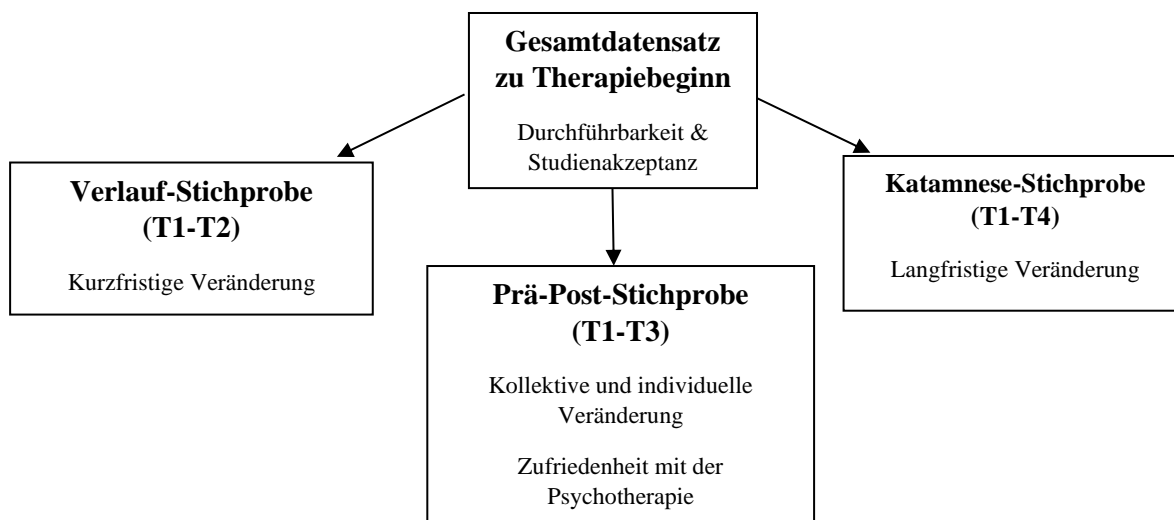
### ***Gruppenstatistische Veränderung***

Durch das Vorliegen einer unterschiedlichen Anzahl an Datensätzen zu den unterschiedlichen Testzeitpunkten wurden für jeden Messzeitpunkt Teilstichproben aus dem Gesamtdatensatz gezogen, welche die vollständigen Fälle zu den jeweiligen Zeitpunkten

beinhalteten. Es entstanden so drei Stichproben mit verwertbaren Datensätzen (*Available-Case-Analyse*), die in Abbildung 2 veranschaulicht sind.

**Abbildung 2**

*Gezogene Teilstichproben für die Analyse*



**Stichprobenunterschiede.** Um mögliche statistisch signifikante Unterschiede der einzelnen Stichproben mit der restlichen Gesamtstichprobe zu ermitteln, wurden Gruppenvergleiche hinsichtlich demographischer Variablen wie dem Alter, dem Geschlecht und dem formalen Bildungsstand sowie den erhobenen Skalenwerten zu Therapiebeginn durchgeführt. Dafür wurde für metrische Variablen ein *t-Test für unabhängige Stichproben* und für kategoriale Variablen der *Chi<sup>2</sup>-Test* angewandt.

Die Voraussetzungen für einen unabhängigen t-Test sind einerseits die Varianzhomogenität in beiden Gruppen, die miteinander verglichen werden und eine Normalverteilung des Merkmals in beiden Gruppen. Ab einer Gruppengröße von 30 Personen ist aufgrund des allgemeinen Grenzwerttheorems die Normalverteilungsvoraussetzung weniger relevant (Bortz & Schuster, 2010). Wichtiger ist die Varianzhomogenität, die mit dem *Levene-Test* überprüft werden kann. Bei fehlender Varianzhomogenität wird der *Welch-Test* zur Überprüfung der Unterschiede der Gruppenmittelwerte herangezogen. Einige Autor\*innen empfehlen, unabhängig vom Vorliegen von Varianzhomogenität, immer den Welch-Test statt des „normalen“ t-Tests zu verwenden, da dieser bei Varianzhomogenität eine vergleichbare und bei fehlender Homogenität der Varianzen eine höhere Testmacht aufweist (Kubinger, Rasch & Moder, 2009). Im Ergebnisteil wurden die Voraussetzungsüberprüfungen dennoch berichtet und im Fall homogener Varianzen der t-Test verwendet.

Der  $\chi^2$ -Test (Chi-Quadrat-Test) kann zur Überprüfung einer möglichen Abhängigkeit zweier oder mehrerer kategorialer Variablen eingesetzt werden. In vorliegendem Fall sollte er

prüfen, ob ein statistisch signifikanter Unterschied zwischen den Häufigkeiten soziodemografischer Daten wie Geschlecht oder Bildungsstand in den Teilstichproben und der restlichen Gesamtstichprobe vorlag. Für den  $\chi^2$ -Test sollen die Beobachtungen voneinander unabhängig sein und die Anzahl der erwarteten Häufigkeiten pro Zelle sollte größer gleich fünf sein. Ist dies nicht der Fall sollte für den Signifikanztest nicht die *asymptotische Signifikanz* des Chi-Quadrat-Tests nach Pearson verwendet werden, sondern die *exakte Signifikanz* des exakten Tests nach Fisher (Bortz & Schuster, 2010).

Der *Wilcoxon-Test* wurde zum Vergleich der Unterscheide auf der ordinalen Stimmungsskala eingesetzt, die unmittelbar vor und nach dem Ausfüllen des Fragebogens abgefragt wird. Dadurch sollen mögliche Einflüsse des Ausfüllens auf die unmittelbare Stimmung der Klient\*innen untersucht werden. Beim Wilcoxon-Test handelt es sich um ein non-parametrisches Verfahren, das bei ordinalen Variablen angewendet wird (Bortz & Schuster, 2010).

**Untersuchung von Therapieeffekten.** Um mögliche Effekte der Therapie zu ermitteln, wurden die Gruppenmittelwerte zu Therapiebeginn und dem jeweiligen Messzeitpunkt auf Unterschiede hin untersucht. Zur Überprüfung der Mittelwertsdifferenzen auf statistische Signifikanz wurde der *t-Test bei verbundenen Stichproben* verwendet. Stellte sich dieser Differenzwert als statistisch signifikant heraus ( $p < .05$ ), konnte von einer statistisch signifikanten Veränderung der untersuchten Stichprobe ausgegangen werden. Dabei handelt es sich um eine indirekte Veränderungserfassung (Stieglitz & Hiller, 2014). Die Voraussetzungen für die adäquate Durchführung eines abhängigen t-Tests sind die Abhängigkeit der Messungen und die Normalverteilung der Differenzwerte der Untersuchungspaare. Der t-Test für Beobachtungspaare ist jedoch robust gegenüber Verletzungen der Voraussetzungen. Ähnlich wie beim t-Test für unabhängige Stichproben haben beim abhängigen t-Test ab einer Stichprobengröße von 30 Personen Voraussetzungsverletzungen der Normalverteilung wenig Einfluss auf das Ergebnis (Bortz & Schuster, 2010). Im Ergebnisteil wurden die Voraussetzungsüberprüfungen dennoch berichtet.

**Multiple Testungen.** Die mehrfache Durchführung von zusammenhängenden (t-) Tests zur Überprüfung einer übergeordneten Hypothese führt zu einer Kumulierung des  $\alpha$ -Fehlers jedes einzelnen Tests, was die Wahrscheinlichkeit von Fehlern 1. Art erhöht, also dem fälschlichen Ablehnen der Nullhypothese (Bortz & Schuster, 2010). Beim t-Test würde ein Gruppenmittelwertsunterschied angenommen werden, obwohl keiner vorhanden ist. Aus diesem Grund sollte der p-Wert angepasst werden. Die *Bonferroni-Korrektur* ist eine

konservative statistisches Methode, um die Alphafehlerkumulierung auszugleichen und wird in vorliegender Arbeit bei Mehrfach-Tests angewandt. Das korrigierte Alpha ergibt sich aus dem ursprünglichen Alpha geteilt durch die Anzahl der ausgeführten Tests (Armstrong, 2014; Bortz & Schuster, 2010).

**Effektstärken.** Bei statistisch signifikanten Ergebnissen wurde die Stärke des Effektes nach Cohen's  $d$  und deren 95 %-Konfidenzintervall angegeben. Die Effektstärke nach Cohen (1988) errechnet sich nach der Formel:

$$d = \frac{\bar{x}_{TE} - \bar{x}_{TB}}{SD_{TB}}$$

Wobei  $\bar{x}_{TE}$  der Mittelwert zu Therapieende,  $\bar{x}_{TB}$  der Mittelwert zu Therapiebeginn und  $SD_{TB}$  die Standardabweichung zu Therapiebeginn darstellen (Cohen, 1988). Effektstärken ermöglichen eine Vergleichbarkeit von Forschungsergebnissen über verschiedene Studiendesigns und Verfahren hinweg (Jacobi, 2020). Es werden Werte von Cohen's  $d < 0.2$  als sehr kleine Effekte,  $d < 0.5$  als kleine Effekte, Werte  $d > 0.5$  als mittlere Effekte und Werte  $d > 0.8$  als große Therapieeffekte interpretiert (Cohen, 1988; Hiller et al., 2011). Eine Effektstärke von Cohen's  $d = 1.0$  würde bedeuten, dass sich der Mittelwert des Klient\*innenkollektives um eine Standardabweichung verändert hat (Jacobi, 2020). Auf die Angabe der Richtung des Effekts durch ein Minuszeichen wird verzichtet.

**Testpower.** Eine mit *G\*Power* durchgeführte a-priori-Poweranalyse ergab, dass beim t-Test mit abhängigen Stichproben und bei einer zweiseitigen Testung, einer 95 %-Alpha-Fehlerwahrscheinlichkeit, einer erwünschten Teststärke von  $(1-\beta) = .95$  und zu erwartbaren, mittleren Effektstärke von  $d = 0.5$ , ein Stichprobenumfang von 54 Personen benötigt wird. Bei den aus der Literatur abgeleiteten zu erwartbaren Effektstärken von  $d = 0.8$  wird ein Stichprobenumfang von  $n = 23$  benötigt, um beim t-Test mit abhängigen Stichproben eine Testpower von  $(1-\beta) = .95$  zu erzielen (Faul, Erdfelder, Lang & Buchner, 2007).

**Umgang mit fehlenden Werten.** Die Analyse fehlender Datenpunkte auf Ebene der errechneten Skalen auf Systematik erfolgte über alle Testzeitpunkte hinweg mit Einschluss aller verwendeter Skalen und soziodemographischer Daten. Der *Little's MCAR-Test* nach Little ergab kein systematisches Fehlen von Daten,  $\chi^2(337) = 367.69, p = .12$  (Little, 1988). Fehlende Skalenwerte wurden nicht durch die Ausprägung oder das Fehlen anderer, in die Analyse eingeschlossener, Variablen bestimmt.

Für die Analyse dieser Arbeit werden nur vollständige bzw. laut Manualen verwertbare Skalen in den jeweiligen Instrumenten einbezogen (vgl. Kapitel Erhebungsinstrumente). Dabei handelt es sich um eine *Available-Case-Analyse*, bei der es zu einer unterschiedlichen Anzahl an verrechenbaren Skalen auch innerhalb einer Stichprobe

kommen kann (Little & Rubin, 2020). Aufgrund der ausreichenden Stichprobengröße für die vorgesehenen statistischen Analysen und einem zufälligen Fehlen von Datenpunkten, dessen Ausmaß sich innerhalb der Stichproben unter und knapp um 5 % bewegt, kann davon ausgegangen werden, dass die Ergebnisse davon nicht beeinflusst sind und auch ohne Datenimputation eine ausreichende Generalisierbarkeit für die Stichprobe aufweisen (Jakobsen et al., 2017; Jensen, Mortensen & Lotz, 2017; Madley-Dowd, Hughes, Tilling & Heron, 2019).

### ***Veränderung auf individueller Ebene***

Hohe Effektstärken können von starken Veränderungswerten weniger Personen beeinflusst werden und geben einen Überblick über die tendenzielle Entwicklung einer Gruppe. Ein statistisch signifikantes Ergebnis einer Gruppe lässt keinen Schluss darüber zu, bei wie vielen Personen dieser Gruppe sich das Befinden reliabel bzw. klinisch bedeutsam verändert hat (Jensen & Corralejo, 2017; Senft, Fischer-Hansal & Schosser, 2020). Zur Beurteilung der Wirksamkeit einer Therapie empfiehlt sich daher, zusätzlich zu gruppenstatistischen Analysen, auch Veränderungswerte auf individueller Ebene zu ermitteln (Jensen & Corralejo, 2017; Lambert & Ogles, 2009). Um das Ausmaß der Veränderung einzelner Klient\*innen zu bestimmen, wurde der *Reliable Change Index (RCI)* verwendet. Mit dem RCI kann der Anteil jener Klient\*innen ermittelt werden, deren Zustand sich zwischen zwei Testzeitpunkten, kontrolliert um den Messfehler des Erhebungsinstrumentes, entweder nicht verändert, reliabel gebessert oder verschlechtert hat (Morley & Dowzer, 2014). Dabei wird die Differenz der Messwerte einer Person zu, beispielsweise, Therapieende ( $x_{TE}$ ) und Therapiebeginn ( $x_{TB}$ ) ins Verhältnis zum Standardfehler beider Werte ( $S_{diff}$ ) gesetzt. Der Standardfehler leitet sich von der Reliabilität des Verfahrens ( $r_\alpha$ ) und der Standardabweichung einer Referenzpopulation ( $S_E$ ) ab (Jacobson & Truax, 1991; Senft et al., 2020):

$$RC = \frac{x_{TE} - x_{TB}}{S_{diff}} \quad \text{wobei} \quad S_{diff} = \sqrt{2(S_E)^2} \quad \text{und} \quad S_E = SD \cdot \sqrt{1 - r_\alpha}$$

Neben einer *reliablen Veränderung* kann eine Person auch eine *klinisch signifikante Veränderung* aufweisen. *Klinische Signifikanz* liegt dann vor, wenn sich zusätzlich zu einer reliablen Veränderung auch die Symptomausprägung einer Person aus einem klinisch auffälligen Wertebereich (zu Therapiebeginn) zu einem klinisch unauffälligen Wertebereich (z.B. zu Therapieende) bewegt. Die klinische Signifikanz hängt somit einerseits vom Ausmaß der individuellen Veränderung und andererseits von der Stärke der Symptomausprägung zu Therapiebeginn und -ende ab (Jacobi, 2020). Es müssen somit zwei Kriterien bzw. Schwellenwerte erreicht werden, um klinische Signifikanz aufzuweisen. Der Schwellenwert für ein klinisch auffälliges bzw. unauffälliges Testergebnis wird anhand zweier ausreichend

großer Stichproben aus pathologischen und gesunden Personen errechnet (Hansen & Lambert, 1996).

**Tabelle 2**

*Cut-Off-Werte für klinische Auffälligkeit der symptombezogenen Skalen*

Skala	Mindestausmaß für eine reliable Veränderung	Schwellenwerte für klinisch bedeutsame Veränderung
GSI <sup>a</sup>	+/- 0.26	0.75
DEPR <sup>a</sup>	+/- 0.64	0.96
ANGS <sup>a</sup>	+/- 0.69	0.67
IIP <sup>b</sup>	+/- 0.38	1.21

*Anmerkung.* GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskaala der SCL-90; IIP = Inventar zur Erfassung interpersonaler Probleme; a. Skalen der SCL und RCI-Maße nach Franke (2014); b. Schwellenwert und kritische Differenz der IIP-Gesamtskala nach Berechnungen von Hanse und Lambert (1996).

Der RCI wurde in vorliegender Arbeit für die Klient\*innen der Prä-Post-Stichprobe (T1-T3) berechnet. Dafür wurden die Werte der symptombezogenen Skalen der SCL und des IIP verwendet. Das für eine reliable Veränderung notwendige Mindestausmaß der Skalendifferenz (kritische Differenz) und der Schwellenwert für eine klinisch (un-)auffällige Symptomausprägung wurden für die Skalen der SCL dem Verfahrensmanual (Franke, 2014) und für die Gesamtskala des IIP einer Studie von Hansen und Lambert (1996) entnommen und sind in Tabelle 2 dargestellt. Für die Berechnungen wurden Personen mit klinisch auffälliger Symptomatik gesondert von Personen mit klinisch unauffälliger Symptomatik in den jeweiligen Skalen zu Therapiebeginn ausgewertet. Das Befinden von Personen mit einer klinisch unauffälligen Symptomatik zu Therapiebeginn kann sich nicht klinische signifikant, jedoch reliabel bessern bzw. verschlechtern. Die Verläufe der Personen wurden den Kategorien *verschlechtert*, *unverändert*, *reliabel gebessert* und *klinisch signifikant gebessert* zugeordnet (Senft et al., 2020). In Studien werden vorwiegend die Ergebnisse der signifikant gebesserten (Response) und klinisch signifikant gebesserten Personen (Remission) angegeben (Hiller, Schindler, & Lambert, 2011; Victor et al., 2018).

### ***Zufriedenheit mit der Psychotherapie***

In den Therapeut\*innen- und Klient\*innenfragebögen wurde eine *direkte Veränderungsmessung* eingesetzt (Stieglitz & Hiller, 2014). Hierfür wurden den Klient\*innen und Therapeut\*innen Fragen zur Therapiebeurteilung gestellt, die den Behandlungserfolg, die Zielerreichung und die Zufriedenheit zum Thema haben. Diese Daten werden mittels deskriptiver Statistik aufbereitet.

## Ergebnisse

### *Teilnahmebereitschaft und Studienakzeptanz der Therapeut\*innen*

**Rücklauf der Therapeut\*innen.** Es wurde keine systematische Erhebung all jener Therapeut\*innen durchgeführt, die angefragt wurden, an der Studie teilzunehmen. Somit liegen keine Daten über den Rücklauf der Psychotherapeut\*innen vor. Beworben wurde die Studie in regelmäßigen Abständen per schriftlichen Aussendungen via E-Mail, mittels mündlicher Anwerbung auf Kongressen und durch Vorträge in den Ausbildungsseminaren der Gesellschaft für Logotherapie und Existenzanalyse. Außerdem wurden die Lehrtherapeut\*innen bzw. Ausbildungsleiter\*innen gebeten, in ihren Seminaren von der Möglichkeit einer Studienteilnahme im Rahmen der Abschlussarbeit zu berichten.

Als Referenzrahmen können die 224 Psychotherapeut\*innen in Ausbildung unter Supervision der GLE-Ö zum Stichtag 1.6.2020 herangezogen werden (Sagerschnig & Valady, 2021). Mit 37 Anmeldungen zur Studie konnten rund 17 % der existenzanalytischen Psychotherapeut\*innen im Supervisionsstatus erreicht werden. Der prozentuelle Anteil könnte noch größer sein, wenn nur jene Psychotherapeut\*innen gezählt werden, die in freier Praxis tätig sind, worüber aber keine Informationen gefunden wurden.

**Studienabbruch von Therapeut\*innen.** Von insgesamt 37 angemeldeten Therapeut\*innen hatten bislang 16 Personen die Studie abgeschlossen, acht Therapeut\*innen hatten die Studienteilnahme abgebrochen. 13 Therapeut\*innen haben einen laufenden Status inne. Von den acht abgebrochenen hatten drei Therapeut\*innen Datensätze von jeweils zwei Klient\*innen geliefert, von diesen insgesamt sechs Klient\*innen wiesen alle ebenfalls einen Abbruchsstatus auf. Bei den restlichen fünf Therapeut\*innen mit Abbruchsstatus lagen keine Daten von Klient\*innen vor. Die Geschlechterverteilung der Therapeut\*innen, die die Studienteilnahme abgebrochen hatten, lag bei 75 % Frauen ( $n = 6$ ) und 25 % Männer ( $n = 2$ ). Das Durchschnittsalter lag in dieser Gruppe bei  $M = 43.1$  Jahren ( $SD = 14.0$ ).

### *Beschreibung der Therapeut\*innenstichprobe*

Von den verbliebenen 29 Therapeut\*innen weisen 13 einen laufenden (45 %) und 16 einen abgeschlossenen Studienstatus (55 %) auf. Die Geschlechterverteilung liegt bei 80 % weiblichen ( $n = 23$ ) bzw. 20 % männlichen ( $n = 6$ ) Teilnehmer\*innen. Der Altersdurchschnitt beträgt  $M = 41,7$  Jahren ( $SD = 8.7$ ). 18 Personen (67 %) gaben einen Universitätsabschluss, sechs Therapeut\*innen (22 %) eine Lehre ohne Matura und drei Teilnehmer\*innen (11 %) einen Maturaabschluss als ihre höchste, abgeschlossene Ausbildung an.

Knapp über die Hälfte der Therapeut\*innen (52 %,  $n = 15$ ) kamen aus dem Raum Vorarlberg und Tirol, gefolgt von der Steiermark mit 17 % ( $n = 5$ ) und Wien mit 10 % ( $n = 3$ ).

Die restlichen 14 % ( $n = 4$ ) verteilten sich auf die Regionen Ober- und Niederösterreich ( $n = 3$ ) bzw. Kärnten / Osttirol ( $n = 1$ ). Diese Informationen wurden über die angegebenen Postleitzahlen ermittelt.

Als häufigste Quellberufssparte wurde mit 17 % Psychologie angegeben ( $n = 5$ ), gefolgt von Wirtschaft mit 14 % ( $n = 4$ ) und Soziale Arbeit bzw. Pädagogik mit jeweils 10 % der Therapeut\*innenstichprobe ( $n = 3$ ). Weitere Quellberufe waren Krankenpflege, andere Gesundheitsberufe, Beratung und Lehrberuf mit jeweils 7 % ( $n = 2$ ) bzw. Ärzt\*innenberuf, Sprachwissenschaft und Technik mit jeweils 3 % ( $n = 1$ ). In drei Fällen (10 %) lag keine Antwort vor.

### ***Qualitative Analyse zur Studienteilnahme der Therapeut\*innen***

Mit Stand Dezember 2021 lagen Erfahrungsberichte von 15 Therapeut\*innen vor, deren Inhalte sich über rund 30 Seiten erstreckte. Im Anschluss werden die Ergebnisse der bislang vorhandenen Erfahrungsberichte hinsichtlich ihrer eigenen Studienteilnahme berichtet.

**Gründe für eine Studienteilnahme der Therapeut\*innen.** Der in den Erfahrungsberichten der Therapeut\*innen am häufigste angesprochene Studienteilnahmegrund war der Wunsch einen Beitrag zur Qualitätssicherung bzw. Wirksamkeitsuntersuchung der Existenzanalyse zu leisten, um die berufspolitische Stellung der eigenen Therapierichtung zu festigen.

Als weiterer Grund wurde die Möglichkeit der beruflichen Weiterentwicklung in diagnostischen und wissenschaftlichen Prozessen angeführt. Zudem wurden von einigen Therapeut\*innen die zusätzlichen Informationsquellen durch die eingesetzten Verfahren und die Rückmeldungen positiv hervorgehoben.

Die wenigen Angaben zu den möglichen Gründen von Kolleg\*innen nicht an der Studie teilzunehmen betrafen die Sorge nicht rechtzeitig die Studienanforderungen (ausreichende Anzahl an teilnehmenden Klient\*innen) zu erfüllen, um eine Abschlussarbeit verfassen zu können. Dadurch wurde eine Verzögerung des fachspezifischen Abschlusses befürchtet. Außerdem wurde die Befürchtung vor einer möglichen negativen Beeinflussung der therapeutischen Beziehung bzw. des Prozesses durch die Studienanforderungen geäußert.

**Studienintegration in den therapeutischen Alltag.** Nahezu alle teilgenommenen Therapeut\*innen merkten in den bisherigen Erfahrungsberichten an, dass sich die Studienteilnahme gut in den therapeutischen Arbeitsalltag integrieren ließ. Schwierigkeiten traten gehäuft zu Beginn der Teilnahme auf, weswegen besonders in dieser Phase Studieninformationen zu Ablauf und Durchführung, Einschulungen und regelmäßigem Austausch im Zuge von Vernetzungstreffen als sehr hilfreich erlebt wurden.

**Schwierigkeiten.** Bei der Studienanwerbung von Klient\*innen war es für einige Therapeut\*innen herausfordernd den geeigneten Zeitpunkt dafür zu finden. Andere berichteten hingegen, dass es für sie kein Problem dargestellt habe, dieses Prozedere im Erstgespräch bzw. in den ersten Stunden unterzubringen. In der Phase zu Beginn einer Therapie, in der noch viel allgemeine Aufklärung zu Setting, Ablauf, Arbeitsweise, etc. erfolgte, konnte auch über eine mögliche Studienteilnahme gesprochen werden. Ein deutlicher Konsens in den bisherigen Erfahrungsberichten der Therapeut\*innen lag darin, dass die Betonung der ausnahmslosen Freiwilligkeit einer Teilnahme und die damit verbundene Möglichkeit eines Abbruches zu jeder Zeit und ohne jedwede Konsequenzen die meisten Bedenken der Klient\*innen ausgeräumt hätte.

Häufig wurde auch von (anfänglichen) Schwierigkeiten bei der Durchführung der IDCL berichtet. Die Dauer wäre zu Beginn schwer einzuschätzen und ein zeitlicher Mehraufwand wäre dadurch spürbar gewesen. Manche Therapeut\*inne befürchteten das Diagnoseverfahren falsch anzuwenden. Nach einer Einarbeitungsphase würden sich die Schwierigkeiten verkleinern und die IDCL konnte als ergänzende Information zum Befinden der Klient\*innen genutzt werden. Die Anwendung der Checklisten hätte außerdem dazu angestoßen, einzelne Symptome achtsamer wahrzunehmen und das Wissen über Standarddiagnosen nach ICD-10 zu festigen, was wiederum den diagnostischen Blick bei einigen geschärft hätte.

Vereinzelt Befürchtungen, dass die therapeutische Beziehung durch eine Studienteilnahme negativ beeinflusst werden könnte, wurden durch die mehrheitlich geäußerten Erfahrungen nicht bestätigt. Zwar wurde von einzelnen Personen ein Spannungsverhältnis beschrieben, das zwischen dem phänomenologischen Zugang der Existenzanalyse und der nomothetischen Herangehensweise von quantitativen Verfahren bestünde. Für die meisten Therapeut\*innen war es jedoch möglich, die quantitative Wissensinhalte mit einem existenzanalytischen Verständnis zu vereinbaren und beide als gegenseitige Ergänzung zu betrachten.

**Umgang mit Rückmeldungen.** Aus den Erfahrungsberichten der Therapeut\*innen lassen sich teils sehr unterschiedliche Umgangsformen mit der Möglichkeit des Erhalts von Rückmeldungen ableiten. Während manche nahezu bei allen Klient\*innen eine Rückmeldung wünschten, gab es andere, die kein Interesse an einer Rückmeldung der Fragebogenergebnisse hatten. Die Mehrheit nutzte die Rückmeldung bei einzelnen Klient\*innen mit diagnostischen Unklarheiten und schwierigeren Verläufen. Ähnlich wie bei der IDCL wäre auch durch die Rückmeldungen ein achtsamerer Umgang mit einzelnen Symptomen angestoßen worden. Vereinzelt wurde auch von Irritationen berichtet, wenn der therapeutische Eindruck nicht mit den Fragebogenergebnissen übereinstimmte. Durch Gespräche mit dem Forschungsteam oder

in der Supervision konnte der Umgang mit Ergebnissen aus Selbsteinschätzungsfragebögen gut besprochen werden.

**Studienorganisation und Veränderungsvorschläge.** Die Kommunikation mit dem Forschungsteam wurde in den bisherigen Erfahrungsberichten mehrheitlich als niederschwellig und unkompliziert beschrieben. Die Therapeut\*innen schätzten die Möglichkeit studienrelevante Fragen oder Unklarheiten per E-Mail oder Telefon rasch und unkompliziert klären zu können. In direktem Austausch mit dem Forschungsteam stehen zu können, wurde als wertschätzend und motivierend erlebt. Die regelmäßigen Vernetzungstreffen, bei denen sich die Therapeut\*innen und das Forschungsteam in der Gruppe austauschen konnten, wurde als hilfreich rückgemeldet.

Veränderungsvorschläge der Psychotherapeut\*innen betrafen Wünsche nach detaillierteren Informationen zu Beginn der Studienteilnahme bezüglich der IDCL und dem Umgang mit Therapieunterbrechungen. Auch das Verhältnis von Aufwand und Nutzen wurde erwähnt. Nach Ende der Abschlussarbeit sei der Nutzen einer Studienteilnahme, abgesehen von den Rückmeldungen, gering. Kürzere Fragebögen für die Klient\*innen, sowie eine Digitalisierung der Studiendurchführung wurden ebenso als Verbesserung vorgeschlagen.

### ***Teilnahmebereitschaft und Studienakzeptanz der Klient\*innen***

**Rücklauf der Klient\*innen.** Um die Rücklaufquote der Klient\*innen zu ermitteln, wurden die Therapeut\*innen gebeten in einer Liste der Strukturierungshilfe die Zu- und Absagen zu vermerken. Es konnten Strukturierungshilfen von bislang 22 Therapeut\*innen gesammelt und ausgewertet werden. Aus den Strukturierungshilfen lagen Informationen über die Zu- und Absage bzw. Voraussetzungserfüllung von 291 Personen vor. Die Voraussetzungen nicht erfüllt hatten 27 Klient\*innen (9 %). Dafür wurden von den Therapeut\*innen vereinzelt Gründe wie Minderjährigkeit, Intelligenzminderung und der (psychische) Zustand der Klient\*innen im Anmerkungsfeld angegeben. Von den 264 angefragten Klient\*innen, die die Voraussetzungen erfüllten, hatten sich 182 Personen bereit erklärt bei der Studie mitzumachen. Daraus ergab sich eine vorläufige Rücklaufquote von rund 69 % aller geeigneter Personen (vgl. Tabelle 3).

Die Geschlechterverteilung lag in der Gruppe der Studienteilnehmer\*innen bei 71 % weiblichen und 29 % männlichen Personen bzw. bei 69 % weiblichen und 31 % männlichen Personen in der Gruppe der Nicht-Teilnehmer\*innen. Die Gruppe der Personen, die die Voraussetzungen nicht erfüllt hatten, bestand aus 56 % Frauen und 44 % Männern. Weitere Geschlechterkategorien wurden nicht erfragt. Das Durchschnittsalter bei den

Teilnehmer\*innen lag bei  $M = 36.5$  Jahren ( $SD = 17.2$ ) und bei den Nicht-Teilnehmer\*innen bei  $M = 46.2$  Jahren ( $SD = 13.6$ ) (vgl. Tabelle 3).

**Tabelle 3**

*Rücklauf der Klient\*innen*

Rücklauf-Status	<i>n</i>	% <sup>1</sup>	% <sup>2</sup>
aufgezeichnete Personen gesamt	291	-	100
Voraussetzungen nicht erfüllt	27	-	9
angefragte geeignete Personen gesamt	264	100	91
teilnahmebereite Personen	182	69	63
nicht teilnahmebereite Personen	82	31	28

*Anmerkung.* *n* = absolute Häufigkeit der jeweiligen Gruppen; %<sup>1</sup> = prozentueller Anteil an allen geeigneten Personen; %<sup>2</sup> = prozentueller Anteil an allen aufgezeichneten Personen.

Es wurde ein t-Test bei unabhängigen Stichproben durchgeführt, um zu überprüfen, ob sich die beiden Gruppen der Teilnehmer\*innen bzw. Nicht-Teilnehmer\*innen hinsichtlich ihres Alters signifikant voneinander unterscheiden. Die Variable Alter war nach dem Kolmogorov-Smirnov-Test in beiden Gruppen nicht normalverteilt ( $p < .05$ ). Aufgrund der ausreichend großen Stichprobengröße wird die Analyse dennoch fortgesetzt. Für beide Gruppen lag nach dem Levene-Test ( $p = .32$ ) Varianzhomogenität vor.

Es gab einen statistisch signifikanten Unterschied mit mittlerer Effektstärke zwischen den Personen, die sich zu einer Teilnahme bereit erklärt hatten und jenen, die eine Studienteilnahme ablehnten hinsichtlich des Durchschnittsalters,  $t(225) = -4,513$ ,  $p < .001$ ,  $d = 0,67$ , 95 %-CI [0.373; 0.967]. Das Durchschnittsalter war in der Gruppe der Studienteilnehmer \*innen um zehn Jahre niedriger (95 %-CI [5,506; 14,042]).

Ein  $\chi^2$ -Test wurde für die Variablen Geschlecht und der Teilnahmebereitschaft durchgeführt. Es gab keinen statistisch signifikanten Zusammenhang zwischen dem Geschlecht und der Teilnahmebereitschaft,  $\chi^2(1) = .103$ ,  $p = .749$ . Alle erwarteten Zelhäufigkeiten waren größer als 5.

### ***Qualitative Analyse zur Teilnahmebereitschaft der Klient\*innen***

Im Anschluss werden die Ergebnisse der bisherigen Erfahrungsberichte der Therapeut\*innen hinsichtlich der Gründe für eine Studienablehnung von Klient\*innen beschrieben.

Die am meisten genannten, wahrgenommenen Prädiktoren für eine Studienteilnahme bzw. keine Studienteilnahme waren der (psychische) Zustand und das Alter der Klient\*innen. Der symptomatische Zustand der Klient\*innen hätte, nach Meinung einiger Therapeut\*innen, einen Einfluss auf die Entscheidung, bei der Studie mitzumachen, gehabt. Insbesondere Personen mit Symptomaten einer affektiven Störung hätten eine Studienteilnahme

abgelehnt. Ausgeprägte Symptome, wie mangelnde Konzentration oder erschwertes Auffassungsvermögen, hätten dazu geführt, dass eine Studienteilnahme als zusätzliche Belastung erlebt wurde. Auch Angstsymptome bzw. paranoide Tendenzen wurden als hinderlich für eine Teilnahme beschrieben. Bedeutend wichtiger als das Störungsbild an sich, wäre jedoch die Intensität der Symptome gewesen. Je stärker sich der Leidensdruck dargestellt habe, desto eher wurde eine Studienteilnahme als zusätzliche Belastung erlebt. Einen weiteren beschriebenen Einfluss auf eine mögliche Studienteilnahme der Klient\*innen hätte das Alter dargestellt. Ältere Personen hätten nach Angaben der Therapeut\*innen eine Studienteilnahme eher abgelehnt, während jüngere sich schneller zu einer Teilnahme bereit erklärt hätten. Auch Klient\*innen, die eine andere Erstsprache als Deutsch hatten, wären einer Teilnahme eher abgeneigt gewesen. Vereinzelt Therapeut\*innen berichteten von der Wahrnehmung, dass zum Zeitpunkt der Studienanwerbung den Klient\*innen noch nicht immer klar gewesen wäre, ob sie überhaupt eine Therapie (bei ihnen) beginnen wollen würden.

#### ***Vorläufige Abbruchsanalyse und bisheriger Fragebogenrücklauf***

Aufgrund des laufenden Charakters der Studie und einer fehlenden, systematischen Erhebung von Abbruchsfällen besteht die Abbruchsanalyse vorwiegend aus den Informationen der Therapeut\*innenfragebögen jener Personen, die trotz Therapieabbruch bereit waren, einen Fragebogen zu Therapieende auszufüllen ( $n = 19$ ).

Zudem machten wenige Therapeut\*innen vereinzelt Anmerkungen zu Abbrüchen in den Strukturierungsbögen. Dadurch konnten vorläufig 37 Fälle einem gesicherten Abbruchsstatus zugeordnet werden (vgl. Tabelle 4). Nach *Therapiebeginn* brachen bisweilen 17 Personen ab. Dabei handelte es sich in elf Fällen um *Therapieabbrüche*, die in den ersten zehn Einheiten stattfanden. Drei Fälle konnten davon als *nicht-qualitätsrelevante Therapieabbrüche*, wie fehlende finanzielle Mittel und Wohnortwechsel, charakterisiert werden. Des Weiteren konnten sechs *Studienabbrüche* ausgemacht werden. In fünf Fällen handelte es sich um einen Studienabbruch von Therapeut\*innen und in einem Fall um einen Klient\*innen-Studienabbruch. Es konnten bislang drei Abbrüche nach der Verlaufsmessung identifiziert werden (vgl. Tabelle 4). Dabei handelte es sich in einem Fall um einen Studienabbruch durch Therapeut\*innen und in zwei Fällen um einen Studienabbruch durch Klient\*innen.

**Tabelle 4**

Vorhandene Informationen zu Abbruchsfällen

Zeitpunkt	Anzahl	Gründe ( <i>n</i> )
<b>Therapieabbruch (11)</b>		
Nach Therapiebeginn (t1)	<i>n</i> = 17	<b>qualitätsrelevant</b> (8) in der Anfangsphase <b>nicht-qualitätsrelevant</b> (3) finanzielle Probleme & Umzug
<b>Studienabbruch (6)</b>		
<b>Klient*innen</b> (1) Fragebogen wird nur halb ausgefüllt		
<b>Therapeut*innen-Abbruch</b> (5)		
<b>Studienabbruch (3)</b>		
Nach Verlaufsmessung (t2)	<i>n</i> = 3	<b>Klient*innen</b> (2) wollen keine Testung zu Therapieende <b>Therapeut*innen-Abbruch</b> (1)
<b>Therapieabbruch (17)</b>		
Vorzeitiges Therapieende (t3)	<i>n</i> = 17	<b>qualitätsrelevant</b> (2) schwere Arbeitsbeziehung (1) keine weiteren Informationen (1) <b>nicht-qualitätsrelevant</b> (15) Mutterschutz der Therapeut*in (4) Bewilligung Reha/Kur/Kassenplatz (4) Umzug (3) Mangel an zeitlichen und finanziellen Ressourcen (2) Praktikumsende der Therapeut*innen (2)

*Anmerkung.* Informationen aus den Strukturierungshilfen und den Therapeut\*innenfragebögen zu Therapieende.

Zu Therapieende lagen Informationen von 17 vorzeitigen Therapieabschlüssen vor. Zwei Fälle konnten einem qualitätsrelevanten Abbruchgrund zugeordnet werden. In 15 Fällen handelte es sich um nicht-qualitätsrelevante Therapieabbrüche (vgl. Tabelle 4).

Bis März 2022 lagen zur Verlaufsmessung Daten von 40 % (*n* = 105) der Fälle zu Therapiebeginn (*n* = 263) vor. Zu Therapieende lagen Daten von 30 % (*n* = 79) und zum Katamnesezeitpunkt lagen Daten von 12 % (*n* = 31) aller Fälle zu Therapiebeginn vor.

#### ***Klient\*innenbewertungen zum Erleben der Fragebögen***

Die Teilnehmer\*innen benötigten für das Ausfüllen des Fragebogens zu Therapiebeginn durchschnittlich 58 Minuten (*SD* = 31.4). Rund 5 % (*n* = 13) der Gesamtstichprobe fand den Fragebogen zu lang und 28 % (*n* = 74) etwas zu lang. Die Mehrheit, nämlich 65 % (*n* = 172), empfand die Länge des Fragebogens als angemessen. Für

eine Person (0,4 %) war der Fragebogen zu kurz und drei Personen enthielten sich bei dieser Frage.

Rund 72 % ( $n = 189$ ) der Gesamtstichprobe gab an, den Fragebogen komplett verstanden zu haben, weitere 26 % ( $n = 69$ ) haben den Fragebogen eher verstanden und drei Personen (1,1 %) haben den Fragebogen kaum verstanden. Niemand gab an, den Fragebogen gar nicht verstanden zu haben und zwei Personen (0,8 %) haben keine Antwort angegeben.

Während des Ausfüllens konnten sich rund 24 % ( $n = 62$ ) ausgezeichnet und 60 % ( $n = 158$ ) gut konzentrieren. 36 Personen (13,7 %) konnten sich weniger gut und 6 Personen (2,3 %) schlecht konzentrieren.

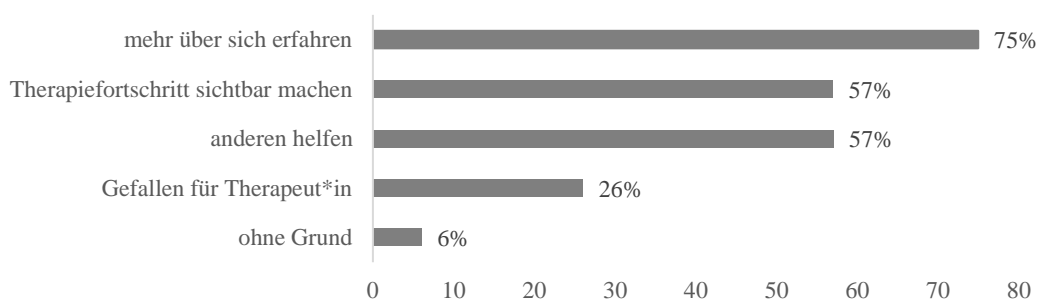
Nach dem Ausfüllen des Fragebogens fühlte sich die Mehrheit, nämlich 143 Personen (54,4 %), weder müde noch aktiviert. 22 % der Gesamtstichprobe ( $n = 59$ ) fühlten sich nach dem Ausfüllen eher müde und rund 7 % ( $n = 17$ ) gaben an sich danach müde gefühlt zu haben.

Mit rund 80 % ( $n = 209$ ) hat die Mehrheit der Gesamtstichprobe angegeben, den Fragebogen korrekt beantwortet zu haben. Weitere 19 % ( $n = 51$ ) haben den Fragebogen eher korrekt beantwortet und eine Person (0,4 %) gab an, den Fragebogen eher nicht korrekt beantwortet zu haben. Niemand gab an, den Fragebogen nicht korrekt beantwortet zu haben.

Die Frage nach dem Teilnahmegrund konnte mehrfach beantwortet werden. Am häufigsten, rund 75 % ( $n = 197$ ) der Gesamtstichprobe, wurde der Grund *mehr über sich erfahren* angekreuzt.

### Abbildung 3

Studienteilnahmegrund der Klient\*innen zu Therapiebeginn ( $n = 263$ )



Anmerkung. Bei der Frage zum Teilnahmegrund der war eine Mehrfachnennung möglich.

Von jeweils 57 % ( $n = 150$ ) wurden die Gründe *den Therapiefortschritt abzubilden* und *anderen damit zu helfen* angegeben. Zusätzlich wurde von 26 % ( $n = 69$ ) angeführt, bei der Studie mitzumachen, um *den Therapeut\*innen einen Gefallen zu tun* und weitere 6 % ( $n = 16$ ) gaben an, *ohne Grund* bei der Studie mitzumachen (vgl. Abbildung 3).

Es wurde ein *Wilcoxon-Test* durchgeführt, um eine mögliche Stimmungsveränderung durch das Ausfüllen des Fragebogens zu untersuchen. Dafür wurde die ordinale Stimmungsskala zu Beginn des Fragebogens mit derselben Stimmungsskala am Ende des Fragebogens auf Unterschiede überprüft. Es zeigte sich ein statistisch signifikanter Unterschied zwischen den beiden Messwerten zu Therapiebeginn ( $Mdn = 4$ ;  $M = 3.84$ ;  $SD = 1.32$ ) und Therapieende ( $Mdn = 4$ ;  $M = 4.22$ ;  $SD = 1.40$ ) mit kleinem Effekt,  $z = -5,02$ ;  $p < .001$ ;  $r = .32$ ,  $d = 0.30$ ,  $n = 241$ . Das Ausfüllen des Fragebogens hob die Stimmung der Klient\*innen mit kleinem Effekt.

**Qualitative Analyse der Erfahrungsberichte zum Erleben der Fragebögen.** Der Umfang des Fragebogens, insbesondere zu Therapiebeginn, wurde in manchen mündlichen Rückmeldungen der Klient\*innen an ihre Therapeut\*innen als zu lang und zeitaufwändig beschrieben. Die Therapeut\*innen berichteten, dass tendenziell eher ältere Personen den Fragebogen als zu lang und in der Formulierung zu kompliziert gefunden hätten. Verständnisschwierigkeiten bezogen sich einerseits auf uneindeutige Formulierungen und andererseits auf das sprachliche Niveau, das teilweise als hoch erlebt worden wäre. Insbesondere Personen mit nicht-deutscher Erstsprache meldeten Schwierigkeiten zurück.

Neben den Berichten von Klient\*innen, die Schwierigkeiten beim Ausfüllen der Fragebögen hatten, gab es auch Berichte von Personen, welche die Fragen als bereichernd und zur Selbstreflexion anregend erlebt hätten. Die Inhalte der Fragen wären gut in den Therapieprozess integrierbar gewesen und die zusätzlichen Informationen durch die Rückmeldungen wären von einigen Klient\*innen geschätzt worden.

Manche Klient\*innen hätten den Wunsch geäußert, durch eine Studienteilnahme etwas zur gesellschaftspolitischen Stellung von psychischen Erkrankungen und psychotherapeutischen Behandlungen beitragen zu können.

### ***Beschreibung der Klient\*innenstichproben***

Im Zeitraum von April 2018 bis März 2022 (Laufzeit knapp vier Jahre) konnten Datensätze von 266 Personen gesammelt werden. Eine Person war mit 17 Jahren noch minderjährig und wurde somit von jeglicher Analyse ausgeschlossen. Von 2 Klient\*innen wurden Testbögen abgegeben, jedoch ohne ausgefüllte Fragebögen. Die Gesamtstichprobe beinhaltet somit Daten von 263 Personen.

### ***Gesamtstichprobe der Klient\*innen***

**Skalenkennwerte und interne Konsistenzen.** Die internen Konsistenzen der verwendeten Skalen lagen alle bis auf die Subskalen *narzisstische* und *antisoziale* des ADP-IV Cluster B im guten bis exzellenten Bereich (vgl. Tabelle 5) (Cohen, 1988).

**Tabelle 5***Kennwerte und interne Konsistenzen der Skalen*

Skala	Items	$\alpha$ lt. Literatur	Therapiebeginn	
			$n$	Erreichtes $\alpha$
RYFF	18	.33-.56 (Subskalen)	252	.81
ESK	46	.93	231	.94
GSI	90	.98	226	.97
DEPR	13	.89	257	.91
ANGS	10	.88	258	.88
IIP	64	.71-.82 (Subskalen)	228	.94
ADP-IV (Cluster B)	35	.91	215	.92
antisozial	8	.85	248	.67
Borderline	10	.80	248	.84
histrionisch	8	.72	246	.80
narzisstisch	9	.68	242	.77

*Anmerkung.*  $\alpha$  = Cronbach's Alpha;  $n$  = Anzahl der Fälle mit vollständigen Skalenitems; RYFF = Gesamtskala der Ryff-Skala; ESK = Gesamtskala der Existenz-Skala; GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; ADP-IV = Assessment of DSM-IV Personality Disorders.

**Soziodemographische Daten.** In den folgenden Beschreibungen werden die Daten der Gesamtstichprobe sowie der Prä-Post-Stichprobe angeführt, da die Prä-Post-Stichprobe in den statistischen Analysen eine bedeutsame Rolle einnimmt.

Das Durchschnittsalter zu Therapiebeginn betrug in der Gesamtstichprobe rund 36 Jahre und in der Prä-Post-Stichprobe 37 Jahre (vgl. Tabelle 6).

Während der Anteil weiblicher Personen in der Gesamtstichprobe bei 73 % lag, war dieser Wert in der Prä-Post-Stichprobe mit 75 % ein wenig darüber. Die Gruppe der ledigen Personen (36 %,  $n = 95$ ) stellte in der Gesamtstichprobe den häufigsten Familienstand dar, gefolgt von verheirateten (30 %,  $n = 79$ ) und in Partnerschaft lebenden Personen mit rund 26 % ( $n = 68$ ). Dagegen bildete die größte Gruppe in der Prä-Post-Stichprobe verheiratete Personen mit 42 % ( $n = 33$ ), gefolgt von der Gruppe lediger Personen mit rund 35 % ( $n = 28$ ) (vgl. Tabelle 6). Mit rund 40 % stellte in beiden Stichproben Personen mit einem Lehr- bzw. Fachschulabschluss ohne Matura die größte Gruppe bei den formalen Bildungsabschlüssen dar. In der Prä-Post-Stichprobe war die nächsthäufige angegebene Bildungskategorie ein Studienabschluss mit 24 % ( $n = 19$ ), gefolgt von Maturaabschluss mit 20 % ( $n = 16$ ). In der Gesamtstichprobe lag der Maturaabschluss (26 %,  $n = 67$ ) an zweiter Stelle gefolgt von einem Studienabschluss mit rund 21 % ( $n = 55$ ). Die Daten sind in Tabelle 6 aufgeführt.

**Tabelle 6***Soziodemographische Daten der Klient\*innen zu Therapiebeginn*

		Prä-Post-Stichprobe (n = 79)		Gesamtstichprobe (n = 263)	
Alter	Mittelwert	37.39		36.28	
	Median	33.00		32.00	
	Standardabweichung	13.18		13.24	
	Minimum	18		18	
	Maximum	68		75	
Geschlecht		<i>n</i>	% <sup>a</sup>	<i>n</i>	% <sup>a</sup>
	weiblich	59	74.7	190	72.8
	männlich	20	25.3	73	27.2
Familienstand	ledig	28	35.4	95	36.1
	verheiratet	33	41.8	79	30.0
	in Partnerschaft	13	16.5	68	25.9
	geschieden	4	5.1	16	6.1
	verwitwet	1	1.3	2	.8
Bildungsabschluss	Pflichtschule	11	13.9	27	10.3
	Lehre / Fachschule ohne Matura	31	39.2	104	39.5
	höhere Schule mit Matura	16	20.3	67	25.5
	Studium / Universität / Fachhochschule	19	24.1	55	20.9
	sonstiges	2	2.5	9	3.4

Anmerkung. a. prozentueller Anteil an der jeweiligen Stichprobe; *n* = absolute Häufigkeit.

**Skalenmittelwerte.** Die Skalenmittelwerte bzw. Standardabweichungen des GSI betragen zu Therapiebeginn  $M = 1.10$  ( $SD = 0.66$ ) in der Gruppe der Prä-Post-Stichprobe und  $M = 1.11$  ( $SD = 0.67$ ) in der Gesamtstichprobe. Der Mittelwert der Depressionsskala der SCL betrug in der Prä-Post-Stichprobe  $M = 1.57$  ( $SD = 0.94$ ) und in der Gesamtstichprobe  $M = 1.55$  ( $SD = 0.90$ ). Auf der Angstskala der SCL konnte in der Prä-Post-Stichprobe ein Gruppenmittelwert von  $M = 1.09$  ( $SD = 0.75$ ) und in der Gesamtstichprobe ein Gruppenmittelwert von  $M = 1.14$  ( $SD = 0.83$ ) ermittelt werden. Die Gruppenmittelwerte der Gesamtskala des IIP waren mit einem Wert von  $M = 1.48$  ( $SD = 0.54$ ) in beiden Stichproben ident (vgl. Tabelle 7).

In den ressourcenbezogenen Skalen, bei denen höhere Werte ein besseres Befinden beschreiben, betragen die Gruppenmittelwerte der ESK-Gesamtskala  $M = 189.11$  ( $SD = 35.66$ ) in der Prä-Post-Stichprobe und  $M = 186.94$  ( $SD = 36.54$ ) in der Gesamtstichprobe. Die Ryff-Skala wies in der Prä-Post-Stichprobe einen Mittelwert von  $M = 66.12$  ( $SD = 12.65$ ) und  $M = 68.67$  ( $SD = 11.46$ ) in der Gesamtstichprobe auf. Auf der Fremdeinschätzungsskala, GAF, die von den Psychotherapeut\*innen ausgefüllt wurde,

ergaben sich Mittelwerte von  $M = 61.67$  ( $SD = 15.45$ ) für die Prä-Post-Stichprobe und  $M = 63.66$  ( $SD = 15.44$ ) in der Gesamtstichprobe. Auch hier werden höhere Werte mit besserem Befinden assoziiert (vgl. Tabelle 7).

**Tabelle 7**

*Gruppenmittelwerte auf den Skalen zu Therapiebeginn*

Skalen	<i>n</i>	Prä-Post-Stichprobe		Gesamtstichprobe	
		<i>(n = 79)</i>		<i>(n = 263)</i>	
		<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>	
GSI	79	1.10 (0.66)	263	1.11 (0.67)	
DEPR	79	1.57 (0.94)	263	1.55 (0.90)	
ANGS	79	1.09 (0.75)	263	1.14 (0.83)	
IIP	79	1.48 (0.54)	263	1.48 (0.54)	
ESK	77	189.11 (35.66)	258	186.94 (36.54)	
RYFF	72	66.19 (12.57)	252	68.73 (11.41)	
GAF	76	61.67 (15.45)	252	63.66 (15.44)	

*Anmerkung.* *M* = Mittelwert; *SD* = Standardabweichung (*standard deviation*); *n* = Anzahl verrechenbarer Skalen bzw. Anzahl an Personen im Datensatz; GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskaala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; ESK = Gesamtskala der Existenz-Skala (höhere Werte gehen mit besserem Befinden einher); RYFF = Gesamtskala der Ryff-Skala Skala (höhere Werte gehen mit besserem Befinden einher); GAF = Global Assessment of Functioning Skala (höhere Werte gehen mit besserem Befinden einher).

**Anteil an Personen mit klinisch (un-)auffälliger Symptomatik.** Auf der GSI-Gesamtskala der SCL ergaben sich zu Therapiebeginn klinisch auffällige Werte bei rund 66 % der Gesamtstichprobe und für 67 % der Personen in der Prä-Post-Stichprobe. Klinisch auffällige Testwerte auf der Depressionsskala des SCL wiesen rund 70 % der Gesamtstichprobe und 68 % der Prä-Post-Stichprobe auf. 65 % bzw. 66 % der Teilnehmer\*innen wurden in den beiden Stichproben auf der Angstskaala der SCL als klinisch auffällig eingestuft. Für die IIP-Gesamtskala ergaben sich klinisch auffällige Werte für 67 % der Personen in der Prä-Post-Stichprobe und für 68 % der Gesamtstichprobe (vgl. Tabelle 8).

Insgesamt lag zu Therapiebeginn bei 84 % der Klient\*innen in mindestens einem der symptombezogenen Instrumente ein klinisch auffälliges Ergebnis vor.

**Tabelle 8***Personen mit klinisch (un-)auffälliger Symptomatik zu Therapiebeginn*

Skala	Psychisches Befinden	Prä-Post-Stichprobe (n = 79)		Gesamtstichprobe (n = 263)	
		n	%	n	%
GSI	klinisch auffällig	53	67.1	173	65.8
	klinisch nicht auffällig	26	32.9	90	34.2
DEPR	klinisch auffällig	54	68.4	185	70.3
	klinisch nicht auffällig	25	31.6	78	29.7
ANGST	klinisch auffällig	51	64.6	173	65.8
	klinisch nicht auffällig	28	35.4	90	34.2
IIP	klinisch auffällig	53	67.1	178	67.7
	klinisch nicht auffällig	26	32.9	85	32.3
GAF	klinisch auffällig	48	60.8	142	54.0
	klinisch nicht auffällig	28	35.4	110	41.8

*Anmerkung.* Klinische Schwellenwerte (*c*) nach Jacobson: *c* = 0.75 für den GSI; *c* = 0.96 für die Depressionsskala; *c* = 0.67 für die Angstskala; *c* = 1.21 für die IIP-Gesamtskala (Franke, 2014; Hansen & Lambert, 1996); Schwellenwert für die GAF-Skala: *c* = 70 (Cramer et al., 2014; von Wyl et al., 2013); GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; GAF = Global Assessment of Functioning.

**Diagnosen.** Die IDCL-Checklisten ermöglichte es den Therapeut\*innen anhand standardisierter Symptomlisten vorhandene Störungsbilder der Klient\*innen valide und reliabel zu erheben. Die zwei größten Diagnosegruppen bei den Hauptdiagnosen (erste Diagnose) in beiden Stichproben waren jene der Affektiven Störungen (F3-Diagnosen) mit 39 % in der Prä-Post-Stichprobe und rund 42 % in der Gesamtstichprobe. Auf die Diagnose der neurotischen, belastungsreaktiven und somatoformen Störungen (F4-Diagnosen) entfielen 33 % in der Prä-Post-Stichprobe und 31 % in der Gesamtstichprobe.

Die restlichen Diagnosegruppen, Psychische und Verhaltensstörungen durch psychotrope Substanzen (F1) und Essstörungen (F5), nahmen einen Anteil von 3 % in der Prä-Post-Stichprobe und 4 % in der Gesamtstichprobe ein. Keine Diagnosen anhand der IDCL wiesen 24 % der Personen in der Prä-Post-Stichprobe und 23 % der Gesamtstichprobe auf (vgl. Tabelle 9).

**Tabelle 9***Hauptdiagnosen zu Therapiebeginn*

Diagnosegruppe	Prä-Post-Stichprobe (n = 79)		Gesamtstichprobe (n = 263)	
	n	%	n	%
F3	31	39.2	110	41.8
F4	26	32.9	81	30.8
Rest	3	3.8	11	4.2
Gesamt	60	75.9	202	76.8
Keine Diagnose	19	24.1	61	23.2

*Anmerkung.* Diagnosen ermittelt durch die IDCL nach ICD-10. F3 = Affektive Störungen; F4 = Neurotische, Belastungs- und somatoforme Störungen; Rest = F1 (Psychische und Verhaltensstörungen durch psychotrope Substanzen) und F5 (Essstörungen).

**Komorbidität.** Bei rund 27 % der Klient\*innen der Gesamtstichprobe wurden mindestens eine komorbide, psychische Störung festgestellt (vgl. Tabelle 10).

**Tabelle 10***Anzahl an Diagnosen pro Person zu Therapiebeginn*

Anzahl an Diagnosen	Prä-Post-Stichprobe (n = 79)		Gesamtstichprobe (n = 263)	
	n	%	n	%
mind. 1 Diagnose	60	75.9	202	76.8
1 Diagnose	38	48.1	129	49.1
2 Diagnosen	22	27.9	71	27.0
3 Diagnosen	7	8.9	17	6.5
4 Diagnosen	2	2.5	7	2.7
5 Diagnosen	-	-	1	0.4

*Anmerkung.* Diagnosen ermittelt durch die IDCL nach ICD-10.

Von den 110 Personen mit einer affektiven Störung als Hauptdiagnose wiesen 31 Klient\*innen (28 %) mindestens eine weitere Diagnose auf. Die häufigste komorbide Störung war dabei jene der neurotischen, belastungsreaktiven und somatoformen Störungen (F4, n = 25).

Bei den Personen mit neurotischen, belastungsreaktiven und somatoformen Störungen (n = 81) als Hauptdiagnose gab es 32 Personen (41 %) mit mindesten einer weiteren Störung. Davon fielen rund 50 % auf affektive Störungen als Zweitdiagnose und die andere Hälfte auf eine weitere Form der F4-Diagnosen.

**Verdacht auf Persönlichkeitsstörung.** In der Gesamtstichprobe ergab sich bei rund 25 % ( $n = 66$ ) der Klient\*innen ein Verdacht auf das Vorliegen einer oder mehrerer Persönlichkeitsstörung(en). Die Ergebnisse in der Prä-Post-Stichprobe legten bei 19 % ( $n = 15$ ) den Verdacht auf das Vorhandensein mindestens einer Persönlichkeitsstörung nahe (vgl. Tabelle 11). Die häufigste Verdachtsdiagnose bei den Persönlichkeitsstörungen im *Cluster-B* war die Borderline-Persönlichkeitsstörung mit 57 Verdachtsdiagnosen (22 %) in der Gesamtstichprobe.

**Tabelle 11**

*Verdacht auf Persönlichkeitsstörungen zu Therapiebeginn*

Anzahl Verdachtsdiagnosen	Prä-Post-Stichprobe ( $n = 79$ )		Gesamtstichprobe ( $n = 263$ )	
	$n$	%	$n$	%
Mind. 1 Verdacht	15	19.0	66	25.1
1 Verdacht	9	11.4	43	16.3
2 Verdachtsdiagnosen	6	7.6	17	6.5
3 Verdachtsdiagnosen	-	-	6	2.3
kein Verdacht	64	81.0	197	74.9

*Anmerkung.* Cluster-B Diagnosen ermittelt durch das Verfahren *Assessment of DSM-IV Personality Disorders (ADP-IV)*.

**Ergänzende Informationen zur Gesamtstichprobe.** In der Gesamtstichprobe gaben rund 47 % der Klient\*innen ( $n = 123$ ) an, dass sie durch Eigeninitiative zu einer psychotherapeutischen Behandlung gefunden hätten. 21 % ( $n = 56$ ) hätten von Allgemeinmediziner\*innen und 18 % ( $n = 46$ ) von Fachärzt\*innen für Psychiatrie und/oder Neurologie eine Psychotherapie empfohlen bekommen. Weitere 13 % ( $n = 33$ ) gaben klinische Psycholog\*innen bzw. Beratungsstellen und die Kategorie *andere* auf die Frage an, wer Ihnen eine Psychotherapie empfohlen habe. Rund 54 % ( $n = 141$ ) der Gesamtstichprobe habe schon zuvor psychosoziale Hilfe in Anspruch genommen. Weitere 15 % ( $n = 40$ ) aller Studienteilnehmer\*innen gaben frühere stationäre Aufenthalte an. Nach Informationen von 198 Personen wurde der Fragebogen zu Therapiebeginn sowohl im Durchschnitt als auch im Median vor der 3.Einheit ausgefüllt.

**Anteil an Therapeut\*innen im Gesamtdatensatz.** Insgesamt verteilten sich die 263 Klient\*innen auf 31 Therapeut\*innen. Dabei zeigen sich sehr unterschiedliche Häufigkeiten. 44 % der Therapeut\*innen ( $n = 14$ ) haben bis zu fünf Klient\*innen zu einer Studienteilnahme bewegen können, 28 % der Therapeut\*innen ( $n = 9$ ) haben zwischen sechs und zehn ihrer Klient\*innen im Zuge der Studie behandelt. Vier Therapeut\*innen (13 %) haben zwischen 11

und 25 ihrer Klient\*innen zu einer Studienteilnahme bewogen. Weitere drei Therapeut\*innen haben 16 bis 20 Klient\*innen innerhalb der Studie behandelt und eine therapeutische Person hat 36 Klient\*innen behandelt, die an der Studie teilgenommen haben.

### ***Prä-Post-Stichprobe (T1-T3)***

Zu Therapieende lagen Datensätze von 79 Klient\*innen vor, die von 23 Therapeut\*innen behandelt worden waren. Die Prä-Post-Stichprobe beinhaltete alle vollständigen Datensätze zu Therapiebeginn und -ende und umfasste nicht nur regulär abgeschlossene, sondern auch vorzeitig beendete Therapien. 62 Personen (78 %) haben nach Angaben der Therapeut\*innen die Therapie regulär beendet. In 17 Fällen kam es zu einem vorzeitigen Therapieende. Dabei handelte es sich bei zwei Klient\*innen um qualitätsrelevante Abbrüche. In einem Fall wurde eine schwere Arbeitsbeziehung beschrieben und in einem anderen Fall kam es zu einem frühzeitigen Therapieende durch den\*die Klient\*in ohne Abschlussgespräch und weiteren Informationen. Die 15 restlichen Therapieabbrüche wiesen einen nicht-qualitätsrelevanten Hintergrund auf. Die häufigsten Gründe für ein frühzeitiges und nicht-qualitätsrelevantes Therapieende waren die Karenzierung der Psychotherapeut\*innen, ein Wohnortwechsel oder der Beginn eines Rehabilitationsaufenthaltes oder einer kassenfinanzierten Psychotherapie.

**Dauer der Therapien bzw. Anzahl der Einheiten.** Durchschnittlich dauerten die Therapien rund ein Jahr (vgl. Tabelle 12).

**Tabelle 12**

*Dauer der Psychotherapie in Monaten und Anzahl der Einheiten*

	Therapiedauer in Monaten (n = 68)	Anzahl der Einheiten (n = 74)
Mittelwert	12.48	30.18
Median	9.43	23.00
Std.-Abweichung	9.43	21.06
Minimum	0.36	4.00
Maximum	40.90	88.00

*Anmerkung.* Informationen der Fälle mit vollständigen Datensätzen zu Therapieende.

Der längste Therapieprozess dauerte knapp dreieinhalb Jahre (41 Monate) und der kürzeste wenige Wochen. Die Therapien umfassten durchschnittlich 30 Einheiten, wobei vier Einheiten den geringsten und 88 Therapiestunden den höchsten Umfang darstellten (vgl. Tabelle 12).

***Vorhandene Zwischenmessungen in der Prä-Post-Stichprobe (T1-T2).*** Von den 79 vorhandenen Prä-Post-Datensätzen wiesen 31 Fälle (39 %) eine Verlaufsmessung (T2) auf.

Diese Therapien dauerten, wie zu erwarten länger als 25 Einheiten. Die kürzeste Therapie, die auch einen Zwischenmesszeitpunkt aufwies, lag bei 29 Einheiten. Die längste Therapie, die wiederum keine Verlaufsmessung aufwies, lag bei 30 Einheiten.

### ***Gruppenstatistische Veränderung zwischen Therapiebeginn und Therapieende (T1-T3)***

Es wurde ein t-Test für unabhängige Stichproben durchgeführt, um zu überprüfen, ob sich die Prä-Post-Stichprobe von der verbleibenden Gesamtstichprobe hinsichtlich des Alters und den Werten auf den psychometrischen Skalen zu Therapiebeginn signifikant unterschied. Nach den Ergebnissen des Kolmogorov-Smirnov-Tests war eine Normalverteilung in keiner Variable für beide Gruppen gegeben, da  $p < .05$ . Aufgrund der ausreichend großen Stichprobengröße in beiden Gruppen ( $n = 184; 79$ ) wurde die Analyse dennoch mit diesem Verfahren gerechnet. In allen Gruppenpaaren lag laut dem Levene-Test Varianzhomogenität ( $p > .05$ ) vor, weswegen der t-Test für unabhängige Stichproben verwendet wurde.

Es gab keinen statistisch signifikanten Unterschied zwischen der Prä-Post-Stichprobe und der verbleibenden Gesamtstichprobe hinsichtlich des Alters und den Werten den Skalen GSI, DEPR, ANGS, IIP-Gesamt, ESK-Gesamt und GAF ( $p > .05$ ). Jedoch unterschied sich die Gruppe mit Daten zu Therapieende (Prä-Post-Stichprobe  $M = 66.19$ ;  $SD = 12.57$ ) von der Gruppe ohne Daten zu Therapieende ( $M = 69.74$ ;  $SD = 10.78$ ) statistisch signifikant im psychologischen Wohlbefinden (Ryff-Skala) zu Therapiebeginn mit kleinem Effekt,  $t(250) = -2.247$ ,  $p = .025$ ,  $d = 0.30$ , 95 %-CI [0.03; 0.59], wobei die Personen ohne Daten zu Therapieende ein höheres psychologische Wohlbefinden zu Therapiebeginn angegeben haben. Nach der Bonferroni-Korrektur war auch dieser Unterschied nicht signifikant ( $p_{\text{korr}} = .006$ ).

Ein Chi-Quadrat-Test ergab keine statistisch signifikanten Zusammenhänge zwischen dem Geschlecht,  $\chi^2(1) = 0.335$ ,  $p = .563$  bzw. dem formalen Bildungsstand,  $\chi^2(4) = 3.440$ ,  $p = .489$  (exakter Test nach Fisher) und der Gruppenzugehörigkeit.

Bei der Überprüfung der Voraussetzungen für den abhängigen t-Test konnte anhand des Kolmogorov-Smirnov-Tests eine Normalverteilung für die Differenzwerte der Skalen GSI, Depression und Ryff-Gesamtskala festgestellt werden ( $p > .05$ ). Die Differenzwerte der Skalen Angst, IIP, ESK und GAF wiesen keine Normalverteilung auf ( $p < .05$ ). Diese Voraussetzung ist jedoch angesichts der Stichprobengröße ( $n = 79$ ) vernachlässigbar. Die Ergebnisse des t-Tests für abhängige Stichproben für den Zeitraum Therapiebeginn und Therapieende sind in Tabelle 13 abgebildet und werden in den folgenden Absätzen beschrieben.

**Veränderung der Skalen zwischen Therapiebeginn und -ende (T1-T3).** Alle ermittelten Prä-Post-Effektstärken können in den symptombezogenen Skalen als groß

eingestuft werden. Während die Gesamtskala des IIP mit Cohen's  $d = 0.86$  die geringste Effektstärke aufweist, ergab sich in der Depressions-Skala des SCL mit Cohen's  $d = 1.04$  der größte Effekt. Die Effektstärken der Skalen GSI (Cohen's  $d = 0.97$ ) und Angst (Cohen's  $d = 0.90$ ) des SCL reihten sich dazwischen ein.

Auch die Effektstärken der ressourcenbezogenen Skalen lagen im großen Bereich. Während die Existenzskala eine Prä-Post-Effektstärke von Cohen's  $d = 0.85$  aufwies, lag jene der Ryff-Skala bei Cohen's  $d = 1.01$ . Die Therapeut\*innen beurteilten das allgemeine Funktionsniveau ihrer Klient\*innen zu Therapieende signifikant höher als zu Therapiebeginn. Mit einem Cohen's  $d = 1.02$  konnte dieser Effekt als groß eingestuft werden. Alle Ergebnisse waren auch nach der Bonferroni-Korrektur signifikant ( $p_{\text{korr}} = .007$ ).

**Tabelle 13**

*Ergebnisse des gepaarten t-Tests in der Prä-Post-Stichprobe*

Skala	n	Therapiebeginn	Therapieende	t	df	p	d	95 %-CI
		M (SD)	M (SD)					
GSI	79	1.10 (0.66)	0.46 (0.45)	9.58	78	< .001	0.97	[0.64; 1.30]
DEPR	79	1.57 (0.94)	0.60 (0.57)	9.74	78	< .001	1.04	[0.70; 1.37]
ANGS	79	1.09 (0.75)	0.42 (0.49)	8.57	78	< .001	0.90	[0.57; 1.23]
IIP	79	1.48 (0.54)	1.02 (0.58)	7.94	78	< .001	0.86	[0.53; 1.18]
ESK	74	188.98 (36.27)	219.90 (34.11)	-7.74	73	< .001	0.85	[0.51; 1.19]
RYFF	71	65.80 (12.21)	78.17 (10.81)	-10.11	70	< .001	1.01	[0.66; 1.37]
GAF	73	61.74 (15.41)	77.53 (15.73)	-7.83	72	< .001	1.02	[0.67; 1.38]

*Anmerkung.* t-Test für abhängige Stichproben; n = Anzahl der Proband\*innen mit verrechenbaren Datensätzen; M = Mittelwert; SD = Standardabweichung; t = t-Wert; df = Freiheitsgrade; p = p-Wert; d = Effektstärke nach Cohen (1988); 95 % CI = 95 %-Konfidenzintervall der Effektstärke; GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; ESK = Gesamtskala der Existenz-Skala; RYFF = Gesamtskala der Ryff-Skala; GAF = Global Assessment of Functioning; Bonferroni-Korrektur ( $p_{\text{korr}} = .007$ ).

### ***Individuelle Veränderung zwischen Therapiebeginn und Therapieende***

Für die Berechnungen des Reliable Change Index (RCI) wurden die Personen mit klinisch auffälligen bzw. unauffälligen Symptomen zu Therapiebeginn auf der jeweiligen Skala gesondert untersucht.

Von den 79 Klient\*innen der Prä-Post-Stichprobe wiesen auf der GSI-Skala der SCL zu Therapiebeginn 53 Personen eine klinisch auffällige bzw. 26 Personen eine klinisch unauffällige Symptomatik auf. Von den Personen mit klinisch auffälliger blieben nach den

Berechnungen des RCI rund 8 % unverändert und weitere 2 % verschlechterten sich. Eine reliable Verbesserung wiesen rund 91 % auf und 76 % haben sich klinisch bedeutsam verändert und lagen zu Therapieende im klinisch unauffälligen Bereich (vgl. Tabelle 14). Der Anteil an Personen mit reliabel gebessertem Befinden lag im Zeitraum zwischen Therapiebeginn und Therapieende auf der Depressionsskala bei 80 % und auf der Angstskaala bei 71 %. Eine klinisch bedeutsame Veränderung wies auf der Depressionsskala das Befinden von 63 % und auf der Angstskaala von 57 % aller Personen mit klinisch auffälliger Symptomatik zu Therapiebeginn auf.

Auf der IIP-Gesamtskala hatten von den zu Therapiebeginn klinisch auffälligen Personen ( $n = 53$ ) zu Therapieende rund 60 % eine ausreichend große Veränderung im Befinden, um als reliabel gebessert eingestuft zu werden. Bei 45 % dieser Personen veränderten sich die interpersonalen Probleme zusätzlich klinisch signifikant und lagen zu Therapieende im klinisch unauffälligen Bereich. Bei rund 38 % dieser Personen blieben ihre interpersonalen Probleme unverändert und bei einer Person (2 %) haben sie sich verschlechtert (vgl. Tabelle 14).

**Tabelle 14**

*Ergebnisse des RCI bei den Personen mit zu Therapiebeginn klinisch auffälliger Symptomausprägung*

Skala	$n^a$	reliabel verschlechtert	reliabel unverändert	reliabel gebessert	klinisch signifikant gebessert
		$n$ (%)	$n$ (%)	$n$ (%)	$n$ (%)
GSI	53	1 (1.9)	4 (7.5)	48 (90.6)	40 (75.5)
DEPR	54	1 (1.9)	10 (18.5)	43 (79.6)	34 (63.0)
ANGS	51	-	15 (29.4)	36 (70.6)	29 (56.9)
IIP	53	1 (1.9)	20 (37.7)	32 (60.4)	24 (45.3)

*Anmerkung.* RCI = Reliable Change Index;  $n^a$  = Anzahl an Personen mit klinisch auffälliger Symptomatik zu Therapiebeginn;  $n$  = absolute Häufigkeit; % = prozentueller Anteil an den Personen mit klinisch auffälliger Symptomatik zu Therapiebeginn; GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskaala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme.

Bei den Personen mit zu Therapiebeginn klinisch unauffälliger Symptomatik hat sich auf der GSI-Gesamtskala der Zustand von rund 27 % signifikant gebessert (vgl. Tabelle 15). Beim Großteil blieb dieser unverändert und keine Person wies ein reliabel verschlechtertes Befinden auf. Ebenso gab es keine Person mit Verschlechterung im Befinden auf der Depressionsskala der SCL-90. Hingegen hat sich der Zustand von 12 % reliabel gebessert und

das Befinden der restlichen 88 % blieb hinsichtlich der depressiven Symptomausprägung unverändert. Bei keiner Person der Gruppe mit klinisch unauffälliger Angstsymptomatik konnte im Befinden eine reliable Verbesserung festgestellt werden. Bei 96 % blieb das Befinden bezogen auf die Ängstlichkeit unverändert und bei einer Person hat es sich reliabel verschlechtert. Ebenso reliabel verschlechtert haben sich die interpersonalen Probleme einer Person. Dagegen haben sich rund 31 % der Gruppe mit klinisch unauffälliger Eingangssymptomatik in ihren interpersonalen Fähigkeiten zu Therapieende reliabel gebessert, während 65 % darin unverändert blieben (vgl. Tabelle 15).

**Tabelle 15**

*Ergebnisse des RCI bei den Personen mit zu Therapiebeginn klinisch unauffälliger Symptomausprägung*

Skala	$n^a$	reliabel	unverändert	reliabel gebessert
		verschlechtert		
		$n$ (%)	$n$ (%)	$n$ (%)
GSI	26	0 (0)	19 (73.1)	7 (26.9)
DEPR	25	0 (0)	22 (88.0)	3 (12.0)
ANGS	28	1 (3.6)	27 (96.4)	0 (0)
IIP	26	1 (3.8)	17 (65.4)	8 (30.8)

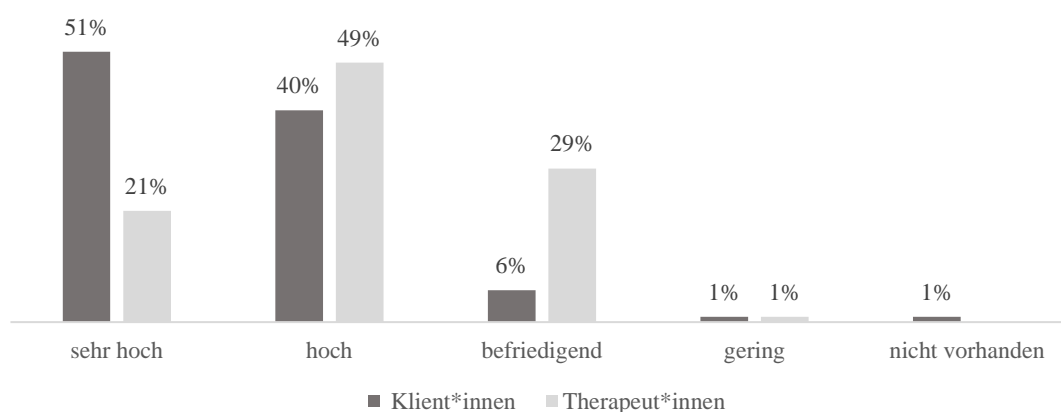
*Anmerkung.* RCI = Reliable Change Index;  $n^a$  = Anzahl der Personen mit klinisch unauffälliger Symptomatik zu Therapiebeginn;  $n$  = absolute Häufigkeit; % = prozentueller Anteil an den Personen mit klinisch unauffälliger Symptomatik zu Therapiebeginn; GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme.

### ***Therapiebeurteilung der Klient\*innen und Therapeut\*innen***

50 % der Klient\*innen ( $n = 39$ ) der Prä-Post-Stichprobe schätzten den Erfolg ihrer Psychotherapie als *sehr hoch* ein. Bei 19 % der Klient\*innen beurteilten die Therapeut\*innen den Psychotherapieerfolg als *sehr hoch* (vgl. Abbildung 4).

**Abbildung 4**

*Beurteilung des Therapieerfolges durch die Klient\*innen und deren Therapeut\*innen zu Therapieende ( $n = 79$ )*



Weitere 40 % der Klient\*innen ( $n = 31$ ) gaben an, einen *hohen* Therapieerfolg erzielt zu haben. Bei 49 % der Klient\*innen gaben auch ihre Therapeut\*innen einen *hohen* Therapieerfolg an. Einen *befriedigenden* Therapieerfolg gaben 8 % der Klient\*innen ( $n = 6$ ) an und in 29 % der Fälle gaben die Therapeut\*innen einen *befriedigenden* Therapieerfolg (vgl. Abbildung 4).

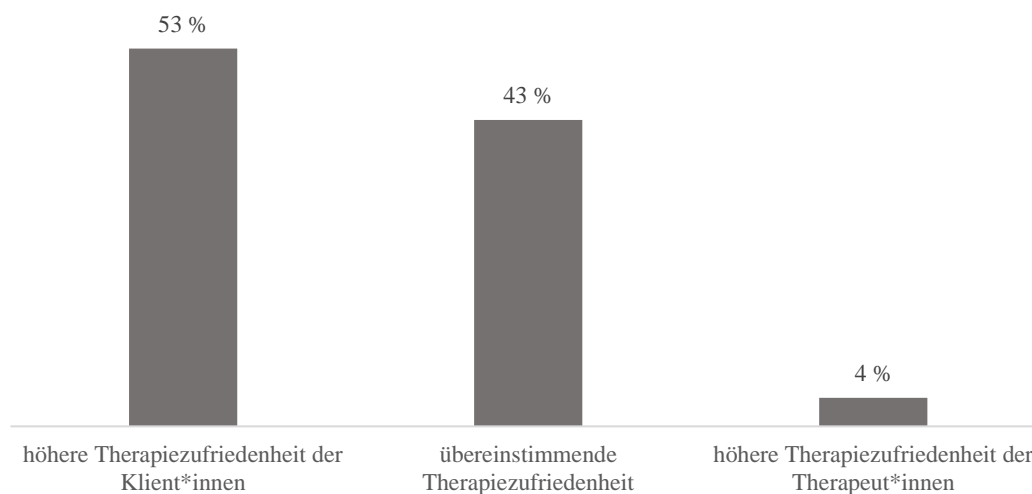
Bei 50 % der Behandlungspaare ( $n = 39$ ) kam es in der Beurteilung des Erfolges der Therapie zu einer Übereinstimmung und in 46 % der Fälle gaben die Klient\*innen einen höheren Therapieerfolg an als deren Therapeut\*innen. Bei einem Behandlungspaar (4 %) wurde der Therapieerfolg von der\*dem Therapeut\*in höher eingeschätzt als von der\*dem Klient\*in.

79 % der Klient\*innen in der Prä-Post-Stichprobe ( $n = 62$ ) waren *sehr zufrieden* mit ihrer Psychotherapie. Bei 33 % der Klient\*innen ( $n = 26$ ) gaben die Psychotherapeut\*innen an, *sehr zufrieden* mit der durchgeführten Psychotherapie gewesen zu sein. Weitere 19 % der Klient\*innen ( $n = 15$ ) waren mit der Psychotherapie *zufrieden* und in 52 % der Fälle ( $n = 42$ ) gaben die Therapeut\*innen an, mit der Psychotherapie *zufrieden* gewesen zu sein.

Bei 43 % der Behandlungspaare ( $n = 34$ ) zeigte sich eine Übereinstimmung in den Zufriedenheitsangaben und bei 53 % der Fälle ( $n = 42$ ) gaben die Klient\*innen eine höhere Zufriedenheit mit der Psychotherapie an als deren Therapeut\*innen. Bei drei Behandlungspaaren (4 %) waren die Therapeut\*innen mit der Psychotherapie zufriedener als ihre Klient\*innen (vgl. Abbildung 5).

#### Abbildung 5

Übereinstimmung der Behandlungspaare in den Beurteilungen zur Therapiezufriedenheit ( $n = 79$ )

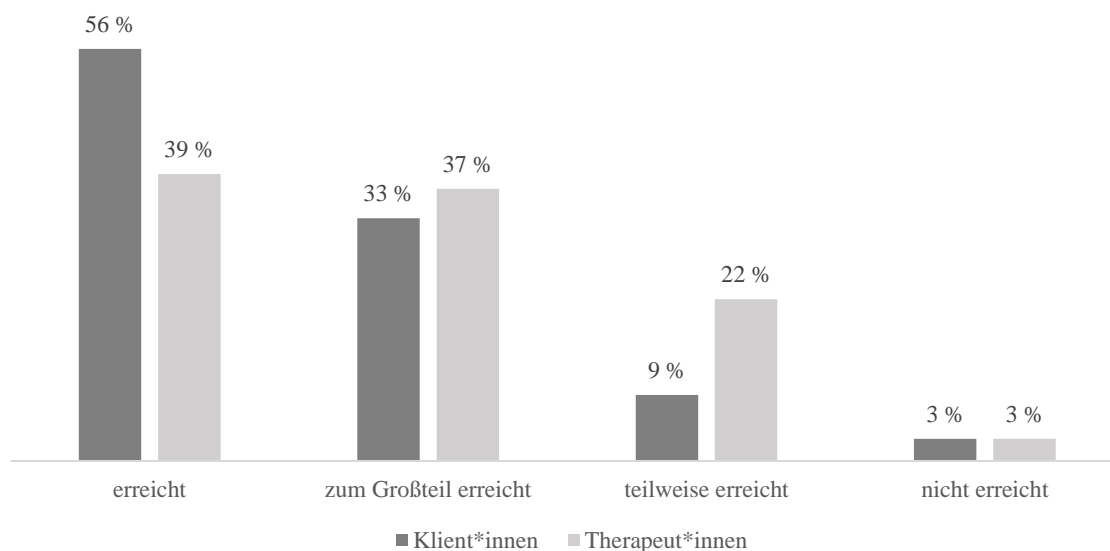


56 % der Klient\*innen ( $n = 44$ ) gaben an, die Therapieziele erreicht zu haben. In 39 % der Fälle ( $n = 31$ ) gaben die Therapeut\*innen an, die Therapieziele erreicht zu haben. Zum

Großteil wurden die Therapieziele für 33 % der Klient\*innen erreicht und in 37 % der Fälle wurden für die Therapeut\*innen die Ziele der Behandlung erreicht (vgl. Abbildung 6). Die restlichen Prozent entfallen auf die Antworten die Therapieziele teilweise erreicht (Klient\*innen = 9 %; Therapeut\*innen = 22 %) und nicht erreicht (Klient\*innen = 3 %; Therapeut\*innen 3 %).

#### Abbildung 6

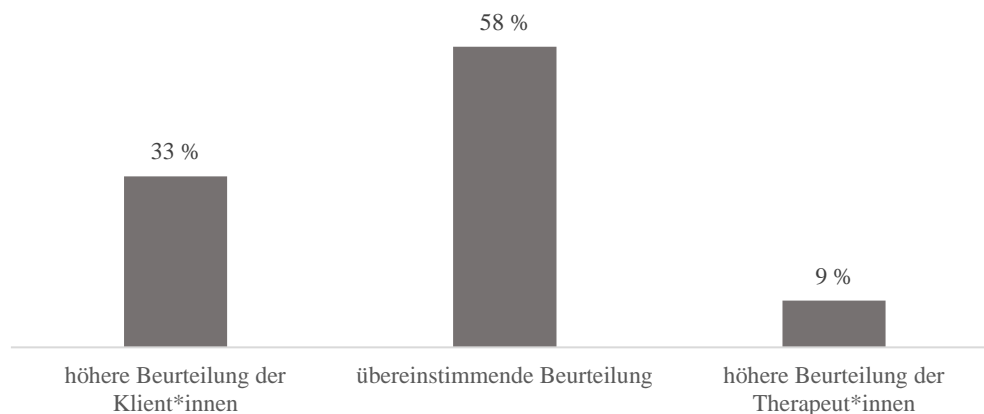
Beurteilung der Zielerreichung durch die Klient\*innen und deren Therapeut\*innen zu Therapieende ( $n = 79$ )



Bei 58 % der Behandlungspaare ( $n = 46$ ) kam es zu einer Übereinstimmung in der Einschätzung der therapeutischen Zielerreichung. In 33 % der Fälle ( $n = 26$ ) schätzten die Klient\*innen die Zielerreichung der Psychotherapie besser ein als deren Therapeut\*innen und bei sieben Behandlungspaaren (9 %) beurteilten die Therapeut\*innen die Zielerreichung besser als ihre Klient\*innen (vgl. Abbildung 7).

#### Abbildung 7

Übereinstimmung der Behandlungspaare in den Beurteilungen zum Erreichen der Therapieziele ( $n = 79$ )



### ***Kurzfristige Therapieeffekte (T1-T2)***

Zur Verlaufsmessung (T2), 25 Einheiten nach Therapiebeginn, lagen Datensätze von 105 Personen vor, die von 23 Therapeut\*innen behandelt wurden. Davon wiesen 31 Fälle Daten zu den späteren Testzeitpunkten auf. Bei 74 Personen handelte es sich um den vorläufig letzten Testzeitpunkt.

Es wurde ein t-Test für unabhängige Stichproben gerechnet, um zu überprüfen, ob sich die Verlauf-Stichprobe von der verbleibenden Gesamtstichprobe hinsichtlich ihres Alters und ihren Ausprägungen in den psychometrischen Skalen zu Therapiebeginn signifikant voneinander unterschied. Nach den Ergebnissen des Kolmogorov-Smirnov-Tests war eine Normalverteilung in beiden Gruppen in den Skalen IIP, ESK und Ryff ( $p > .05$ ) gegeben. In den anderen Skalen wies mindestens eine der Gruppen keine Normalverteilung auf ( $p < .05$ ). Aufgrund der ausreichend großen Stichprobengröße in beiden Gruppen ( $n = 158; 105$ ) wird die Analyse dennoch mittels t-Tests für unabhängige Stichproben durchgeführt. In den Gruppenpaaren der Variablen GSI und Angst lag nach dem Levene-Test keine Varianzhomogenität ( $p < .05$ ) vor, weswegen der Welch-Test verwendet wurde. Bei allen anderen Variablenpaaren wurde der t-Test für unabhängige Stichproben verwendet.

Es gab keinen statistisch signifikanten Unterschied zwischen der Verlauf-Stichprobe und der restlichen Gesamtstichprobe hinsichtlich des Alters und der Mittelwerte der Skalen GSI, Depression, Angst, IIP und ESK und RYFF ( $p > .05$ ). Jedoch unterschied sich die Gruppe mit Daten zur Verlaufsmessung ( $M = 59.94; SD = 14.99$ ) zu Therapiebeginn statistisch signifikant und mit kleinem Effekt von den Personen ohne Daten zur Verlaufsmessung ( $M = 66.11; SD = 15.29$ ) in den Ausprägungen der therapeutischen Fremdbeurteilung, ermittelt durch die GAF,  $t(250) = -3,156, p = .002, d = 0,41, 95\% -CI [0.15; 0.66]$ . Die Therapeut\*innen beurteilten die Klient\*innen mit einer Verlaufsmessung hinsichtlich ihres allgemeinen Funktionsniveaus schlechter als die Klient\*innen ohne Verlaufsmessung. Die Bonferroni-Korrektur änderte nichts an diesen Ergebnissen ( $p_{korr} = .006$ ). Laut den Ergebnissen eines Chi-Quadrat-Tests ergaben sich keine signifikanten Zusammenhänge zwischen dem Geschlecht bzw. dem formalen Bildungsstand und der Gruppenzugehörigkeit ( $p > .05$ ).

Bei der Überprüfung der Voraussetzungen für den abhängigen t-Test konnte bei den Differenzwerten der Angst- ( $p = .04$ ) und GAF-Skala ( $p = .001$ ) keine Normalverteilung festgestellt werden. Für die Differenzwerte der anderen Variablen konnte eine Normalverteilung festgestellt werden ( $p > .05$ ). Diese Voraussetzung ist angesichts der Stichprobengröße ( $n = 105$ ) vernachlässigbar. Die Ergebnisse des abhängigen t-Tests für die

Verlauf-Stichprobe sind in Tabelle 16 abgebildet. Alle überprüften Gruppenmittelwertspaare wiesen signifikante Unterschiede zwischen Therapiebeginn und der Verlaufsmessung im Sinne einer Verbesserung auf. Die Bonferroni-Korrektur änderte nichts an diesen Ergebnissen ( $p_{\text{korr}} = .007$ ). Die Effekte waren dabei in ihrer Stärke als gering bis mittel einzustufen (vgl. Tabelle 16).

**Tabelle 16**

*Ergebnisse des gepaarten t-Tests in der Verlauf-Stichprobe*

Skala	n	Therapiebeginn	Verlaufsmessung	t	df	p	d	95 %-CI
		M (SD)	M (SD)					
GSI	105	1.13 (0.60)	0.73 (0.54)	8.985	104	< .001	0.66	[0.38; 0.94]
DEPR	105	1.60 (0.85)	1.01 (0.75)	7.763	104	< .001	0.70	[0.42; 0.98]
ANGS	105	1.14 (0.72)	0.73 (0.70)	7.136	104	< .001	0.56	[0.28; 0.84]
IIP	104	1.53 (0.51)	1.33 (0.55)	5.750	103	< .001	0.40	[0.13; 0.68]
ESK	95	186.65 (35.90)	199.22 (35.97)	-4.552	94	< .001	0.35	[0.06; 0.64]
RYFF	95	67.80 (11.28)	73.14 (11.17)	-5.286	94	< .001	0.47	[0.18; 0.76]
GAF	98	59.63 (14.83)	68.74 (14.43)	-6.455	97	< .001	0.61	[0.32; 0.90]

*Anmerkung.* t-Test für abhängige Stichproben; n = Anzahl der Proband\*innen mit verrechenbaren Datensätzen; M = Mittelwert; SD = Standardabweichung; t = t-Wert; df = Freiheitsgrade; p = p-Wert; d = Effektstärke nach Cohen (1988); 95 %-CI = 95 %-Konfidenzintervall der Effektstärke; GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskaala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; ESK = Gesamtskala der Existenz-Skala; RYFF = Gesamtskala der Ryff-Skala; GAF = Global Assessment of Functioning; Bonferroni-Korrektur ( $p_{\text{korr}} = .007$ ).

### **Langfristige Therapieeffekte (T1-T4)**

Zur Katamnesemessung, 6 Monate nach Therapieende, lagen Datensätze von 31 Personen vor, die von 15 Therapeut\*innen behandelt wurden.

Es wurde ein t-Test für unabhängigen Stichproben durchgeführt, um zu überprüfen, ob sich die Katamnese-Stichprobe von der verbleibenden Gesamtstichprobe hinsichtlich des Alters und den Ausprägungen auf den psychometrischen Skalen zu Therapiebeginn und -ende signifikant voneinander unterschied. Laut den Ergebnissen des Kolmogorov-Smirnov-Tests war eine Normalverteilung zu Therapiebeginn in beiden Gruppen bei den Skalen GSI, Depression, Angst, IIP, ESK und Ryff und zu Therapieende in beiden Gruppen der Skalen IIP und Ryff gegeben ( $p > .05$ ). In allen anderen Variablen lag in mindestens einer der beiden Gruppen keine Normalverteilung vor ( $p < .05$ ). Aufgrund der ausreichend großen Stichprobengröße in beiden Gruppen ( $n = 232; 31$ ) wird die Analyse dennoch mittels parametrischem Verfahren fortgesetzt. In den Gruppenpaaren zu Therapiebeginn lag laut den

Ergebnissen des Levene-Tests für jede Skala Varianzhomogenität ( $p < .05$ ) vor. Hier wurde der t-Test für unabhängige Stichproben verwendet. In den Gruppenpaaren zu Therapieende lag für keine Skala Varianzhomogenität ( $p < .05$ ) vor, weswegen der Welch-Test verwendet wurde.

Die Katamnese-Stichprobe unterscheidet sich zu Therapiebeginn statistisch signifikant von der restlichen Gesamtstichprobe auf der ESK-Gesamtskala mit kleinem Effekt,  $t(256) = 2,435$ ,  $p = .02$ ,  $d = 0.47$ , 95 %-CI [0.09; 0.86]. Nach der Bonferroni-Korrektur war auch dieser Unterschied nicht signifikant ( $p_{\text{korr}} = .006$ ).

Die Gruppe mit Datensätzen zur Katamnesebewertung unterschied sich zu Therapieende statistisch signifikant von jener Gruppe ohne Katamnesebewertung hinsichtlich der Gruppenmittelwerten aller Skalen, außer der IIP-Gesamtskala. Jedoch war nach der Bonferroni-Korrektur der Unterschied in den Gruppenmittelwerten der Angst- ( $p = .05$ ), RYFF- ( $p = .01$ ) und GAF-Skala ( $p = .02$ ) nicht mehr signifikant ( $p_{\text{korr}} = .006$ ). Personen mit Datensätzen zum Katamnesezeitpunkt wiesen zu Therapieende nach der Alphaadjustierung signifikant bessere Werte mit jeweils knapp großem Effekt in der angegebenen Globalbelastung (GSI,  $d = 0.76$  (95%-CI [0.29; 1.24]), der Depressivität (DEPR,  $d = 0.74$  (95%-CI [0.27; 1.22])) und in ihren existenziellen Fähigkeiten (ESK,  $d = 0.76$  (95%-CI [0.29; 1.26])) auf als jene Personen, die an der Katamnesebewertung nicht mehr teilnahmen.

Die Ergebnisse eines Chi-Quadrat-Tests ergaben keine signifikanten Zusammenhänge zwischen dem Geschlecht bzw. dem formalen Bildungsstand und der Gruppenzugehörigkeit ( $p > .05$ ).

Bei der Überprüfung der Voraussetzungen für den abhängigen t-Test konnte für alle Skalen bis auf die Gesamtskala der ESK ( $p < .05$ ) eine Normalverteilung der Differenzwerte zwischen Therapiebeginn und Katamnese festgestellt werden. Diese Voraussetzung ist angesichts der Stichprobengröße vernachlässigbar. Die Ergebnisse der statistischen Überprüfung der Mittelwertsdifferenzen der Skalen zwischen Therapiebeginn und Katamnese sind in Tabelle 17 abgebildet. Alle überprüften Gruppenmittelwertspaare wiesen signifikante Unterschiede zwischen Therapiebeginn und der Katamnesebewertung im Sinne einer Verbesserung auf. Die Bonferroni-Korrektur änderte nichts an diesen Ergebnissen ( $p_{\text{korr}} = .008$ ). Die Effekte waren dabei mit Cohen's  $d > 1.0$  als groß einzustufen und sind im Vergleich zu Therapieende noch gestiegen (vgl. Tabelle 17).

90 % der Personen der Katamnese-Stichprobe gaben an, dass sie den Therapieerfolg als sehr bzw. ziemlich nachhaltig einschätzen würden. Knapp über die Hälfte gab zudem an, dass sich seit Therapieende etwas an ihrer Wohnungs-, Arbeits-, oder Familiensituation

geändert habe und knapp über drei Viertel hätten seit Therapieende keine weitere Psychotherapie mehr in Anspruch genommen.

**Tabelle 17**

*Ergebnisse des gepaarten t-Tests in der Katamnese-Stichprobe*

Skala	n	Therapiebeginn	Katamnese	t	df	p	d	95 %-CI
		M (SD)	M (SD)					
GSI	31	0.90 (0.54)	0.27 (0.25)	6.352	30	< .001	1.17	[0.62; 1.72]
DEPR	31	1.41 (0.91)	0.37 (0.50)	6.392	30	< .001	1.15	[0.60; 1.70]
ANGS	31	0.96 (0.66)	0.23 (0.29)	6.247	30	< .001	1.11	[0.56; 1.66]
IIP	30	1.40 (0.52)	0.77 (0.43)	7.259	29	< .001	1.23	[0.67; 1.79]
ESK	29	201.83 (33.82)	237.38 (23.91)	-6.272	28	< .001	1.05	[0.49; 1.61]
RYFF	29	67.93 (11.31)	84.28 (7.44)	-8.647	28	< .001	1.44	[0.85; 2.03]

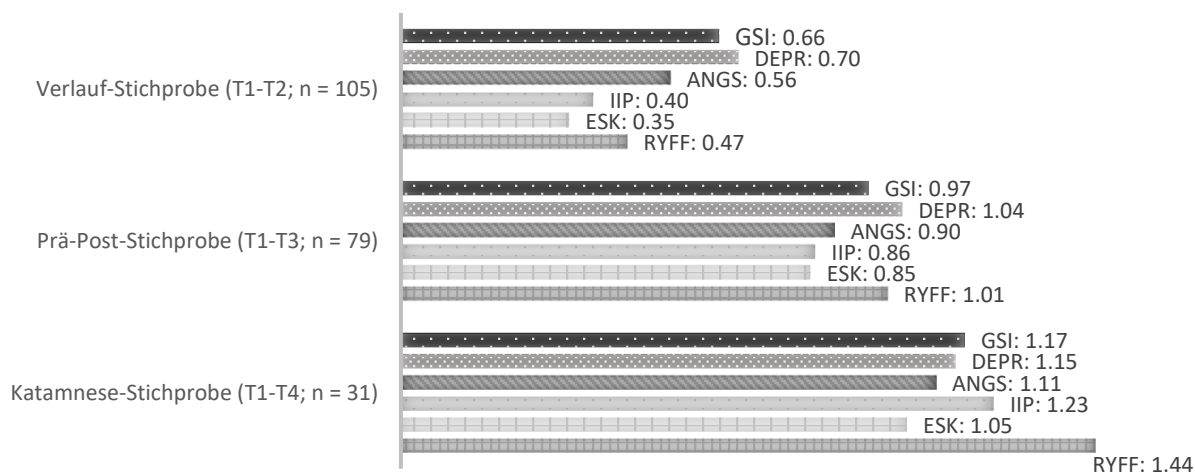
*Anmerkung.* t-Test für abhängige Stichproben; n = Anzahl verrechenbarer Skalen; M = Mittelwert; SD = Standardabweichung; t = T-Wert; df = Freiheitsgrade; p = p-Wert; d = Effektstärke nach Cohen (1988); 95 % CI = 95 %-Konfidenzintervall der Effektstärke; GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskaala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; ESK = Gesamtskala der Existenz-Skala; RYFF = Gesamtskala der Ryff-Skala; Bonferroni-Korrektur ( $p_{\text{korrt}} = .008$ ).

### **Zusammenfassung der Therapieeffekte**

Abbildung 8 fasst die ermittelten Effektstärken in den einzelnen Verfahren der gezogenen Teilstichproben visuell zusammen.

**Abbildung 8**

*Effektstärken in den einzelnen Verfahren über die Testzeitpunkte hinweg*



*Anmerkung.* GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskaala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; ESK = Gesamtskala der Existenz-Skala; RYFF = Gesamtskala der Ryff-Skala; Effektstärken nach Cohen's d.

## Explorative Analysen

**Completer- und ITT-Analysen der GSI-Werte.** Es wurde ein explorativer t-Test für abhängige Stichproben durchgeführt, um die Veränderung im globalen Symptomerleben einer *Completer*-Stichprobe, welche nur Personen mit regulär abgeschlossenen Therapieerläufen beinhaltet ( $n = 62$ ), und die Veränderung einer *ITT*-Stichprobe, welche alle in die Studie aufgenommenen Personen beinhaltet ( $n = 263$ ), zu untersuchen. Die fehlenden Datensätze zu Therapieende der ITT-Stichprobe wurden mit der LOCF-Methode imputiert (Hiller et al., 2009; Steffanowski et al., 2011; Stieglitz & Hiller, 2014; Strauss et al., 2015). Die Ergebnisse des gepaarten t-Tests der Completer- bzw. ITT-Stichprobe sind in Tabelle 18 abgebildet.

**Tabelle 18**

*Ergebnisse eines explorativen gepaarten t-Tests in der Completer- bzw. ITT-Stichprobe*

Skala	n	Therapiebeginn	Therapieende	t	df	p	d	95 % CI
		M (SD)	M (SD)					
GSI <sub>c</sub>	62	1.06 (0.66)	0.34 (0.33)	10.28	61	< .001	1.09	[0.709; 1.471]
GSI <sub>ITT</sub>	263	1.11 (0.67)	0.82 (0.67)	9.83	262	< .001	0.43	[0.257; 0.603]

*Anmerkung.* t-Test für abhängige Stichproben; n = Anzahl der Proband\*innen mit verrechenbaren Datensätzen; M = Mittelwert; SD = Standardabweichung; t = t-Wert; df = Freiheitsgrade; p = p-Wert; d = Effektstärke nach Cohen (1988); 95 %-CI = 95 %-Konfidenzintervall der Effektstärke; GSI<sub>c</sub> = Global Severity Index einer Completer-Stichprobe; DEPR = Depressionsskala der SCL-90; ANGS = Angstskala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; ESK = Gesamtskala der Existenz-Skala; RYFF = Gesamtskala der Ryff-Skala; GAF = Global Assessment of Functioning.

In beiden Gruppen kam es zu einer statistisch signifikanten Veränderung zwischen Therapiebeginn und Therapieende. Die Effektstärken unterschieden sich jedoch deutlich voneinander. Es konnte ein großer Effekt bei den Klient\*innen mit regulären Therapieabschlüssen und ein kleiner Effekt in der ITT-Stichprobe gefunden werden.

**Prädiktoren des Therapieergebnisses.** Mit einer explorativen, multiplen-linearen Regressionsanalyse wurde untersucht, ob die therapeutische Beziehung (WAI), die Therapiedauer (Anzahl an Einheiten), die einzelnen Therapeut\*innen, wobei jedem\*r Therapeut\*in eine Zahl zugeordnet wurde und die Ausprägung auf der GSI-Skala zu Therapiebeginn Prädiktoren für das Therapieergebnis (GSI<sub>TE-TB</sub>) waren.

Das Modell hatte mit einem  $R^2 = .59$  (korrigiertes  $R^2 = .57$ ) eine hohe Anpassungsgüte und erklärt statistisch signifikant rund 60 % der Ergebnisvarianz,  $F(4,66) = 24.04$ ,  $p < .001$ . Sowohl die Eingangsbelastung als auch die therapeutische Beziehung sagten statistisch signifikant das Kriterium Symptomveränderung GSI<sub>TE-TB</sub> voraus. Insbesondere die Globalbelastung zu Therapiebeginn (GSI<sub>TB</sub>) hatte einen bedeutenden Einfluss auf die Größe

der Symptomveränderung,  $\beta = -0.66$ ;  $t(70) = -9,12$ ;  $p < .001$ . Weniger stark, aber statistisch ebenso signifikant wurde die Symptomveränderung durch die therapeutische Beziehung beeinflusst,  $\beta = -0.28$ ;  $t(70) = -2.67$ ;  $p = .01$ . Bei Zunahme des Wertes der beiden Prädiktoren verringert sich der Wert der Symptomdifferenz. Die Symptomdifferenz konnte Werte zwischen -4 (größtmögliche positive Veränderung) bis 4 (größtmögliche negative Veränderung) annehmen. Die beiden anderen Prädiktoren, Therapeut\*innen und Anzahl an Therapieeinheiten hatten keinen signifikanten Einfluss auf die Symptomveränderung.

**Zusammenhang der ESK-Gesamtskala mit den anderen Skalen.** Die Ergebnisse einer Korrelationsanalyse sprachen für einen signifikanten Zusammenhang der ESK-Gesamtskala mit den restlichen symptom- bzw. ressourcenbezogenen Skalen mit hohem Effekt (vgl. Tabelle 19).

**Tabelle 19**

*Korrelativer Zusammenhang der ESK-Gesamtskala mit den anderen Skalen zu Therapiebeginn (N = 258)*

Skala	ESK	GSI	DEPR	ANGS	IIP
GSI	-.749**	-			
DEPR	-.725**	.905**	-		
ANGS	-.578**	.863**	.719**	-	
IIP	-.718**	.666**	.632**	.490**	-
RYFF	.756**	-.642**	-.646**	-.538**	-.655**

*Anmerkung.* GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ANGS = Angstskala der SCL-90, IIP = Gesamtskala des Inventars zur Erfassung interpersonaler Probleme; ESK = Gesamtskala der Existenz-Skala; RYFF = Gesamtskala der Ryff-Skala; \*\*  $p < .001$ ;  $r > .50$  = großer Effekt nach Cohen (1988).

**Qualitative Ergänzungen - Verstehen durch Verstanden werden.** Bei einer qualitativen Analyse der bisher ausgefüllten offenen Eingabefelder im Klient\*innenfragebogen zu negativen Veränderungen durch die Therapie ( $n = 9$ ) bezogen sich die Angaben auf *indirekt* negative Nebenwirkungen durch positive Veränderungen. Ein neuer Umgang mit sich selbst (z.B. „*wachsendes Selbstwertgefühl, steigendes Selbstbewusstsein*“) veränderte auch die eigene Verortung in der Umwelt, durch deutlichere Abgrenzung oder Äußerung der eigenen Meinung („*das stört manchen Personen aus meinem näheren Umkreis*“). Zudem wurde eine wahrgenommene Erschöpfung durch die Auseinandersetzung mit teils als negativ erlebten Inhalten („Beschäftigung mit negativen Gedanken“) beschrieben.

Die Fragen, welche Besonderheiten allgemein bzw. welche Ereignisse außerhalb der Therapie für den Therapieverlauf entscheidend gewesen sein könnten, wurde von 92 Personen beantwortet. Des Öfteren fand die Beschreibung einer gelungenen therapeutischen Arbeitsbeziehung Eingang in den Kommentaren der Klient\*innen. Es wäre entscheidend

gewesen, sich in der Therapie gut aufgehoben gefühlt zu haben, um vertrauen und sich dadurch öffnen zu können. Manche beschrieben die therapeutische Beziehung auch als Begegnung auf Augenhöhe, in der gemeinsam eruiert wurde anstatt Lösungen vorgegeben zu bekommen. Auch ein ausgewogenes Maß an Nähe und Distanz und das Einfühlungsvermögen der Therapeut\*innen wurden als bedeutsam für eine gelungene Therapie hervorgehoben. Durch diese Ingredienzen sei ein Verstehen durch Verstanden-werden möglich geworden. Hilfreich seien auch regelmäßige Therapieeinheiten gewesen.

Weitere Antworten betrafen Gegebenheiten im sozialen Umfeld der Klient\*innen. Sowohl positive als auch negative Veränderungen, wie eine Geburt, ein Todesfall, eine Trennung, eine neue Partnerschaft, die Pflege von Angehörigen oder ein Wohnortwechsel wurden als wichtige Ereignisse angeführt. Auch berufliche Belange, wie Jobverlust bzw. -wechsel, Arbeitsplatzprobleme, Stress und Pensionierung sowie der Beginn oder die Beendigung von Ausbildungen wurden als wichtige Themen beschrieben. Im Laufe der Studie wurden immer öfter auch Themen rund um die *Covid-19*-Pandemie angeführt. Dabei fanden die Veränderung des Alltags, die niedrigere, höhere oder andere Arbeitsbelastung und das intensive familiäre Zusammenleben Erwähnung in den offenen Eingabefeldern.

## **Diskussion**

Die Ergebnisinterpretation wird in zwei Teile gegliedert. Zum einen erfolgt eine ergebnisspezifische Diskussion, die die zwei großen Fragestellungen der Arbeit und die dazugehörigen Fragen bzw. Hypothesen schrittweise erörtert. Im zweiten Teil werden studienübergreifende Überlegungen angestellt, die die Ergebnisse dieser Arbeit in einen größeren Kontext setzen sollen.

### **Ergebnisspezifische Diskussion**

#### ***Therapeut\*innenstichprobe***

Wegen der fehlenden Rückläuferhebung bei den Therapeut\*innen können zur besseren Einschätzung der erreichten Teilnehmer\*innen die 424 fachspezifischen Ausbildungskandidat\*innen bzw. die 224 Psychotherapeut\*innen in Ausbildung unter Supervision der GLE-Ö (Stichtag 01.06.2020) als Referenzrahmen herangezogen. Mit bislang 37 Anmeldungen zur Studie konnten rund 17 % der Psychotherapeut\*innen mit Supervisionsstatus der GLE-Ö zu einer Teilnahme bewogen werden. Der prozentuelle Anteil dürfte noch größer sein, wenn nur jene Psychotherapeut\*innen berücksichtigt werden, die in freier Praxis tätig sind. Dazu wurden jedoch keine spezifischen Informationen gefunden. Die teilnahmebereiten Therapeut\*innen der Studie entsprachen hinsichtlich der erhobenen soziodemographischen Daten in etwa der Gruppe von Ausbildungskandidat\*innen der GLE-Ö

im Jahr 2020. Es ließen sich sowohl eine ähnliche Geschlechterverteilung mit rund 80 % weiblichen Personen sowie vergleichbare Altersgruppen mit einem Durchschnitt von knapp über 40 Jahren finden. Als positiv hervorzuheben ist, dass sich die teilnehmenden Therapeut\*innen auf ganz Österreich verteilten, wobei ein Schwerpunkt im westlichen Raum (Tirol / Vorarlberg) festzustellen ist. Die Quellenberufe der Therapeut\*innen hatten eine ähnliche Vielfalt wie es bei den Psychotherapeut\*innen generell in Österreich üblich ist (Sagerschnig & Valady, 2021).

### ***Berichte der Therapeut\*innen zur Studienteilnahme***

**Teilnahmegründe.** Das Gelingen von naturalistischen Psychotherapiestudien hängt in hohem Maße von teilnahmebereiten und motivierten Therapeut\*innen ab (Castonguay et al., 2015; Figl & Laireiter, 2022; Strauss et al., 2015; Taubner et al., 2016). Aus diesem Grund ist es angeraten, sich mit den Vorbehalten und Motiven für eine Studienteilnahme von Psychotherapeut\*innen auseinanderzusetzen. Die Ergebnisse der qualitativen Inhaltsanalyse der Erfahrungsberichte der Therapeut\*innen sollen Aufschluss über die Studienakzeptanz vorliegender Studie geben und werden in den nächsten Absätzen behandelt.

Eine häufig erwähnte Motivation der Therapeut\*innen war, einen Beitrag zur Wissenschaft leisten zu wollen, der gleichzeitig ihren eigenen therapeutischen Verfahrenshintergrund nach innen und nach außen stärken sollte. Persönlich erhofften sich einige Therapeut\*innen eine berufliche Weiterentwicklung durch diagnostische und forschungsbezogene Inhalte sowie zusätzliche Informationsquellen zum therapeutischen Prozess. Zu ähnlichen Ergebnissen bei österreichischen Psychotherapeut\*innen kamen auch Felber und Margreiter (2007). Neben subjektiven Anreizen, wie z.B. der Erhalt von Rückmeldungen oder das Schreiben der fachspezifischen Abschlussarbeit im Zuge der Forschung, können auch die übergeordneten Studienziele, wie die wissenschaftliche Positionierung der Psychotherapie im Allgemeinen und jene des eigenen Verfahrens im Besonderen das Interesse für eine Studienteilnahme wecken (Felber & Margreiter, 2007).

Befürchtungen, die gegen eine Studienteilnahme sprachen, betrafen neben den Anforderungen des Studiendesigns und einer damit einhergehenden Ressourcenfrage auch die durch eine Teilnahme mögliche Beeinflussung des therapeutischen Prozesses und die (Un-) Vereinbarkeit von quantitativer Forschung mit dem phänomenologischen Hintergrund des existenzanalytischen Verfahrens. Auch andere Studien zur Einstellung von österreichischen Psychotherapeut\*innen zur Forschung bestätigten eine wahrgenommene Diskrepanz zwischen den Weltbildern vieler psychotherapeutischer Verfahren und dem quantitativen Forschungsmainstream. Zusätzlich wurden ein zu hoher (zeitlicher) Ressourcenaufwand und

eine Mindestanforderung an erbrachten Leistungen – ebenso wie bei anderen Studien – als bedeutender Grund für eine Nicht-Teilnahme von Psychotherapeut\*innen diskutiert (Castonguay et al., 2015; Figl & Laireiter, 2022; Strauss et al., 2015; Taubner et al., 2016).

Die meisten Therapeut\*innen der vorliegenden Studie berichteten, dass die Studienintegration in den klinischen Alltag nach einer Eingewöhnungsphase gut möglich gewesen wäre. Zusätzlich bot die Studie sowohl quantitative als auch qualitative Formen der Erhebung und der Auswertung an, was von den Therapeut\*innen mehrheitlich als positiv erlebt wurde.

**Herausforderungen.** Schwierigkeiten traten gehäuft zu Beginn der Teilnahme auf. Sie bezogen sich auf den passenden Zeitpunkt und die angemessene Form der Studienanwerbung von Klient\*innen, die Verwendung der IDCL und bei manchen Therapeut\*innen auch auf den Umgang mit quantitativen Testergebnissen. Aufgrund der Möglichkeit der niederschweligen Kontaktaufnahme mit dem Forschungsteam und den regelmäßigen Vernetzungstreffen zu studienrelevanten Themen konnte den Herausforderungen jedoch gut begegnet werden, und es stellte sich bald Routine ein. Auch in anderen Studien wird die große Bedeutung einer transparenten und fortwährenden Kommunikationsmöglichkeit unter den Forschungsbeteiligten für eine erfolgreiche Studiendurchführung thematisiert. Dabei ist ein kollegiales Verhältnis, das gegenseitiges Vertrauen schafft und Meinungen auf gleicher Ebene austauschen lässt, als förderlich hervorzuheben. Praxisorientierte Forschung soll einem *Bottom-Up*-Ansatz entsprechen, bei dem die Beteiligten einer Studie sich gegenseitig beeinflussen können und auch sollen (Castonguay et al., 2015; Figl & Laireiter, 2022; Strauss et al., 2015). Erlebte Schwierigkeiten bei der Studienanwerbung von Klient\*innen sind auch in der Psychotherapieforschungsliteratur bekannt. Psychotherapeut\*innen haben das Gefühl, dass nicht sie es sein dürfen, die ein Anliegen haben. Dieser Eindruck kann bei Psychotherapeut\*innen in Ausbildung zusätzlich noch stärker entstehen (Castonguay, 2013; Strauss et al., 2015). Einige Therapeut\*innen beschrieben, dass die Betonung der absoluten Anonymität und Freiwilligkeit mit der Möglichkeit eines jederzeitigen Abbruches vielen Klient\*innen die Entscheidung erleichtert hätte.

**Umgang mit Rückmeldungen der Fragebogenergebnisse.** Anhand der Erfahrungsberichte entstand der Eindruck, dass die meisten Therapeut\*innen Rückmeldungen gezielt bei Klient\*innen mit unklaren Diagnosen und schwierigen Verläufen verwendeten. Auch in den Forschungsprojekten der TK- und der QS-PSY-BAY-Studie wurden Rückmeldungen eingesetzt. In letzterer wurden, wie in vorliegender Studie auch, keine

weiteren Untersuchungen dazu angestellt, ob Rückmeldungen einen Einfluss auf Therapieeffekte hätten. Im Gegensatz dazu war diese Frage das zentrale Anliegen der TK-Studie. Es konnten jedoch keine statistisch signifikanten Unterschiede in den Therapieeffekten jener Klient\*innen gefunden werden, deren Therapeut\*innen Rückmeldungen erhielten, im Vergleich zu jener Gruppe, die keine Rückmeldung erhielt, wobei auch nicht erhoben wurde, ob und wie diese Rückmeldungen verwendet wurden und wie sehr der klinische Eindruck der Therapeut\*innen mit diesen übereinstimmten (Steffanowski et al., 2011; Wittmann & Steffanowski, 2011).

In anderen Studien konnte jedoch ein positiver Effekt der Rückmeldungen auf die Therapieverläufe gefunden werden (Castonguay et al., 2015). Durch Rückmeldungen der Fragebogenergebnisse ergibt sich eine weitere Informationsquelle zum Zustand der Klient\*innen und zum bisherigen Behandlungsprozess, die klinisch genutzt werden kann (Strauss et al., 2015). So berichteten Therapeut\*innen der vorliegenden Studie, dass durch die Verwendung der Rückmeldungen, auch der IDCL, ein achtsamerer Umgang mit und ein gefestigteres Wissen über klinische Störungsbilder und die einzelnen Symptome entstanden wäre.

Durch die Rückmeldungen der Fragebögen erhalten die Daten nicht erst nach abgeschlossener Erhebungszeit bei der Studiendatenauswertung einen Nutzen. Dieses Vorgehen vermittelt bestenfalls, dass Studienbeteiligte nicht nur Daten generieren, sondern diese auch ehest möglich rückgemeldet bekommen. Für ein aussichtsreiches, naturalistisches Forschungsvorhaben sollte Motivationsarbeit bei teilnehmenden Psychotherapeut\*innen und Klient\*innen als unerlässlich gelten (Castonguay et al., 2015).

### ***Klient\*innenstichprobe***

Die vorläufige Rücklaufquote, welche durch die bisher vorliegenden Aufzeichnungen von 22 Therapeut\*innen ermittelt wurde, belief sich auf 69 % all jener Personen, die die Einschlusskriterien (Volljährigkeit und intellektuell nicht beeinträchtigt) erfüllt hatten. Der Anteil der teilnahmebereiten Personen an allen Personen, die in den Praxen vorstellig und in den Listen vermerkt wurden, betrug 63 %. In der Literatur werden Rücklaufquoten bei prospektiven Längsschnittstudien zwischen 25 % und 38 % berichtet (Gallas et al., 2008; Strauss et al., 2015; von Wyl, Cramer, Koemeda, Tschuschke & Schulthess, 2013). Die Rücklaufquote in einer vorherigen, retrospektiven Querschnittstudie der GLE-Ö belief sich auf rund 78 % (Laireiter et al., 2013).

Aus den Rückmeldungen einiger Therapeut\*innen ging hervor, dass die Klient\*innen zum Teil vorselegiert wurden. So wurde stark belasteten Personen trotz erfüllter

Einschlusskriterien eine Studienteilnahme eher nicht angeboten. Über das Vorgehen von Therapeut\*innen, bei Klient\*innen aufgrund ihres klinischen Zustandsbildes auf eine Studienanwerbung zu verzichten, berichteten auch die Studienautor\*innen der TK-Studie (Wittmann et al., 2011). Der am häufigsten angegebene Grund vonseiten der Therapeut\*innen keine Studienteilnahme anzubieten war die aktuelle Belastung der Klient\*innen gefolgt von der Befürchtung, dass eine Studienteilnahme negative Auswirkungen auf die Therapie haben könnte (Strauss et al., 2015; Wittmann et al., 2011).

Jedenfalls hat sich in vorliegender Studie gezeigt, dass teilnahmebereite Personen statistisch signifikant durchschnittlich um zehn Jahre jünger waren als nicht-teilnahmebereite Klient\*innen. Das Ergebnis der statistischen Analyse wurde durch die Erfahrungsberichte der Therapeut\*innen gestützt, aus denen ebenso hervorging, dass ältere Personen tendenziell eine Studienteilnahme ablehnten. Zudem wurde von den Therapeut\*innen berichtet, dass das Störungsbild und die Symptomstärke einen Einfluss auf die Studienteilnahme der Klient\*innen gehabt hätten. In der TK-Studie wurden die Gründe für eine Studienablehnung vonseiten der Klient\*innen erhoben. Dabei zeigte sich, dass 34 % keine zusätzliche Zeit aufwenden wollten oder konnten, 20 % dem Datenschutz misstrauten und 19 % keine Fragebögen ausfüllen wollten. 11 % der befragten Klient\*innen führten ihren schlechten Zustand als Grund für eine Studienablehnung an (Altmann et al., 2016; Strauss et al., 2015; Wittmann et al., 2011). Ähnliche Punkte wurden auch in vorliegender Studie in den Erfahrungsberichten der Therapeut\*innen angegeben, wobei daraus nicht hervorgeht, wie häufig diese Gründe angeführt werden.

**Fragebogenrücklauf und Abbruch.** Es ist erneut zu betonen, dass der bisherige Rücklauf der Fragebögen und eine erste Abbruchdiagnostik nur vor dem Hintergrund der laufenden Studie interpretiert werden können. Erschwerend kommt hinzu, dass keine systematische Erhebung von Abbruchfällen durch eine gezielte Befragung der Therapeut\*innen erfolgte. Informationen dazu lagen nur bei Personen vor, die ein vorzeitiges Therapieende aufwiesen und dennoch bereit waren, an der Abschlussdiagnostik teilzunehmen. Vereinzelt wurden Anmerkungen zu Therapie- oder Studienabbrüchen von den Therapeut\*innen in den Strukturierungslisten, in der sie den Rücklauf vermerkten, gemacht.

Die vorläufig eindeutig identifizierbaren Abbrüche deuten auf qualitätsrelevante Gründe vor allem in den ersten Einheiten der Therapie hin. In der Anfangsphase der Therapie orientieren sich Klient\*innen und Therapeut\*innen, ob die Indikation einer Therapie angeraten und eine Zusammenarbeit passend wäre. Bei den qualitätsrelevanten Abbrüchen zu Therapiebeginn könnte es sich in vorliegender Studie auch um Abbrüche mit ähnlichem

Hintergrund gehandelt haben. Zu Therapieende lagen die meisten Informationen zu Abbrüchen vor. In dieser Phase wurde vor allem von *nicht-qualitätsrelevanten Therapieabbrüchen* berichtet, also beispielsweise Wohnortwechsel oder Bewilligung von kassenfinanzierten Versorgungsleistungen. Denkbar ist, dass Klient\*innen mit einem nicht-qualitätsrelevanten Abbruchgrund eher dazu bereit waren, dennoch den Fragebogen zu Therapieende auszufüllen.

Abbruchraten werden in Studien aus der Literatur sehr unterschiedlich berichtet. In einer Überblicksarbeit von Altmann und Kolleg\*innen (2016) zu naturalistischen Studien im deutschsprachigen Raum wurden nur problematische, also qualitätsrelevante, Abbrüche angeführt. Diese können jedoch sehr unterschiedlich definiert werden, was die Vergleichbarkeit erschwert (Altmann et al., 2016). Eindeutiger sind die Zahlen zum Fragebogenrücklauf. Während die PAP-S- und TRANS-OP-Studien einen Prä-Post-Stichprobenanteil von rund 75 % des Anfangsdatensatzes aufwiesen, lag dieser Wert in der QS-PSY-BAY-Studie mit 18 % und in der TK-Studie mit 37 % weit darunter. Die Autor\*innen erklären sich den hohen Schwund unter anderem mit den zum Teil noch laufenden Therapien zum Zeitpunkt der (Zwischen-) Analyse (Steffanowski et al., 2011; Wittmann & Steffanowski, 2011). Das Problem von Abbrüchen ist in klinischen Längsschnittstudien ein allgemeines und in naturalistischen Studiendesigns im Besonderen ein großes Thema (von Wyl et al., 2013). Vorliegende Studie lag mit einem Prä-Post-Stichprobenanteil von rund 30 % am Gesamtdatensatz zwischen den beiden letztgenannten Studien. Dieser Umstand ist insofern beachtlich, als dass es sich bei allen anderen Beispielstudien um Kooperationsprojekte zwischen Sozialversicherungsträgern und anderen gesundheitsrelevanten Einrichtungen handelt und die Studie der GLE-Ö nur von einem einzelnen Institut durchgeführt wird. Dies hebt die Bedeutung von Forschungsk Kooperation bei Studien im naturalistischen Setting insgesamt hervor (Castonguay et al., 2015; Strauss et al., 2015). Denkbar ist auch, dass Klient\*innen von privat-finanzierten Therapien - wie es in vorliegender Studie der Fall ist - weniger bereit sind, zusätzlich an Studien über einen längeren Zeitraum hinweg teilzunehmen. Hingegen wäre es bei kassenfinanzierten Therapien vorstellbar, dass zukünftig das Ausfüllen von (kurzen) Fragebögen in das therapeutische Prozedere integriert wird (Strauss et al., 2015).

**Studienakzeptanz.** Aufgrund der hohen Schwundraten bei praxisorientierten Studien werden immer wieder auch Studienakzeptanz und Zumutbarkeit des Studiendesigns mituntersucht. Die Autor\*innen der TK- und der QS-PSY-BAY Studie ermittelten, dass der Großteil der Studienteilnehmer\*innen eine Überprüfung der Psychotherapie als wichtig

erachteten und der zusätzliche Zeitaufwand für das Ausfüllen der Fragebögen vertretbar und die Fragen verständlich gewesen wären (Steffanowski et al., 2011; Wittmann et al., 2011). Auch in vorliegender Studie schätzten die Mehrheit der Klient\*innen die Fragebogenlänge als angemessen ein und knapp über ein Viertel fand diese etwas zu lang.

Zur Beurteilung der Teilnahmemotivation der Klient\*innen wurden diese nach dem Grund für ihre Studienzusage befragt. Die meisten Teilnehmer\*innen sahen in einer Studienteilnahme einen persönlichen Vorteil. Sie wollten mit den Fragebögen mehr über sich erfahren und erhofften sich, dadurch den Therapiefortschritt sichtbar machen zu können. Dahinter reihten sich auch soziale Gründe ein, wie etwa andere damit unterstützen zu wollen oder den Therapeut\*innen einen Gefallen zu tun.

Befürchtungen vonseiten der Psychotherapeut\*innen, dass sich das (wiederholte) Vorlegen der Fragebögen negativ auf den Therapieprozess auswirken könnte, lassen sich nicht bestätigen (Strauss et al., 2015; Taubner et al., 2016). In vorliegender Studie konnte sogar eine statistisch signifikante, unmittelbare Verbesserung der Stimmung im Zuge des Ausfüllens des Fragebogens mit kleinem Effekt gefunden werden.

Die Daten aus den QS-PSY-BAY- und TK-Studien und der vorliegenden Studie deuten auf eine gute Akzeptanz von Untersuchungen in der psychotherapeutischen Routineversorgung bei Klient\*innen hin. Ähnlich wie bei einigen Therapeut\*innen werden auch von einigen Klient\*innen die zusätzlichen Informationsquellen geschätzt. Die Mehrheit der an den naturalistischen Studien teilnehmenden Klient\*innen empfanden sowohl die Fragebögen als auch die Rückmeldungen für den therapeutischen Prozess anregend und hilfreich (Altmann et al., 2016; Strauss et al., 2015; Wittmann & Steffanowski, 2011).

**Soziodemographische Daten der Klient\*innen.** In der vorläufigen Gesamtstichprobe der Klient\*innen kann der typische Mehranteil an weiblichen Personen festgestellt werden. Ähnliche Geschlechterverteilungen mit rund 70 % Frauenanteil weisen auch die Stichproben der QS-PSY-BAY-, TK- und PAP-S-Studien auf (Steffanowski et al., 2011; Wittmann & Steffanowski, 2011; von Wyl et al., 2013).

Das Durchschnittsalter in der vorliegenden Studie betrug mit 36 Jahren etwas weniger als jenes in den anderen Studien, das dort zwischen 39 und 40 Jahren lag. Dieser Umstand passt zu dem ermittelten Altersunterschied von Studienteilnehmer\*innen und -ablehner\*innen in vorliegender Studie.

Mit mehrheitlich verheiratete bzw. in Partnerschaft lebenden und über einen Lehrabschluss bzw. einer Fachschulausbildung verfügenden Personen entspricht vorliegende Gesamtstichprobe auch in diesen soziodemographischen Variablen in etwa den Verteilungen

in den angeführten Vergleichsstudien (Laireiter et al., 2013; Steffanowski et al., 2011; Wittmann et al., 2011; von Wyl et al., 2013).

Bei den zum angeführten naturalistischen Studien hebt sich die Stichprobe der TRANS-OP-Studie durch eine für Psychotherapeuten\*innen unübliche 50:50-Geschlechterverteilung und ein auffallend hohes Bildungsniveau (75 % Maturaabschluss) von den anderen berichteten Stichproben ab (Gallas et al., 2008). Dieser Unterschied kann durch den Erhebungskontext erklärt werden. Im Gegensatz zu den anderen Vergleichsstudien wurden in der TRANS-OP-Studie ausschließlich privat versicherte Psychotherapeuten\*innen zur Studie eingeladen (Altmann et al., 2016).

Zwar weist auch vorliegende Studie bei Psychotherapeuten\*innen in Ausbildung unter Supervision einen anderen Kontext als die Vergleichsstudien auf, weil alle durchgeführten Therapien ausnahmslos von den Klient\*innen selbst finanziert werden mussten und von Psychotherapeuten\*innen i.A.u.S. durchgeführt wurden. Dennoch konnten in der Gesamtstichprobe vergleichbare Verteilungen in den Stichproben der angeführten Studien bei fertig ausgebildeten Psychotherapeuten\*innen in freier Praxis hinsichtlich der soziodemographischen Daten gefunden werden.

### ***Skalenmittelwerte und Anteil an Personen mit klinisch auffälliger Symptomatik***

Alle symptombezogenen Skalenmittelwerte der Gesamtstichprobe befanden sich zu Therapiebeginn deutlich im krankheitswertigen Bereich und lagen nur geringfügig über bzw. unter den Gruppenmittelwerten der zum Vergleich gebrachten Studienstichproben und einer Vergleichsstichprobe von 1263 stationären Psychotherapeuten\*innen im Manual der SCL-90 (Franke, 2014; Gallas et al., 2008; Wittmann et al., 2011). Auch der durchschnittliche Wert der externen Beurteilung zum Befinden der Klient\*innen (GAF) befand sich ebenso im klinisch auffälligen Bereich und war nahezu ident mit jenem der PAP-S-Studie (Crameri et al., 2014; Tschuschke et al., 2015; von Wyl et al., 2013). Der Anteil an Personen mit mindestens einem klinisch auffälligen Testwert in einem der symptombezogenen Verfahren war in vorliegender Studie mit 84 % leicht höher als in der PAP-S-Studie mit 80 % (Crameri et al., 2014; Tschuschke et al., 2015; von Wyl et al., 2013).

Für die ressourcenbezogene Ryff-Skala standen keine Referenzwerte zur Verfügung. Der ESK-Gruppenmittelwert lag in vorliegender Studie zu Therapiebeginn etwas über dem ESK-Mittelwert einer Stichprobe von stationären Suchtpatient\*innen, wobei höhere Werte mit besserem Befinden einhergehen (Längle et al., 2005).

**Diagnosen.** In den vorgestellten naturalistischen Studien bildeten F3- und F4-Diagnosen den größten Anteil bei den Hauptdiagnosen der Psychotherapeuten\*innen in

freier Praxis (vgl. Tabelle 20) (Cramer et al., 2014; Gallas et al., 2008; Strauss et al., 2015; Tschuschke et al., 2015; Wittmann & Steffanowski, 2011; Wyl et al., 2013). So auch in vorliegender Gesamtstichprobe.

Der Anteil an Personen mit mindestens einer komorbiden Störung betrug in der QS-BAY-Studie 59 % und 47 % in der Studie an den Ausbildungsambulanzen von Victor und Kolleg\*innen (2018), was deutlich über dem Wert in vorliegender Studie mit 27 % lag (Strauss et al., 2015; Victor et al., 2018). Personen mit Verdacht auf Cluster-B-Persönlichkeitsstörungen waren in vorliegender Studie doppelt so stark vertreten wie in der PAP-S-Studie (12 %) (von Wyl et al., 2013). Die Unterschiede könnten auf verschiedene Vorgangsweisen in der Diagnosestellung zurückgeführt werden.

### ***Prä-Post-Stichprobe***

Für die Untersuchung der praktischen Wirksamkeit existenzanalytischer Ausbildungstherapien wurde eine Prä-Post-Stichprobe nach dem *Available-Case-Prinzip* gezogen. Von den 79 vorhandenen Datensätzen der Prä-Post-Stichprobe wiesen 17 Fälle (22 %) ein vorzeitiges Therapieende auf. Dieser Anteil liegt nur knapp unter den Angaben der Autor\*innen der PAP-S-Studie, die von rund einem Viertel vorzeitige Therapieabbrüche in ihrem Prä-Post-Datensatz berichteten (Cramer et al., 2014).

**Therapiedauer.** Die Therapiedauer der existenzanalytischen Ausbildungstherapien lag mit einer durchschnittlichen Anzahl von 30 Einheiten ( $SD = 22$ ) unter dem Durchschnitt der Studien aus Deutschland und der Schweiz ( $M \sim 40$ ) (vgl. Tabelle 20). Eine Ausnahme bildet die QS-PSY-BAY-Studie mit einer durchschnittlichen Sitzungsanzahl von rund 26 Einheiten ( $SD = 22$ ) (Steffanowski et al., 2011; Tschuschke et al., 2015; Victor et al., 2018; Wittmann et al., 2011). Die TRANS-OP-Studie lässt Vergleiche mit den untersuchten Verfahren zu und berichtet von einem Median von 25 Sitzungen bei der Verhaltenstherapie und somit von einem ähnlichen Therapieumfang wie die existenzanalytischen Ausbildungstherapien in vorliegender Studie ( $Md = 23$ ) (Gallas et al., 2008).

**Tabelle 20***Naturalistische Vergleichsstudien*

Studie	Autor*innen	Land	Jahre	Therapie- verfahren	N <sub>PRÄ</sub>	N <sub>TH</sub>	Diagnose- gruppe (%)	Dauer in Einheiten M <sub>EH</sub> (SD) bzw. M <sub>Mon</sub> (SD)	Ergebnisse
TRANS-OP	Gallas et al. (2008, 2010)	D	1998- 2000	TP/AP/VT	714	471	F3 (47.5) F4 (42.3)	TP: $Md_{EH} = 42$ AP: $Md_{EH} = 101$ VT: $Md_{EH} = 25$	Response <sub>CC</sub> : 62 % aller Teilnehmer*innen Remission: 36 % aller Personen mit klinisch auffälliger Symptomatik ( $n = 542$ )
TK- Modellprojekt	Wittmann & Steffanowski (2011)	D	2005- 2010	TP/AP/VT	1598	403	F3 (44.4) F4 (46.9)	$M_{EH} = 39.12$ (24.10) $M_{Mon} = 25$ (8.7)	Gemittelt Prä-Post <sub>CC/AC</sub> $d = 1.20$ ( $n = 594$ ) 1-Jahres-Katamnese $d = 1.05$ ( $n = 463$ )
QS-PSY-BAY	Steffanowski et al. (2011) Strauss et al. (2015)	D	2007- 2009	TP/AP/VT	1696	217	F3 (48.7) F4 (39.6)	$M_{EH} = 26.4$ (21.8) $M_{Mon} = 12.7$ (5.9)	Prä-Post <sub>ITT</sub> $d = 0.74$ ( $n = 1696$ ) Prä-Post <sub>C</sub> $d = 1.10$ ( $n = 309$ ) 1-Jahres-Katamnese $d = 1.10$ ( $n = 309$ )
PAP-S	Crameri et al. (2014) Tschuschke et al. (2015)	CH	2007- 2012	10 Verfahren (H & TP)	362	81	F3 (39.0) F4 (40,4)	$M_{EH} = 43.2$ (35.9)	Prä-Post <sub>CC/AC</sub> $d = 0,89$ ( $n = 264$ ) 1-Jahres-Katamnese $d = 1.05$ ( $n = 146$ )
Ausbildungs- Ambulanzen	Victor et al. (2018)	D	2010- 2014	VT	719	260	F3 (54.0) F4 (35.7)	$M_{EH} = 43.3$ (19.7)	Prä-Post <sub>C</sub> $d = 1.63$ Response = 80 % Remission = 64 % ( $n = 401$ )
GLE	Längle et al. (2005)	A	1993- 2004	EA	248	12	F3 (27 %) F4 (24 %)	$M_{EH} = 28$ (-)	$d_{AC} = 0.99$
GLE	Laireiter et al. (2013)	A	1996- 1998	EA	108	26	F3 (27.4) F4 (40.6)	$M_{EH} = 48.1$ (57.1)	$d_c > 1.00$

*Anmerkungen.* TP = Tiefenpsychologisch fundierte Psychotherapie; AP = Analytische Psychotherapie; VT = Verhaltenstherapie; H = Humanistische Psychotherapie; EA = Existenzanalyse Md = Median; M = Mittelwert; SD = Standardabweichung; EH = Einheiten; Mon = Monate;  $d$  = durchschnittliche Effektstärken nach Cohen; CC = Complete-Case; AC = Available-Case; C = Completer; ITT = Intention to Treat; (-) = keine Angabe.

### ***Gruppenstatistische Veränderung zwischen Therapiebeginn und Therapieende (T1-T3)***

Alle untersuchten Prä-Post-Mittelwertsdifferenzen waren auch nach der Bonferroni-Korrektur statistisch signifikant und wiesen hohe Effekte auf, was für eine deutliche Verbesserung des Befindens im Zeitraum der Therapie spricht. Beim Vergleich der gemittelten Effektstärken zeigt sich, dass es mit einem Cohens'  $d_{Prä-Post}$  um bzw. teilweise deutlich über 1.00 zu ähnlichen Therapieeffekten wie in den dargestellten Studien zu Therapieende kam (vgl. Tabelle 20) (Altmann et al., 2016).

**Symptombezogene Veränderungen.** Die höchste Effektstärke konnte auf der Depressionsskala der SCL ermittelt werden. Ein Einfluss durch auffallend hohe Gruppenmittelwerte zu Therapiebeginn (Deckeneffekte), die zu einem stärkeren Effektmaß führen könnten, konnte aber durch den Vergleich mit einer Stichprobe von stationären Psychotherapieklient\*innen im Manual der SCL-90 entkräftet werden (Franke, 2014). *Deckeneffekte* treten dann auf, wenn sich der Gruppenmittelwert einer Stichprobe im hohen bzw. klinisch stark auffälligen Skalenbereich befindet, wodurch die Wahrscheinlichkeit steigt, dass sich unabhängig von den Interventionseffekten automatisch höhere Effekte ergeben (Andrade, 2021). Jedenfalls wies die Prä-Post-Stichprobe zu Therapieende nur geringfügig höhere Mittelwerte auf der Depressionsskala der SCL als die Normstichprobe im Manual der SCL-90 auf und befand sich somit zu Therapieende im klinisch unauffälligen Bereich. Auch auf der Angst-Skala und der GSI-Gesamtskala näherten sich die Gruppenmittelwerte zu Therapieende den Zahlen der Normstichprobe ( $N = 2205$ ) im Manuale an (Franke, 2014). Die Effektstärken lagen ein wenig unter jener der Depressionsskala, jedoch im großen Bereich mit Werten von Cohen's  $d > 0.8$  (Cohen, 1988).

Auch die Ergebnisse des IIP zeigen einen bedeutsamen Rückgang von interpersonalen Problemen im Zeitraum der Therapie. Die Effektstärke fiel aber im Vergleich zu den Skalen der SCL etwas geringer aus. Dem könnten zwei Ursachen zugrunde liegen: Einerseits wäre denkbar, dass interpersonale Schwierigkeiten veränderungsresistenter als andere psychische Probleme sind oder diese Veränderung erst zu einem späteren Zeitpunkt des Therapieprozesses auftritt (Wittmann et al., 2011). Andererseits kann aber auch das eingesetzte Instrument weniger veränderungssensitiv sein, was die Höhe der zu generierenden Effekte beeinflussen könnte. Die Mittelwerte zu Therapiebeginn lagen zudem weniger deutlich im klinisch auffälligen Bereich und bereits näher am Schwellenwert zum klinischen unauffälligen Bereich, was auf *Bodeneffekte* hindeuten könnte, die ebenfalls zu geringeren Effekten beitragen könnten. Bodeneffekte treten dann auf, wenn sich der Gruppenmittelwert einer Stichprobe im niedrigen bzw. unauffälligen Wertebereich befindet, wodurch kaum

Effekte durch jedwede Intervention zu erwarten sind (Andrade, 2021; Cámara, Begré & Känel, 2010). Auch in der TK-Studie waren die Prä-Post-Effektstärken der IIP-Gesamtskala geringer als jene der anderen Instrumente (Wittmann et al., 2011).

**Ressourcenbezogene Veränderung und externe Beurteilung.** Auf der Existenzskala kam es im Vergleich zu den anderen Skalen, ähnlich wie beim IIP, zu geringeren Effektstärken, die dennoch als groß zu interpretieren sind (Cohen, 1988). Dies scheint im ersten Moment paradox, denn das verfahrenseigene Erhebungsinstrument sollte die Veränderung durch existenzanalytische Therapie doch besonders sensitiv abbilden. Jedoch ist wie beim Inventar zur Erfassung interpersonaler Probleme (IIP) auch hier denkbar, dass die Fragen stärker auf längerfristige Einstellungen ausgerichtet sind, statt auf kurzfristig erfasstes Befinden. Die Prä-Post-Stichprobe wies zu Therapieende exakt den gleichen Gruppenmittelwert auf wie die Stichprobe der Praxisstudie von Längle et al. (2005) zu Therapieende (Längle et al., 2005). Dieser Wert liegt ein wenig über dem Mittelwert der Normstichprobe und spricht für ein gutes Behandlungsergebnis (Längle, Orgler & Kundi, 2000).

Im Gegensatz zur IIP- und ESK-Gesamtskala wies die Ryff-Skala eine ähnlich hohe Prä-Post-Effektstärke wie die Depressionsskala auf. Die Klient\*innen berichteten somit von einer Veränderung mit starken Effekten hinsichtlich ihres psychologischen Wohlbefindens. Die markante Veränderung könnte ein Hinweis darauf sein, dass bei humanistischer Psychotherapie der Fokus stärker auf Ressourcenorientierung denn auf Symptomreduktion gelegt wird (Angus et al., 2015).

Die externe Einschätzung der Therapeut\*innen zum globalen Funktionsniveau ihrer Klient\*innen veränderte sich ebenso statistisch signifikant mit großem Effekt. Die Stärke der Veränderung ist mit jener der PAP-S-Studie zu vergleichen, in der die GAF-Skala von externen Beurteiler\*innen ausgefüllt wurde (Tschuschke et al., 2015). Aus der Literatur ist bekannt, dass die Skala eine hohe Veränderungssensitivität für Behandlungseffekte aufweist (Heuft & Lange, 2015).

### ***Individuelle Veränderung zwischen Therapiebeginn und Therapieende (T1-T3)***

Durch die Ermittlung des *Reliable Change Index* für die GSI-Skala konnte bei knapp 91 % der 53 klinisch auffälligen Personen zu Therapiebeginn eine reliable Verbesserung festgestellt werden. Zusätzlich veränderten sich knapp über drei Viertel der zu Therapiebeginn klinisch auffälligen Personen hinsichtlich ihrer globalen Gesamtbelastung klinisch signifikant. Diese Personen wiesen zu Therapieende keine klinisch relevante Symptomatik mehr auf. Im Vergleich mit den berichteten RCI-Werten der TRANS-OP-Studie

und der Untersuchung an Ausbildungsambulanzen von Victor und Kolleg\*innen (2018) war der Anteil an Personen mit bedeutsamen Veränderungen in der globalen Gesamtbelastung in vorliegender Studie größer (vgl. Tabelle 20) (Gallas et al., 2008; Victor et al., 2018). Die hohen RCI-Veränderungswerte dieser Studie sind vor dem Hintergrund der verwendeten Available-Case-Stichprobe beachtenswert, da auch frühzeitig beendete Therapieverläufe (22 %) in die Analyse inkludiert wurden.

Der Anteil an Klient\*innen mit (klinisch) signifikant gebessertem Befinden war auf den drei anderen symptombezogenen Skalen kleiner als jener des GSI und vergleichbar mit den Ergebnissen der beiden anderen Studien (TRANS-OP und Ausbildungsambulanzen) (Gallas et al., 2008; Victor et al., 2018). Interessant ist, dass trotz der hohen Prä-Post-Effektstärke der Depressionsskala die RCI-Ergebnisse dieser Skala in ihrer Stärke hinter der GSI-Skala an zweiter Stelle liegen. Die meisten Personen haben sich auf der GSI-Gesamtskala - und nicht auf der Depressionsskala der SCL - reliabel gebessert. Dieses Ergebnis relativiert das Gruppenergebnis. Denkbar wäre, dass sich einige Personen in der Prä-Post-Stichprobe in ihrer Depressionsausprägung besonders stark verändert haben, was wiederum zu der größeren Effektstärke der Gruppe insgesamt geführt haben könnte (Jensen & Corralejo, 2017; Senft et al., 2020). Das gruppenstatistische Ergebnis kann durch die Ermittlung der Veränderung bei den einzelnen Klient\*innen differenziert betrachtet werden.

Mit der gesonderten Berechnung der RCI-Werte für Personen, die zu Therapiebeginn klinisch unauffällig waren, sollten mögliche Verschlechterungen ausgeschlossen werden. Lediglich auf der Angst-Skala der SCL und der IIP-Gesamtskala hat sich das Befinden jeweils einer Person (reliabel) verschlechtert, jedoch wiesen diese Fälle auch zu Therapieende einen klinisch unauffälligen Status auf.

### ***Therapieurteilungen der Klient\*innen und Therapeut\*innen***

Beinahe alle Personen der Prä-Post-Stichprobe, nämlich 98 %, waren zu Therapieende mit dem Behandlungsergebnis *zufrieden* oder *sehr zufrieden*. In der TK-Studie gaben etwas weniger nämlich 77 % der Klient\*innen die beste bzw. zweitbeste Bewertung zur Behandlungszufriedenheit ab. 90 % der Klient\*innen schätzten den Erfolg der Psychotherapie als (*sehr*) *hoch* ein. Ein weniger deutliches Bild zeigte sich in der Studie von Laireiter und Kolleg\*innen (2000), in der knapp über der Hälfte der befragten Klient\*innen die existenzanalytische Psychotherapie als (*sehr*) *erfolgreich* eingeschätzt hatten. Diese Studie verwendete ein retrospektives Studiendesign, bei dem davon auszugehen ist, dass mehr Zeit zwischen dem Ausfüllen des Fragebogens und dem Therapieende lag. Eine länger zurückliegende Therapie wird eventuell anders beurteilt als eine aktuelle. Ebenso ist denkbar,

dass in vorliegender Studie durch Abgabe des ausgefüllten Fragebogens an die Therapeut\*innen eher sozial erwünschte Antworten gegeben wurden und damit deutlich höhere Zufriedenheitswerte zustande kamen. Genauso kann nicht ausgeschlossen werden, dass Daten zu Therapieende vorwiegend von Personen vorlagen, die mit der Therapie zufriedener gewesen waren, was dem Selbstselektions-Bias entsprechen würde.

Insgesamt zeigt sich die aus der Literatur bekannte typische schiefe Verteilung bei den Zufriedenheitswerten von Klient\*innen auch in dieser Studie. Die meisten Klient\*innen den therapeutischen Prozess grundsätzlich eher positiv bewerten mit Ausnahme von deutlich unzufriedenen Klient\*innen (Laireiter et al., 2000).

Gute Zufriedenheitswerte von Klient\*innen sind zwar kein Garant für einen ausreichenden Therapieeffekt, so hängt beispielsweise die subjektive Therapiebeurteilung der Klient\*innen häufig mit der therapeutischen Beziehungsqualität zusammen und in geringerem Ausmaß mit dem tatsächlichen Therapieergebnis. Dennoch können positive Zufriedenheitswerte als bedeutsame Voraussetzung für erfolgreiche Therapieprozesse angesehen werden (Hiller et al., 2009; Laireiter et al., 2000). Auch in der PAP-S-Studie konnte ein korrelativer Zusammenhang zwischen Behandlungszufriedenheit und Therapieergebnis gefunden werden (Tschuschke et al., 2015).

### ***Kurz- und langfristige Veränderungen***

**Verlauf-Stichprobe.** Zur Verlaufsmessung (25 Einheiten nach Therapiebeginn) weisen eher jene Personen Daten auf, die länger in Therapie waren und zu Therapiebeginn tendenziell stärker belastet waren. Das verdeutlicht auch der signifikante Unterschied zu Therapiebeginn zwischen der Verlauf-Stichprobe mit der restlichen Gesamtstichprobe mit kleinem Effekt auf der GAF-Skala. Die ermittelten Effektstärken der auch nach der Bonferroni-Korrektur signifikanten Mittelwertsunterschiede zwischen Therapiebeginn und der Verlaufsmessung befinden sich alle (noch) nicht im großen Bereich (vgl. Abbildung 8) (Cohen, 1988). Wie zu Therapieende weisen die Veränderung der depressiven Symptomatik auch kurzfristig die größten Effekte auf und die Gesamtskalen des IIP und der ESK sind erneut durch geringere Effekte geprägt. Interessant ist, dass im Gegensatz zu Therapieende bei der Verlaufsmessung das psychologische Wohlbefinden, gemessen durch die Ryff-Skala, einen eher geringeren Effekt aufweist (vgl. Abbildung 8). Das könnte darauf hindeuten, dass sich durch die aktuelle Bearbeitung der Problematiken bereits Einiges an der wahrgenommenen Symptomausprägung verändert hatte, aber das erlebte psychologische Wohlbefinden sich erst nach beendeter Therapie bzw. zu einem späteren Zeitpunkt verändert. Die Ergebnisse und Effektgrößen der Zwischenmessung verdeutlichen, dass Therapien bei

gegebener Indikation ausreichend lange dauern sollten, bis sich ein zufriedenstellender und möglichst großer Effekt einstellt.

**Katamnese-Stichprobe.** Die Ergebnisse der Katamnese-Stichprobe geben Hinweise auf einen nachhaltigen Therapieeffekt. In allen Erhebungsverfahren konnten auch nach der Bonferroni-Korrektur statistisch signifikante Veränderungen gegenüber Therapiebeginn festgestellt werden. Diese Resultate müssen jedoch vor dem Hintergrund einer deutlich selektiven Stichprobe interpretiert werden. Die Klient\*innen, die bereit waren, die Katamnesefragebögen auszufüllen, wiesen bessere Werte zu Therapiebeginn auf der ESK-Gesamtskala und zu Therapieende in allen Verfahren - bis auf die der IIP-Gesamtskala - auf.

Bei den Klient\*innen der Katamnese-Messung konnten sechs Monate nach Therapieende insbesondere in der Ryff-Skala sehr hohe Effektstärken festgestellt werden. Interessant ist, dass die IIP-Gesamtskala hier die zweithöchste Effektstärke aller Verfahren aufweist. Dieses Ergebnis ist insofern erstaunlich, als dass sich die Effekte des IIP zuvor immer im unteren Bereich befanden. Die Werte der IIP-Gesamtskala der Katamnese-Stichprobe unterschieden sich weder zu Therapiebeginn noch zu Therapieende von jenen der restlichen Gesamtstichprobe, weswegen dieser Befund geringer von einer Verzerrung betroffen sein dürfte. Es könnte ein Hinweis darauf sein, dass erlebte interpersonale Probleme (erst) längere Zeit nach einer Psychotherapie deutlich abnehmen.

Verglichen mit den Effektstärken zu Therapieende vergrößerten sich die Therapieeffekte zum Katamnesezeitpunkt, was auf einen *nachhaltigen Therapieeffekt* existenzanalytischer Behandlung hindeutet (Laireiter et al., 2013; Strauss et al., 2015). Insgesamt entspricht die Vergrößerung der Effekte mit jedem weiteren Messzeitpunkt den Ergebnissen anderer Psychotherapiestudien (vgl. Abbildung 8) (Längle et al., 2000; Wittmann et al., 2011).

### ***Explorative Analysen***

**Complete- vs. ITT-Analysen.** Es wurde eine *ITT*-Stichprobe erstellt, bei der alle Studienteilnehmer\*innen mit Daten zu Therapiebeginn eingeschlossen wurden. Für die fehlenden Datensätze zu Therapieende wurde nach Empfehlung von Hiller und Kolleg\*innen (2011) als Imputationstechnik die *LOCF*-Methode (*last-observation-carried-forward*) verwendet, bei der der zuletzt vorhandene Messzeitpunkt eines Falles für den fehlenden Messzeitpunkt zu Therapieende eingesetzt wird (Hiller et al., 2011). Die *LOCF*-Methode gilt als konservative Schätzung von Therapieeffekten, da Personen mit Testdaten ausschließlich zu Therapiebeginn keinerlei Veränderungen aufweisen, und dennoch in die Analyse miteinfließen. Somit sind die korrigierten Effektstärken von der Anzahl vorhandener

Datensätze zu den jeweiligen Testzeitpunkten abhängig und sollten nur vor diesem Hintergrund interpretiert werden. Die Autor\*innen der QS-PSY-BAY-Studie gaben bei der ITT-Analyse an, wie hoch der Anteil an Datensätzen zu den jeweiligen Testzeitpunkten war. So flossen in deren ITT-Analyse 15 % an Fällen ein, die ausschließlich Daten zum ersten Messzeitpunkt enthielten und deren Therapieeffekt somit bei  $d = 0.00$  lag (Steffanowski et al., 2011; Strauss et al., 2015). In vorliegender Studie lag der Anteil an Personen mit Daten ausschließlich zu Therapiebeginn bei 42 % und somit deutlich über dem berichteten Anteil der QS-PSY-BAY-Studie. Der ITT-Stichprobe wurde, wie in der Literatur empfohlen, eine *Completer*-Stichprobe gegenübergestellt, die nur Fälle mit regulär beendeten Therapien umfasste. Die ermittelte Effektstärke der Completer-Stichprobe lag für den GSI mit  $d_C = 1.09$  im großen Bereich und jene der ITT-Analyse mit  $d_{ITT} = 0.43$  im kleinen Bereich. Grundsätzlich ist davon auszugehen, dass der wahre Therapieeffekt irgendwo zwischen diesen beiden Werten liegt (Hiller et al., 2011; Strauss et al., 2015).

**Prädiktoren des Therapieergebnisses.** Eine explorative multiple lineare Regression sollte den Einfluss möglicher Prädiktoren (die therapeutische Beziehung, die Anzahl der Einheiten, die Therapeut\*innen, die Eingangssymptomatik) auf das Therapieergebnis ( $GSI_{TE-TB}$ ) untersuchen. Nahezu 60 % der Ergebnisvarianz konnte vor allem mit der Ausgangsbelastung (Ausprägung des GSI zu Therapiebeginn) und etwas weniger mit der therapeutischen Beziehung erklärt werden. Je stärker die Personen zu Therapiebeginn belastet waren, desto mehr verbesserten sie sich im Verlauf der Therapie. Dieser Befund wird durch die Ergebnisse der PAP-S-Studie gestützt (Tschuschke et al., 2015). Dabei könnte das Phänomen der *Regression zur Mitte* eine Rolle spielen, welches besagt, dass Personen mit hohen Testwerten (zu Therapiebeginn) bei der nächsten Messung mit einer viel höheren Wahrscheinlichkeit einen Testwert näher dem Durchschnittswert aufweisen (Victor et al., 2018).

Ebenso hatte die Ausprägung der therapeutischen Beziehung einen Einfluss auf den Behandlungserfolg. Mit steigender Beziehungszufriedenheit vergrößerte sich auch das Therapieergebnis. Während dieser häufig postulierte Einflussfaktor in der TK-Studie nicht bestätigt werden konnte, sprechen die Befunde der PAP-S-Studie für einen positiven Einfluss der therapeutischen Beziehung auf den Behandlungsausgang (Horvath et al., 2011; Tschuschke et al., 2015; Wittmann et al., 2011). Die therapeutische Beziehung wurde auch von einigen Klient\*innen in den offenen Eingabefeldern der Fragebögen als bedeutsam für den Therapieprozess beschrieben. Dabei wurde besonders die *Empathiefähigkeit* der

Therapeut\*innen hervorgehoben. Durch einen dialogischen Prozess auf Augenhöhe könne ein *Verstehen* durch ein *Verstanden-werden* einsetzen (Längle, 2014; Launeanu et al., 2019).

Weder in der PAP-S-Studie noch in vorliegender Studie konnte eine signifikante Ergebnisvarianzaufklärung durch die Sitzungsanzahl oder die einzelnen Therapeut\*innen festgestellt werden (Tschuschke et al., 2015).

**Korrelativer Zusammenhang der ESK mit psychischem Befinden.** Eine explorative Korrelationsanalyse wurde durchgeführt, um den Zusammenhang der ESK-Gesamtskala mit den eingesetzten symptom- und ressourcenbezogenen Skalen zu untersuchen. Die signifikanten und teils bemerkenswert starken Zusammenhänge der ESK-Gesamtskala mit den Skalen der eingesetzten, psychometrischen Instrumente geben Anlass zu der Annahme, dass die innere Erfüllung eine erhebliche Bedeutsamkeit für das psychische Befinden aufweist. Dabei zeigte sich, dass höheres existenzielles Erleben mit niedrigeren Werten auf den Angst- und Depressionsskalen sowie auf der Globalskala der SCL-90 und mit geringeren interpersonalen Problemen einhergeht. Steigende Werte auf der ESK-Skala hängen zudem mit höherem psychologischen Wohlbefindenserleben zusammen, was die Nähe der erhobenen Konzepte stützt (Längle, Orgler & Kundi, 2000; Ryff & Keyes, 1995). Diese Ergebnisse sprechen insgesamt für die Bedeutung von existenziellen Konzepten im psychotherapeutischen Kontext.

## **Studienübergreifende Diskussion**

### ***Verschiedene Veränderungsmaße als Therapiebeurteilung***

Der komplexe Prozess einer psychotherapeutischen Behandlung sollte möglichst multidimensional erhoben, analysiert und diskutiert werden. Dabei sollten die Vor- und Nachteile unterschiedlicher Analysemethoden bei der Ergebnisinterpretation mitberücksichtigt und diskutiert werden (Wittmann et al., 2011).

**Effektstärkenmaß.** Mit Effektstärken kann über unterschiedliche Instrumente, Verfahren und Studien hinweg eine vergleichende Einschätzung der Therapieergebnisse angestellt werden (Jacobi, 2020). Bei der Interpretation von Effektstärken sollten jedoch einige Einflussfaktoren mitberücksichtigt werden.

Den zum Teil sehr unterschiedlichen Effektstärken der einzelnen Verfahren innerhalb einer Studie und während eines Testzeitpunktes können verschiedene Ursachen zugrunde liegen. Einerseits wäre denkbar, dass es sich dabei tatsächlich um eine *differenzielle Wirkung der Therapieform* zu verschiedenen Testzeitpunkten und bei verschiedenen Symptomen handelt. So könnten die hohen Effektstärken auf der Depressionsskala der SCL in dieser Studie darauf hindeuten, dass die existenzanalytische Psychotherapie bei affektiven Störungen

wirksamer ist als beispielsweise bei Angstsymptomen. Ebenso könnte die merkliche Zunahme der IIP-Effektstärken zum Katamnesezeitpunkt ein Anzeichen dafür sein, dass interpersonale Probleme eines längeren Wirkzeitraums bedürfen.

Andererseits könnte aber auch die variierende *Veränderungssensitivität der einzelnen Instrumente* einen Einfluss auf die Ausprägungen von Effektstärken haben. Auch in vorliegender Studie könnten die geringeren Effektstärken der ESK-Gesamtskala über die einzelnen Testzeitpunkte hinweg ein Indiz dafür sein.

Beim Vergleich von Effektstärken über Studiendesigns hinweg sind weitere Einflussfaktoren in Betracht zu ziehen. Bei der Berechnung von Effektstärken gilt: je kleiner die Variabilität der Stichprobe, also je homogener die Klient\*innengruppe zu Therapiebeginn ist, desto kleiner ist die Standardabweichung der relevanten Merkmale und somit auch die Zahl im Nenner der Formel für Effektstärken. Eine kleinere Zahl im Nenner hat höhere Werte zur Folge, was gleichbedeutend mit größeren Effektstärken ist (Jacobi, 2020). Auch die Stichprobengröße kann einen Einfluss auf die Höhe der Effektstärken haben. Bei steigender Stichprobengröße wird die Streuung der Werte der Untersuchungsgruppe tendenziell kleiner (Bhude, Shah & Acharya, 2018; Jacobi, 2020; Senft et al., 2020).

Die Empfehlung, nur klinisch auffällige Personen in die Analysen einzuschließen, kann zu einer weiteren Beeinflussung der Effektstärken führen. Bei Personen mit klinisch auffälligen Symptomausprägungen kommt das Phänomen der *Regression zur Mitte* stärker zum Tragen (Hiller et al., 2011; Victor et al., 2018). Zudem bildet eine Stichprobe bestehend aus ausschließlich klinisch auffälligen Personen nicht die Versorgungsrealität von Psychotherapie in freier Praxis ab und bei Ausschluss nicht pathologischer Fälle gehen auch Informationen zu diesen Personen verloren (Bohart et al., 1998). In einer Studie von Hiller (2009) betraf ein Ausschluss von Personen mit klinisch unauffälliger Symptomatik von der Analyse immerhin 30 % der Eingangsstichprobe und in der Studie von Victor et al. (2018) rund 17 %. Ohne Informationen zu diesen Personen kann beispielsweise eine Verschlechterung des Befindens im Zeitraum der Therapie nicht ausgeschlossen werden.

**Maß der reliablen Veränderung und klinische Signifikanz.** Während Effektstärken die Größe eines empirischen Effektes einer Gruppe angeben, lassen sie keine Aussage darüber zu, ob die Stärke der Veränderungen auf große Unterschiede innerhalb weniger Personen zurückzuführen ist und weniger bedeutsame Veränderungen bei anderen Personen dieser Gruppe überlagert wird (Jacobi, 2020; Jensen & Corralejo, 2017). Zusätzlich zu ermitteln, bei wie vielen Personen sich das Befinden reliabel gebessert oder verschlechtert bzw. klinisch signifikant gebessert hat erhöht daher die Aussagekraft von Studienergebnissen (Jensen &

Corralejo, 2017). In der Literatur haben sich dafür der *Reliable Change Index (RCI)* und die *klinische Signifikanz* durchgesetzt. Obwohl eine Verlaufsbeurteilung einzelner Psychotherapeuten\*innen durch eine Zuteilung in die Gruppen *Response* (reliabel gebessert) und *Remission* (klinisch signifikant gebessert) sinnvoll erscheint, wird dieses Maß (bislang) deutlich weniger als das Effektstärkenmaß zur Therapiebeurteilung eingesetzt (Hiller, Schindler, & Lambert, 2011).

Mit dem RCI-Maß ist es möglich, Therapieeffekte differenzierter zu untersuchen. So kann beispielsweise Verschlechterung deutlicher abgebildet werden und schlägt sich nicht nur in geringeren Effektstärken einer Gruppe, sondern in der genauen Zahl an Personen mit verschlechtertem Befinden nieder. Weiterführende Untersuchungen mit dieser Gruppe hinsichtlich möglicher therapierelevanter Einflussfaktoren auf unerwünschte Ergebnisse sind denkbar. Befunde, dass Personen mit einem klinisch signifikanten Veränderungsstatus eine höhere Therapiezufriedenheit aufweisen als jene ohne klinisch signifikante Verbesserung, sprechen zudem für die Validität des RCI-Maßes (Senft et al., 2020).

Neben der gruppenstatistischen und individuellen Veränderungsbeurteilung, welche beide zu den *indirekten* Veränderungsmaßen gezählt werden, kann die Erhebung der Klient\*innenzufriedenheit die statistischen Befunde um eine *direkte* Veränderungseinschätzung zusätzlich erweitern (Hiller et al., 2011; Laireiter et al., 2000).

Durch den Einsatz verschiedener Erhebungsinstrumente (störungsspezifisch / transdiagnostisch, ressourcen- / symptombezogen, direkt / indirekt) und Analyseebenen (gruppenstatistisch / individuell) kann in der Zusammenschau eine sich gegenseitig ergänzende und zuverlässige Beurteilung der Ergebnisqualität von Psychotherapie erfolgen (Wittmann et al., 2011).

### ***Stärken und Limitationen***

Die Befunde vorliegender Studie sollten vor dem Hintergrund folgender Stärken und Schwächen des Studiendesigns betrachtet werden:

Bei der Analyse der Daten stellte sich der anhaltende, prozesshafte Charakter der Studie als erschwerend dar. Insbesondere die Rücklaufs- und die Abbruchsanalyse waren dadurch beeinflusst. Auch durch das Fehlen einer systematischen Befragung (zumindest) der Therapeut\*innen bei Abbruchfällen entfielen grundlegende Informationen zu den Gründen und Zeitpunkten der Abbrüche. In Zukunft sollten Therapeut\*innen gebeten werden, bei jedem Abbruchfall (unabhängig vom Messzeitpunkt) einen kurzen Abbruchsfragebogen auszufüllen. Zudem sollten bei absehbar vorzeitigem Therapieende die Klient\*innen ebenfalls ersucht werden, einen Abschlussfragebogen auszufüllen. Das hätte den Vorteil, dass neben

den zusätzlichen Informationen für die Abbruchanalyse die Personen mit einem Therapieabbruch als Kontrollgruppe fungieren könnten (Laireiter et al., 2013). Mittlerweile gilt zwar ein Ein-Gruppen-Studiendesign zur Ermittlung der praktischen Wirksamkeit eines etablierten Psychotherapieverfahrens als adäquat, dennoch würde eine Kontroll- bzw. Vergleichsgruppe den Ergebnissen mehr Nachdruck verleihen (Lambert & Ogles, 2009).

Die erzielte Stichprobengröße ist zwar im Vergleich zu jenen der vorgestellten naturalistischen Studien als klein zu beurteilen, doch sollte in diesem Fall mitberücksichtigt werden, dass es sich bei vorliegender Studie nicht um ein Kooperationsprojekt mehrerer Institutionen handelt. Die praxisorientierten Vergleichsstudien wurden durch Kooperationen von (Sozial-) Versicherungen, universitären Forschungsinstituten, ambulanten Einrichtungen usw. realisiert. Trotz dieser Ressourcenaggregation kam es bei manchen dieser Studien zu anteilmäßig vergleichbaren Rücklauf- und Abbruchquoten wie in vorliegender Studie (Strauss et al., 2015; Wittmann et al., 2011; Wittmann & Steffanowski, 2011).

Ein Problem, mit dem nicht nur naturalistische, sondern auch randomisiert-kontrollierte Studiendurchführende konfrontiert sind, ist der Bias durch Selbstselektion (*self-selection-bias*). Dieses Phänomen besagt, dass teilnahmewillige Therapeut\*innen und Klient\*innen sich systematisch von Nicht-Studienteilnehmer\*innen unterscheiden können. Die Gefahr einer Ergebnisverzerrung durch Selbstselektivität vergrößert sich mit fortschreitender Studiendauer und kann somit insbesondere bei Therapieabschluss- und Katamneseergebnissen relevant sein (Laireiter et al., 2013; Steffanowski et al., 2011; Strauss et al., 2015; Wittmann & Steffanowski, 2011). Um dieses Problem zumindest kontrolliert zu halten, hilft der Vergleich mit den Stichproben anderer Studien. Zudem sollten gezogene Teilstichproben mit der restlichen Gesamtstichprobe auf statistisch signifikante Unterschiede hin überprüft werden. In vorliegender Studie konnte trotz statistisch signifikanter Unterschiede im Durchschnittsalter der Studienteilnehmer\*innen und -ablehner\*innen eine zu anderen praxisorientierten Psychotherapiestudien vergleichbare Zusammensetzung der Gesamtstichprobe festgestellt werden. Bei den einzelnen Teilstichproben unterschieden sich besonders die Personen der Katamnese-Stichprobe von der restlichen Stichprobe.

Bei naturalistischen Psychotherapiestudien wird häufig der fehlende Einsatz von Behandlungsmanualen kritisiert. Das hat die Konsequenz, dass nicht genau festgestellt werden kann, welche Form der Therapie bzw. welche konkreten Interventionen zu den Therapieeffekten geführt haben. In den vorgestellten Forschungsprojekten wurde diese Thematik nur von der PAP-S-Studie aufgegriffen (von Wyl et al., 2013). Dafür wurden Aufzeichnungen der durchgeführten Einheiten stichprobenartig analysiert, um festzustellen,

welche Interventionen angewandt wurden und in welchem Ausmaß diese dem erlernten Psychotherapieverfahren entsprachen (*adherence*). Es hat sich gezeigt, dass der eindeutig größte Anteil der eingesetzten Methoden in allen zehn untersuchten Therapieverfahren auf unspezifische bzw. allgemeine Interventionen entfiel. Einen geringeren Anteil nahmen Interventionen ein, welche anderen Verfahren zugeordnet werden konnten, und interessanterweise wurde in den untersuchten Behandlungen am wenigsten Interventionen des eigenen Verfahrens eingesetzt. Zudem konnte kein Einfluss der Behandlungsspezifität bzw. der Methodentreue auf das Therapieergebnis gefunden werden (Tschuschke et al., 2015). Aus vorangegangenen Untersuchungen geht hervor, dass, je mehr Berufserfahrung die Therapeut\*innen aufweisen, desto weniger wird nach erlerntem Verfahren bzw. konkretem Manual gearbeitet bzw. umgekehrt, Berufsnovizen halten sich stärker an das erlernte Verfahren (Wampold, 2021). Das kann wiederum durch die Ergebnisse der PAP-S-Studie gestützt werden, aus welcher eine signifikant stärkere Konzepttreue bei berufsneuen Therapeut\*innen hervorging (Tschuschke et al., 2015). Für vorliegende Studie kann angenommen werden, dass sich die in Ausbildung und unter Supervision befindlichen Psychotherapeut\*innen (noch) stärker an existenzanalytischen Behandlungskonzepten orientieren. Dennoch wäre empfehlenswert, in zukünftigen Studien die Therapeut\*innen im Abschlussfragebogen die eingesetzten Methoden beschreiben zu lassen, um gegebenenfalls Interventionen mit den Therapieergebnissen in Verbindung bringen zu können.

Um den Datenschutz und die Anonymität der Klient\*innen gegenüber Dritten zu gewährleisten, erfolgte eine Studienbeteiligung der Klient\*innen ausschließlich mittels ihrer Therapeut\*innen. Die ausgefüllten Fragebögen gehen bei diesem Prozedere wieder zurück an die Therapeut\*innen und sind für diese einsehbar. Bei diesem Vorgehen kann eine Verzerrung durch sozial erwünschtes Antwortverhalten, insbesondere bei der Erhebung der Arbeitsallianz, nicht ausgeschlossen werden (Wilmers et al., 2008).

Hinsichtlich der zeitlichen Komponente sollte auch die, für naturalistische Studien zwar übliche, lange Studiendauer diskutiert werden. Die Zeitpunkte des Therapiebeginns der einzelnen Fälle innerhalb der Stichprobe können Jahre auseinanderliegen. Einflussgrößen, wie gesellschaftliche Krisen am Beispiel der Covid-19-Pandemie, sollten in den Analysen zumindest mitgedacht werden und fanden auch in den offenen Eingabefeldern der Therapeut\*innen- und Klient\*innen-Fragebögen Erwähnung.

Positiv hervorzuheben ist der Einsatz der *Internationalen Diagnose Checklisten* nach der *International Classification of Diagnosis* (IDCL nach ICD-10) zu Therapiebeginn. Durch das validierte Diagnoseverfahren wurde ermöglicht, dass Therapeut\*innen eine

verfahrensübergreifende und klinisch reliable Diagnose nach dem aktuellen Klassifikationssystem stellen konnten. Der für ambulante Psychotherapeuten\*innen typische Mehranteil an F3- und F4-Diagnosen konnte auch in dieser Studie gefunden werden.

Einen weiteren Vorteil stellt die Möglichkeit dar, im Zuge einer Studienteilnahme Rückmeldungen aus den Daten zu erhalten. Das sollte als Anreiz zur Teilnahme und als zusätzliche Ergänzung zum therapeutischen Eindruck dienen (Castonguay et al., 2015; Steffanowski et al., 2011; Wittmann & Steffanowski, 2011). Aus den Erfahrungsberichten der Therapeuten\*innen ging eine gute Annahme der Rückmeldungen hervor, was auch durch Befunde anderer Studien bestätigt werden konnte (Strauss et al., 2015; Wittmann et al., 2011). Einige Therapeuten\*innen berichteten, dass Informationen aus den Rückmeldungen besonders bei unklaren Diagnosen und schwierigen Verläufen verwendet wurden, was eine Unterstützung in der täglichen Praxis bedeutete (Rubel et al., 2017). Die individuellen Klient\*innenverläufe rücken auf diese Weise innerhalb des Studiendesigns stärker in den Fokus, und die Daten erhalten bereits während der Erhebung einen Nutzen. Daraus ergibt sich ein evaluativer Studiencharakter bereits während der Datenerhebung und nicht erst bei der finalen Ergebnisanalyse (Castonguay et al., 2013).

Die Untersuchung der Ergebnisqualität existenzanalytischer Behandlungen bei Psychotherapeuten\*innen in Ausbildung unter Supervision in freier Praxis ermöglicht nicht nur ein unterrepräsentiertes Forschungs- und Berufsfeld stärker in den Fokus der Aufmerksamkeit zu rücken, sondern erhöht auch die Chance eine nachhaltige Forschungskultur zu schaffen. Gerade Ausbildungskandidat\*innen, die in dieser Phase ihre berufliche Identität entwickeln, können durch eine frühe Forschungsbeteiligung zu einem ausgeprägten Bewusstsein für die Wissenschaftspraxis ihres Faches angeregt werden (Castonguay et al., 2015). Zudem können besonders zu Berufsbeginn zusätzliche Informationsquellen, wie Ergebnisrückmeldungen oder Diagnose-Checklisten, als Unterstützung der täglichen Behandlungspraxis dienen (Rubel et al., 2017; Strauss et al., 2015).

### **Konklusion und Ausblick**

Für die Pilotstudie der GLE-Ö zur Untersuchung der praktischen Wirksamkeit von existenzanalytischen Behandlungen durch Psychotherapeuten\*innen i.A.u.S. in freier Praxis konnte vor dem Hintergrund vorangegangener Diskussion die methodische Durchführbarkeit und bei den teilnahmebereiten Klient\*innen und Therapeuten\*innen eine gute Studienakzeptanz festgestellt werden. Die erreichte Gesamtstichprobe der Therapeuten\*innen entsprach hinsichtlich wichtiger soziodemographischer Parameter der untersuchten Zielgruppe (Sagerschnig & Valady, 2021). Auch die Gesamtstichprobe der Klient\*innen wies in

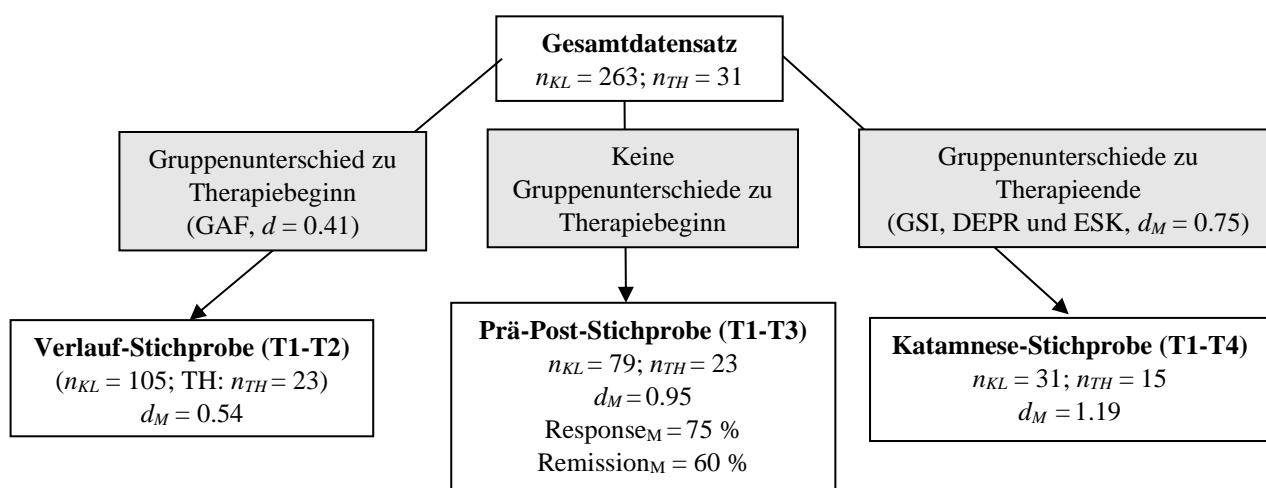
relevanten Variablen, wie Alter, Geschlecht, Diagnoseverteilung, Sitzungszahl und psychischer Belastung zu Therapiebeginn, ähnliche Verteilungen wie vergleichbare naturalistische Studien aus der Literatur auf (Gallas et al., 2008, 2010; Steffanowski et al., 2011; Strauss et al., 2015; Tschuschke et al., 2015; Wittmann et al., 2011).

Die Ergebnisse der gruppenstatistischen, individuellen und direkten Veränderungsmessungen der Prä-Post-Stichprobe stützen sich weitestgehend gegenseitig und deuten auf eine hohe Ergebnisqualität der durchgeführten Behandlungen durch existenzanalytische Psychotherapeut\*innen in Ausbildung unter Supervision in freier Praxis hin (vgl. Abbildung 9). Die ermittelten Behandlungseffektstärken für den Zeitraum zwischen Therapiebeginn und Therapieende sind als hoch einzustufen und vergleichbar mit den Behandlungsergebnissen ausgebildeter Therapeut\*innen anderer praxisorientierter Studien (Gallas et al., 2008; Laireiter et al., 2013; Strauss et al., 2015; Tschuschke et al., 2015; Victor et al., 2018; Wittmann et al., 2011).

Auch die Vergrößerung der Effektstärken mit jedem weiteren Messzeitpunkt entspricht Befunden anderer Psychotherapiestudien (Längle et al., 2005; Wittmann et al., 2011). Die Therapieeffekte im kleinen bzw. mittleren Bereich der Verlaufsstichprobe weisen auf die Notwendigkeit ausreichend langer Therapie bei gegebener Indikation hin. Die Befunde der selektiven Katamnese-Stichprobe deuten einen nachhaltigen und stabilen Therapieeffekt an (vgl. Abbildung 9).

### Abbildung 9

Zusammenfassung der Ergebnisse



Anmerkung.  $n_{KL}$  = Stichprobenanzahl der Klient\*innen;  $n_{TH}$  = Stichprobenanzahl der Therapeut\*innen;  $d_M$  = über die Verfahren gemittelte Effektstärke;  $Response_M$  = über alle Verfahren gemittelte *Response*-Rate;  $Remission_M$  = über alle Verfahren gemittelte *Remission*-Rate; in grau gehaltene Boxen = Unterschiede der Teilstichprobe zur Gesamtstichprobe zu Therapiebeginn; GAF = Global Assessment of Functioning; GSI = Global Severity Index; DEPR = Depressionsskala der SCL-90; ESK = Gesamtskala der Existenz-Skala.

Ausbildungstherapien leisten somit nicht nur quantitativ sondern auch qualitativ einen wesentlichen Beitrag zur psychotherapeutischen Versorgung der Bevölkerung (Löffler-Stastka et al., 2019; Peikert et al., 2014). Die Befunde sprechen indirekt auch für die Behandlungsqualität der Personalen Existenzanalyse im Allgemeinen, da davon auszugehen ist, dass sich besonders berufsneue Therapeut\*innen an die verfahrenseigenen Therapiekonzepte orientieren (Tschuschke et al., 2015; Wampold, 2021).

Im Zuge zukünftiger Studien zur Wirksamkeit existenzanalytischer Psychotherapie sollte versucht werden, größere Klient\*innen- und Therapeut\*innenstichproben zu generieren. Eine größere Stichprobe würde differenziertere Analysen ermöglichen. So wären störungsspezifische oder veränderungsspezifische Untersuchungen von (klinisch) signifikant gebesserten, unveränderten und verschlechterten Personen von besonderem Interesse für praktische Ableitungen. Des Weiteren sollten Daten von vorzeitig beendeten Therapien systematischer ermittelt werden. Diese könnten als Vergleichsgruppe eingesetzt werden. Auch zum Katamnesezeitpunkt sollte eine größere Stichprobe erlangt werden, um Selektionseffekte möglichst gering zu halten. Das Maß der individuellen Veränderung (RCI) sollte auch für die Verlauf- und Katamnese-Stichprobe berechnet werden, um diese Ergebnisse den Befunden der Gruppenanalysen gegenüberzustellen. Interessant wäre weiters, ein möglicher Einfluss von Rückmeldungen auf den Therapieprozess selbst zu untersuchen. Dafür wären Fragen den Umgang betreffend im Therapeut\*innenfragebogen zu empfehlen. Auch Dosis- und Wirkungszusammenhänge könnten in zukünftigen, zeitlich offenen, naturalistischen Studien genauer untersucht werden. Aus vorliegendem Datenbefund ging per se kein Einfluss der Anzahl an Therapieeinheiten auf das Behandlungsergebnis hervor. Jedoch deuteten die geringen Effektstärken in der Verlaufsstichprobe auf die Notwendigkeit einer bedarfsorientierten Psychotherapieversorgung bei gegebener Indikation hin. Es sollte untersucht werden, welche Personen längerer Therapiezeiträume bedürfen. Dafür wären kürzere Abstände bei Verlaufsmessungen mit kürzeren Fragebögen anzuraten.

Durch den Einsatz von nicht nur symptom-, sondern auch ressourcenbezogenen Erhebungsinstrumenten wie der Existenzskala sollten über Mediationsanalysen die Wirkungswege von psychotherapeutischer Behandlung weiter untersucht werden. Es wäre z.B. zu klären, ob eine Symptomreduzierung über die Verbesserung existenzieller Fähigkeiten erfolgt oder umgekehrt. Der starke Zusammenhang der ESK-Gesamtskala mit den Symptomskalen spricht für einen bedeutsamen Einfluss sinnzentrierter Inhalte auf das psychische Befinden.

Neben den inhaltlichen Implikationen lassen sich auch methodische Empfehlungen aus der vorliegenden Studie ableiten. Auch weiterhin sollte das Augenmerk von Durchführenden naturalistischer Studien auf die methodischen Anforderungen und Umsetzungsmöglichkeiten verschiedener Studiendesigns in diesem Setting liegen. Ein gut evaluierter Methoden- und Kriterienkatalog scheint sinnvoll und notwendig, um Orientierung, aber auch Qualitätsstandards bei praxisorientierten Studien zu gewährleisten. Genaue Beschreibungen der methodischen Durchführbarkeit von naturalistischen Forschungsprojekten können zukünftige Studienvorhaben zumindest in der Planung erleichtern. Eine wachsende Zahl an Forschungsbefunden zur Psychotherapie aus dem naturalistischen Kontext verbessert die Interpretierbarkeit von Studienergebnissen (*benchmarking*) (Berg, 2021; Minami et al., 2008; Seligman, 1995). Trotz der Herausforderungen, die eine praxisorientierte Studiendurchführung mit sich bringt, überwiegen die Vorteile eines anpassungsfähigen Studiendesign, das die untersuchten Therapien zeitlich und thematisch nicht einschränkt.

Psychotherapie ist wirksam, das gilt mittlerweile als erwiesen (Datler et al., 2021; Wampold & Imel, 2015). In der öffentlichen Versorgung sollte der Bedarf an Psychotherapie ausfinanziert und psychotherapeutische Behandlung auf Krankenschein nicht nur gesetzlich beschlossen, sondern auch tatsächlich ermöglicht werden. Die Versorgungsleistung von Psychotherapeut\*innen in Ausbildung unter Supervision sollte dahingehend anerkannt und in die gesundheitspolitische Planung integriert werden. Naturalistische Evaluationsstudien können durch ihre methodischen Stärken diesen Prozess realitätsnah und qualitätssichernd begleiten. Die Dringlichkeit dieser Forderungen wird durch die hohe Prävalenz psychischer Störungen und deren gesellschaftlichen Auswirkungen betont und durch die beträchtliche Krankheitslast für Betroffene unterstrichen (Riffer et al., 2018). Die vorherrschende Schiefelage zugunsten psychopharmakologischer Behandlungen sollte im Sinne eines ganzheitlichen Behandlungskonzeptes korrigiert werden (Grabenhofer-Eggerth & Sator, 2020; Löffler-Stastka & Hochgerner, 2021).

Eine bedarfsorientierte Psychotherapieversorgung und -forschung ist aus gesamtgesellschaftlicher und individueller Sicht eine nachhaltige und wesentliche Investition in die Zukunft und sollte umgehend forciert werden (Altmann et al., 2016; Rieß & Löffler-Stastka, 2022; Strauss, 2019; Wampold, 2021).

*Großen Dank allen Teilnehmenden der Studie.*

*Herzlichen Dank an die beiden Betreuer dieser Arbeit  
Univ.-Prof. Dr. Anton-Rupert Laireiter & Dr. Clemens Fartacek.*

### Literaturverzeichnis

- Aas, I. M. (2011). Guidelines for rating Global Assessment of Functioning (GAF). *Annals of General Psychiatry, 10*(1), 2. <https://doi.org/10.1186/1744-859X-10-2>
- Altmann, U., Thielemann, D., & Strauß, B. (2016). Ambulante Psychotherapie unter Routinebedingungen: Forschungsbefunde aus dem deutschsprachigen Raum. *Psychiatrische Praxis, 43*(7), 360–366. <https://doi.org/10.1055/s-0042-115616>
- Andrade, C. (2021). The ceiling effect, the floor effect, and the importance of active and placebo control arms in randomized controlled trials of an investigational drug. *Indian Journal of Psychological Medicine, 43*(4), 360–361. <https://doi.org/10.1177/02537176211021280>
- Angus, L., Watson, J., Elliott, R., Schneider, K., & Timulak, L. (2015). Humanistic psychotherapy research 1990–2015: From methodological innovation to evidence-supported treatment outcomes and beyond. *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 25*, 1–18. <https://doi.org/10.1080/10503307.2014.989290>
- Ardito, R. B., & Rabellino, D. (2011). Therapeutic alliance and outcome of psychotherapy: Historical excursus, measurements, and prospects for research. *Frontiers in Psychology, Volume 2*. <https://doi.org/10.3389/fpsyg.2011.00270>
- Bandelow, B. (2016). *Panik- und Agoraphobieskala (PAS)* (2.Auflage). Hogrefe. <https://www.hogrefe.com/at/shop/panik-und-agoraphobie-skala.html#1+1>
- Barkham, M., & Margison, F. (2007). Practice-based evidence as a complement to evidence-based practice: From dichotomy to chiasmus. In C. Freeman, & M. Power (Hrsg.), *Handbook of evidence-based psychotherapies* (S. 443–476). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470713242.ch25>
- Berg, H. (2021). Why only efficiency, and not efficacy, matters in psychotherapy practice. *Frontiers in Psychology, 12*, 603211. <https://doi.org/10.3389/fpsyg.2021.603211>
- Bernsteiner, M., & Boggatz, T. (2016). Wohlbefinden im Alter: Die Inhaltsvalidität der Ryff-Skala für BewohnerInnen von Pflegeheimen und betreuten Wohneinrichtungen. *Pflege, 29*, 137–149. <https://doi.org/10.1024/1012-5302/a000485>
- Bhide, A., Shah, P. S., & Acharya, G. (2018). A simplified guide to randomized controlled trials. *Acta Obstetricia et Gynecologica Scandinavica, 97*(4), 380–387. <https://doi.org/10.1111/aogs.13309>
- Bohart, A., O'Hara, M., & Leitner, L. (1998). Empirically violated treatments: Disenfranchisement of humanistic and other psychotherapies. *Psychotherapy Research, 8*(2), 141–157. <https://doi.org/10.1080/10503309812331332277>
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Auflage). Springer. <https://doi.org/10.1007/978-3-642-12770-0>
- Bosch, A. (2019). Freude, Glück, Wohlbefinden. In H. Kappelhoff, J.-H. Bakels, H. Lehmann, & C. Schmitt (Hrsg.), *Emotionen: Ein interdisziplinäres Manual* (S. 144–149). J.B. Metzler. [https://doi.org/10.1007/978-3-476-05353-4\\_20](https://doi.org/10.1007/978-3-476-05353-4_20)

- Buchholz, M. (2007). Der Fall von RCT - Kritik einer Hierarchie. *Psycho-News-Letter*, Nr. 55. Deutsche Gesellschaft für Psychoanalyse und Tiefenpsychologie.  
[https://www.researchgate.net/publication/282643084\\_Der\\_Fall\\_von\\_RCT\\_-\\_Kritik\\_einer\\_Hierarchie\\_The\\_Fall\\_of\\_RCT\\_-\\_critic\\_of\\_a\\_hierarchy](https://www.researchgate.net/publication/282643084_Der_Fall_von_RCT_-_Kritik_einer_Hierarchie_The_Fall_of_RCT_-_critic_of_a_hierarchy)
- Cain, D. J., Keenan, K., & Rubin, S. (2016). Introduction. In D. J., Cain, K., Keenan, & S. Rubin (Hrsg.), *Humanistic psychotherapies: Handbook of research and practice* (2. Auflage). American Psychological Association. <https://doi.org/10.1037/14775-000>
- Cámara, R., Begré, S., & von Känel, R. (2010). Der Effekt stressreduzierender Interventionen auf chronisch entzündliche Darmerkrankungen: Qualitätskontrolle von zehn therapeutischen Studien. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 56, 116–135.  
<https://doi.org/10.13109/zptm.2010.56.2.116>
- Caspar, F., & Jacobi, F. (2004). Psychotherapieforschung. In W. Hiller & E. Leibing (Hrsg.), *Lehrbuch der Psychotherapie* (S. 395-410). Springer. [https://doi.org/10.1007/978-3-540-29988-2\\_6](https://doi.org/10.1007/978-3-540-29988-2_6)
- Castonguay, L., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-oriented research: Approaches and applications. In M. Lambert (Hrsg.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6. Auflage, S. 85–133). Wiley.
- Castonguay, L., Barkham, M., Youn, S. J., & Page, A. C. (2021). Practice-based evidence- findings from routine clinical settings. In M. Barkham, W. Lutz, & L. Castonguay (Hrsg.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7. Auflage, S. 191-222). Wiley.
- Castonguay, L. G., Youn, S. J., Xiao, H., Muran, J. C., & Barber, J. P. (2015). Building clinicians-researchers partnerships: Lessons from diverse natural settings and practice-oriented initiatives. *Psychotherapy Research*, 25(1), 166–184. <https://doi.org/10.1080/10503307.2014.973923>
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685–716.  
<https://doi.org/10.1146/annurev.psych.52.1.685>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Auflage). Routledge.  
<https://doi.org/10.4324/9780203771587>
- Cope, D. G. (2015). Conducting pilot and feasibility studies. *Oncology Nursing Forum*, 42(2).  
<https://doi.org/10.1188/15.ONF.196-197>
- Datler, W., Hochgerner, M., Korunka, C., Löffler-Stastka, H., & Pawlowsky, G. (2021). Disziplin, Profession und evidenzbasierte Praxis: Zur Stellung der Psychotherapie im Gesundheitssystem. Eine Bilanz. *Psychotherapie Forum*, 25(1), 7–21. <https://doi.org/10.1007/s00729-021-00167-4>
- Cramer, A., Koemeda, M., Tschuschke, V., Schulthess, P., & von Wyl, A. (2014). Ergebnisqualität ambulanter Psychotherapie: Ergebnisse aus der Grundversorgung in der Schweiz. *Psychotherapie Forum*, 4(2), 96–107. <https://doi.org/10.21256/zhaw-1969>
- Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2–21. <https://doi.org/10.1016/j.socscimed.2017.12.005>

- Desmet, M. (2013). Experimental versus naturalistic psychotherapy research: Consequences for researchers, clinicians, policy makers, and patients. *Psychoanalytische Perspectieven*, 31(1), Article 1. <https://biblio.ugent.be/publication/5645063>
- Dick, F., & Kringler, W. (2007). Evidenzbasierung, Methodik, therapeutische Freiheit und Kreativität. *Zeitschrift Für Neuropsychologie*, 18(1), 41–54. <https://doi.org/10.1024/1016-264X.18.1.41>
- Doering, S., Renn, D., Höfer, S., Rumpold, G., Smrekar, U., Janecke, N., Schatz, D., & Schotte, C. (2007). Validierung der deutschen Version des Fragebogens zur Erfassung von DSM-IV Persönlichkeitsstörungen (ADP-IV). *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 53(2), S.111-128. <https://doi.org/10.13109/zptm.2007.53.2.111>
- Donald, G. (2018). A brief summary of pilot and feasibility studies: Exploring terminology, aims, and methods. *European Journal of Integrative Medicine*, 24, 65–70. <https://doi.org/10.1016/j.eujim.2018.10.017>
- Eckert, J. (2013). Machtmissbrauch in den Psychotherapiewissenschaften. *Persönlichkeitsstörungen. Theorie und Therapie*, 17, 278–287. [https://www.researchgate.net/publication/262313412\\_Machtmissbrauch\\_in\\_den\\_Psychotherapiewissenschaften](https://www.researchgate.net/publication/262313412_Machtmissbrauch_in_den_Psychotherapiewissenschaften)
- Eiling, A., Schlipfenbacher, C., Hörz-Sagstetter, S., & Jacobi, F. (2014). Über die Zukunft der Evidenzbasierten Psychotherapie und die Beziehung zwischen Praxis und Forschung. Aktuelle Studien zur Versorgungsforschung. *Psychotherapeutenjournal*, 2, 2014. [https://www.researchgate.net/publication/269278660\\_Uber\\_die\\_Zukunft\\_der\\_Evidenzbasierten\\_Psychotherapie\\_und\\_die\\_Beziehung\\_zwischen\\_Praxis\\_und\\_Forschung\\_Aktuelle\\_Studien\\_zur\\_Versorgungsforschung](https://www.researchgate.net/publication/269278660_Uber_die_Zukunft_der_Evidenzbasierten_Psychotherapie_und_die_Beziehung_zwischen_Praxis_und_Forschung_Aktuelle_Studien_zur_Versorgungsforschung)
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, 16, 319–324. <https://doi.org/10.1037/h0063633>
- Fäh, M. (2006). Der ‘Grawe-Effekt’: Von der Polemik zum Katalysator von positiven Entwicklungen im Psychotherapiefeld. *Psychotherapie Forum*, 14(1), 42-46. <https://doi.org/10.1007/s00729-006-0118-5>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Felber, M., & Margreiter, U. (2007). Einstellung von Psychotherapeut(inn)en zu Psychotherapieforschung. *Psychotherapie Forum*, 15(4), 192–196. <https://doi.org/10.1007/s00729-007-0214-1>
- Figl, C., & Laireiter, A.-R. (2022). Teilnahmeinteresse, Wünsche und Visionen österreichischer Psychotherapeut\_innen zur Psychotherapieforschung – Ergebnisse einer Online-Befragung. *Psychotherapie Forum*, 26, 73–79 (2022). <https://doi.org/10.1007/s00729-022-00203-x>
- Franke, G. (2014). *SCL®-S – Symptom-Checklist-90®—Standard*. Hogrefe.
- Frieden, T. (2017). Evidence for health decision making-beyond randomized, controlled trials. *New England Journal of Medicine*, 377, 465–475. <https://doi.org/10.1056/NEJMra1614394>

- Fydrich, T., & Schneider, W. (2007). Evidenzbasierte Psychotherapie. *Psychotherapeut*, 52(1), 55–68.  
<https://doi.org/10.1007/s00278-006-0522-x>
- Gadke, D. L., Kratochwill, T. R., & Gettinger, M. (2021). Incorporating feasibility protocols in intervention research. *Journal of School Psychology*, 84, 1–18. <https://doi.org/10.1016/j.jsp.2020.11.004>
- Gallas, C., Kächele, H., Kraft, S., Kordy, H., & Puschner, B. (2008). Inanspruchnahme, Verlauf und Ergebnis ambulanter Psychotherapie. *Psychotherapeut*, 53(6), 414–423.  
<https://doi.org/10.1007/s00278-008-0641-7>
- Gallas, C., Puschner, B., Kühn, A., & Kordy, H. (2010). Dauer und Umfang ambulanter Psychotherapie und Implikationen für die Versorgungspraxis. *Psychotherapie Psychosomatik Medizinische Psychologie*, 60 (1), 5–13. <https://doi.org/10.1055/s-0029-1243258>
- Gelo, O. C. G., Lagetto, G., Dinoi, C., Belfiore, E., Lombi, E., Blasi, S., Aria, M., & Ciavolino, E. (2020). Which methodological practice(s) for psychotherapy science? A systematic review and a proposal. *Integrative Psychological and Behavioral Science*, 54(1), 215–248. <https://doi.org/10.1007/s12124-019-09494-3>
- Görtz, A. (2020). Phänomenologisch-hermeneutische Struktur- und Prozessdiagnostik in der Existenzanalyse. *Psychotherapie Forum*, 24(1), 26–36. <https://doi.org/10.1007/s00729-020-00138-1>
- Grabenhofer-Eggerth, A., & Sator, M. (2020). *Psychotherapie als Sozialversicherungsleistung- Inanspruchnahme und Finanzierung*. Gesundheit Österreich, Wien.  
<https://jasmin.goeg.at/id/eprint/1522>
- Grawe, K. (1992). Psychotherapieforschung zu Beginn der neunziger Jahre. *Psychologische Rundschau*, 43(3), 132–162. <https://psycnet-apa-org.uaccess.univie.ac.at/record/1992-88855-001>
- Hansen, N. B., & Lambert, M. J. (1996). Brief Report: Assessing clinical significance using the inventory of interpersonal problems. *Assessment*, 3(2), 133–136. <https://doi.org/10.1177/107319119600300204>
- Hausler, M., Huber, A., Strecker, C., Brenner, M., Höge, T., & Höfer, S. (2017). Validierung eines Fragebogens zur umfassenden Operationalisierung von Wohlbefinden. *Diagnostica*, 63(3), 219–228.  
<https://doi.org/10.1026/0012-1924/a000174>
- Heekerens, H.-P. (2005). Vom Labor ins Feld. *Psychotherapeut*, 50(5), 357–366.  
<https://doi.org/10.1007/s00278-005-0441-2>
- Helle, M. (2019). Humanistische Psychologie und Humanistische Psychotherapie. In M. Helle (Hrsg.), *Psychotherapie* (S. 61–93). Springer. [https://doi.org/10.1007/978-3-662-58712-6\\_3](https://doi.org/10.1007/978-3-662-58712-6_3)
- Herzog, P., Kaiser, T., & Brakemeier, E.-L. (2022). Praxisorientierte Forschung in der Psychotherapie. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 51(2), 127–148.  
<https://doi.org/10.1026/1616-3443/a000665>
- Heuft, G., & Lange, C. (2015). Die Beeinträchtigungsschwere in der psychosomatischen und psychiatrischen Qualitätssicherung: Global Assessment of Functioning Scale (GAF) vs. Beeinträchtigungs-Schwere-Score (BSS). *Zeitschrift für Psychosomatische Medizin und Psychotherapie*. <https://doi.org/10.13109/zptm.2002.48.3.256>

- Hiller, W., Bleichhardt, G., & Schindler, A. (2009). Evaluation von Psychotherapien aus der Perspektive von Qualitätssicherung und Qualitätsmanagement. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 57(1), 7–22. <https://doi.org/10.1024/1661-4747.57.1.7>
- Hiller, W., Schindler, A., Andor, T., & Rist, F. (2011). Vorschläge zur Evaluation regulärer Psychotherapien an Hochschulambulanzen im Sinne der Phase-IV-Therapieforschung. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, 40(1), 22–32. <https://doi.org/10.1026/1616-3443/a000063>
- Hiller, W., Schindler, A., & Lambert, M. (2011). Defining response and remission in psychotherapy research: A comparison of the RCI and the method of percent improvement. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, 22(1), 1–11. <https://doi.org/10.1080/10503307.2011.616237>
- Hiller, W., Zaudig, M., & Mombour, W. (1995). *IDCL. Internationale Diagnose Checklisten für ICD-10*. Huber. <https://www.testzentrale.de/shop/idcl-fuer-icd-10-icd-10-checklisten-69247.html>
- Horowitz, L. M., Strauss, B., & Kordy, H. (2000). *Das Inventar zur Erfassung interpersonaler Probleme* (2. Auflage). Beltz-Test-Gesellschaft.
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy*, 48(1), 9–16. <https://doi.org/10.1037/a0022186>
- Jacobi, F. (2020). Entwicklung und Beurteilung therapeutischer Interventionen. In J. Hoyer, & S. Knappe (Hrsg.), *Klinische Psychologie & Psychotherapie* (S. 472-503). Springer. [https://doi.org/10.1007/978-3-662-61814-1\\_19](https://doi.org/10.1007/978-3-662-61814-1_19)
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12–19. <https://doi.org/10.1037//0022-006x.59.1.12>
- Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology*, 17(1), 162. <https://doi.org/10.1186/s12874-017-0442-1>
- Jensen, S. A., & Corralejo, S. M. (2017). Measurement Issues: Large effect sizes do not mean most people get better – clinical significance and the importance of individual results. *Child and Adolescent Mental Health*, 22(3), 163–166. <https://doi.org/10.1111/camh.12203>
- Jensen, H. H., Mortensen, E. L., & Lotz, M. (2017). Missing outcome data in a naturalistic psychodynamic group therapy study. *Psychology*, 8 (3), Article 3. <https://doi.org/10.4236/psych.2017.83027>
- Kriz, J. (2014). *Grundkonzepte der Psychotherapie: Mit Online-Materialien* (7. Auflage). Beltz.
- Kriz, J. (2019). «Evidenzbasierung» als Kriterium der Psychotherapie-Selektion? *Psychotherapie-Wissenschaft*, 9(2), Article 2. <https://doi.org/10.30820/1664-9583-2019-2-42>
- Kubinger, K. D., Rasch, D., & Moder, K. (2009). Zur Legende der Voraussetzungen des t-Tests für unabhängige Stichproben. *Psychologische Rundschau*, 60(1), 26–27. <https://doi.org/10.1026/0033-3042.60.1.26>

- Kuckartz, U. (2014). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (2. Auflage). Beltz Juventa.
- Laireiter, A., Schaireiter, M. M., Schirl-Russegger, C., Baumgartner, I., Längle, A., & Sauer, J. (2013). Die Wirksamkeit von Existenzanalyse und Logotherapie: Retrospektive Beurteilungen. *Existenzanalyse*, 30(1), 4–14. [https://www.existenzanalyse.org/service/gle-downloads/forschungspublikationen/quantitative-forschung/?wppcp\\_file\\_download=yes&wppcp\\_private\\_file\\_id=104&wppcp\\_post\\_id=6844](https://www.existenzanalyse.org/service/gle-downloads/forschungspublikationen/quantitative-forschung/?wppcp_file_download=yes&wppcp_private_file_id=104&wppcp_post_id=6844)
- Laireiter, A., Schirl, C. F., Kimeswenger, I., Längle, A., & Sauer, J. (2000). Zufriedenheit mit Existenzanalyse: Ergebnisse einer katamnesticen Feldstudie zur Patientenbeurteilung von existenzanalytisch-logotherapeutischer Psychotherapie. *Existenzanalyse*, 17(3), 40–50. [https://www.existenzanalyse.org/service/gle-downloads/forschungspublikationen/qualitative-forschung/?wppcp\\_file\\_download=yes&wppcp\\_private\\_file\\_id=66&wppcp\\_post\\_id=6809](https://www.existenzanalyse.org/service/gle-downloads/forschungspublikationen/qualitative-forschung/?wppcp_file_download=yes&wppcp_private_file_id=66&wppcp_post_id=6809)
- Laireiter, A.-R., & Kopetzky, U. (2015). Die Wirksamkeit von Psychotherapie in der Praxis. Erkenntnisse aus einer missglückten Praxisstudie. In Koordinationsstelle Psychotherapieforschung an der Gesundheit Österreich GmbH (Hrsg.), *Praxisorientierte Psychotherapieforschung. Verfahrensübergreifende, patientenorientierte Aspekte und Kompetenzentwicklung. Tagungsband zur wissenschaftlichen Tagung 2014* (S. 14-22). Österreichisches Bundesinstitut für Gesundheitswesen (ÖBIG). [https://jas-min.goeg.at/103/1/Tagungsband %20zur %20Fachtagung %20Psychotherapieforschung.pdf](https://jas-min.goeg.at/103/1/Tagungsband%20zur%20Fachtagung%20Psychotherapieforschung.pdf)
- Lambert, M. J., & Ogles, B. M. (2009). Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. *Psychotherapy Research*, 19(4–5), 493–501. <https://doi.org/10.1080/10503300902849483>
- Längle, A. (2014). *Lehrbuch zur Existenzanalyse: Grundlagen* (2. Auflage). Facultas. WUV.
- Längle, A. (2021). *Existenzanalyse und Logotherapie*. Stuttgart: Kohlhammer
- Längle, Silvia (2007), Phänomenologische Forschung in der Existenzanalyse. *Existenzanalyse*, 24, 2, 54 – 61 [https://www.existenzanalyse.org/service/gle-downloads/forschungspublikationen/qualitative-forschung/?wppcp\\_file\\_download=yes&wppcp\\_private\\_file\\_id=36&wppcp\\_post\\_id=6809](https://www.existenzanalyse.org/service/gle-downloads/forschungspublikationen/qualitative-forschung/?wppcp_file_download=yes&wppcp_private_file_id=36&wppcp_post_id=6809)
- Längle, A., & Kriz, J. (2012). The renewal of humanism in European psychotherapy: Developments and applications. *Psychotherapy*, 49(4), 430–436. <https://doi.org/10.1037/a0027397>
- Längle, A., Görtz, A., Probst, C., Probst, M., Lopatka, C., Kubin, M., & Steinert, K. (2005). Wie wirksam ist existenzanalytische Psychotherapie: Ein Projektbericht zur Existenzanalyse. *Psychotherapie Forum*, 13(2), 54–60. <https://doi.org/10.1007/s00729-005-0087-0>
- Längle, A., Görtz, A., Rauch, A., Jarosik, H., & Haller, R. (2000). Effektivitätsstudie zur Existenzanalyse und explorativer Vergleich mit anderen Psychotherapiemethoden im stationären Setting. *Existenzanalyse*, 17(3), 17–29. [https://www.existenzanalyse.org/service/gle-downloads/forschungspublikationen/qualitative-forschung/?wppcp\\_file\\_download=yes&wppcp\\_private\\_file\\_id=68&wppcp\\_post\\_id=6809](https://www.existenzanalyse.org/service/gle-downloads/forschungspublikationen/qualitative-forschung/?wppcp_file_download=yes&wppcp_private_file_id=68&wppcp_post_id=6809)

- Längle, S., & Klaassen, D. (2019). Logotherapy and Existential Analysis. In E. van Deurzen, E. Craig, A. Längle, K. Schneider, D. Tantam, & S. du Plock (Hrsg.), *The Wiley world handbook of Existential Therapy* (S. 341–355). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119167198.ch21>
- Längle, A., Van Deurzen, E., Craig, E., Schneider, K., Tantam, D., & Plock, S. (2019). The History of Logotherapy and Existential Analysis. In E. van Deurzen, E. Craig, A. Längle, K. Schneider, D. Tantam, & S. du Plock (Hrsg.), *The Wiley world handbook of Existential Therapy* (S. 309–323). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119167198.ch19>
- Launeanu, M., Klaassen, D., & Muir, B. A. (2019). Logotherapy and Existential Analysis Therapy Illustration. In E. van Deurzen, E. Craig, A. Längle, K. Schneider, D. Tantam, & S. du Plock (Hrsg.), *The Wiley world handbook of Existential Therapy* (S. 356–368). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119167198.ch22>
- Legewie, H. (2015). Goldstandard für die Psychotherapieforschung: Kontrollierte oder ökologisch valide Studien?. *Gesprächspsychotherapie und Personzentrierte Beratung* 2000, (2), 125-128. <https://doi.org/10.13140/RG.2.1.2102.8004>
- Leichsenring, F., Abbass, A., Beutel, M., Gündel, H., Heuft, G., Hoffmann, S. O., Kächele, H., Kruse, J., Rieger, U., Rudolf, G., Spitzer, C., Salzer, S., Luyten, P., Wampold, B., & Steinert, C. (2019). Vom Sinn des Verfahrenskonzepts und der Verfahrensvielfalt – und warum das Baukasten-System in der Psychotherapie nicht funktioniert. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 65 (4), 321-340. <https://doi.org/10.13109/zptm.2019.65.4.oa1>
- Leichsenring, F., Abbass, A., Hilsenroth, M. J., Luyten, P., Munder, T., Rabung, S., & Steinert, C. (2018). “Gold standards,” plurality and monocultures: The need for diversity in psychotherapy. *Frontiers in Psychiatry*, 9, 159. <https://doi.org/10.3389/fpsy.2018.00159>
- Leichsenring, F., & Rieger, U. (2004). Psychotherapeutische Behandlungsverfahren auf dem Prüfstand der Evidence Based Medicine (EBM). *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 50, 203–217. <https://doi.org/10.13109/zptm.2004.50.2.203>
- Little, R. J. A. (1988). A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association*, 83(404), 1198–1202. <https://doi.org/10.1080/01621459.1988.10478722>
- Little, R. J. A., & Rubin, D. B. (2020). *Statistical analysis with missing data* (3rd edition). Wiley. <https://onlinelibrary.wiley.com/doi/book/10.1002/9781119482260>
- Löffler-Stastka, H. (2013). Geschichte und Entwicklung der Psychotherapieforschung, ihre unterschiedlichen Forschungsparadigmen und Zielsetzung. In Koordinationsstelle Psychotherapieforschung an der Gesundheit Österreich GmbH (Hrsg.), *Psychotherapieforschung. Wissenschaftliche Beratung und Vernetzung. Tagungsband zum Workshop 2012* (S. 13-20). Österreichisches Bundesinstitut für Gesundheitswesen (ÖBIG). <https://jasmin.goeg.at/id/eprint/386>
- Löffler-Stastka, H., & Datz, F. (2016). State of the Art praxisorientierte Psychotherapieforschung. *Psychotherapie Forum*, 21(2), 44–52. <https://doi.org/10.1007/s00729-015-0052-5>

- Löffler-Stastka, H., Gelo, O. C. G., Pleschberger, I., Hartmann, A., Orlinsky, D., Rønnestad, M. H., & Willutzki, U. (2019). Psychotherapie-Berufsausbildung in Österreich: Basis- und soziodemographische Hintergrunddaten aus einer SPRISTAD-Pilotstudie. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, *65*, 341–352. <https://doi.org/10.13109/zptm.2019.65.4.341>
- Löffler-Stastka, H., Gelo, O., Pleschberger, I., Schröder, T., Orlinsky, D. E., Rønnestad, M. H., & Willutzki, U. (2018). Psychotherapieausbildung in Österreich. *psychopraxis neuropraxis*, *21*(5), 227–231. <https://doi.org/10.1007/s00739-018-0508-9>
- Löffler-Stastka, H., & Hochgerner, M. (2021). Versorgungswirksamkeit von Psychotherapie in Österreich. *psychopraxis neuropraxis*, *24*(1), 57–61. <https://doi.org/10.1007/s00739-020-00686-w>
- Lutz, W., Castonguay, L., Lambert, M., & Barkham, M. (2021). Traditions and new beginnings: historical and current perspectives on research in psychotherapy and behavior change. In M. Barkham, W. Lutz, & L. Castonguay (Hrsg.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (7. Auflage, S. 3–18). Wiley.
- Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, *110*, 63–73. <https://doi.org/10.1016/j.jclinepi.2019.02.016>
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12. Auflage). Beltz.
- Minami, T., Wampold, B. E., Serlin, R. C., Hamilton, E. G., Brown, G. S., & Kircher, J. C. (2008). Benchmarking the effectiveness of psychotherapy treatment for adult depression in a managed care environment: A preliminary study. *Journal of Consulting and Clinical Psychology*, *76*(1), 116–124. <https://doi.org/10.1037/0022-006X.76.1.116>
- Munder, T., Wilmers, F., Leonhart, R., Linster, H. W., & Barth, J. (2010). Working Alliance Inventory-Short Revised (WAI-SR): Psychometric properties in outpatients and inpatients. *Clinical Psychology & Psychotherapy*, *17*(3), 231–239. <https://doi.org/10.1002/cpp.658>
- Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy: Noch einmal. In A. Bergin, & S. Garfield (Hrsg.), *Handbook of Psychotherapy and Behavior Change* (4. Auflage, S. 270–376). Wiley.
- Orlinsky, D., Rønnestad, M. H., & Willutzki, U. (2004). Fifty years of process-outcome research: Continuity and change. In M. Lambert (Hrsg.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (5. Auflage, S. 307–390). Wiley.
- Paine, D. R., Bell, C. A., Sandage, S. J., Rupert, D., Bronstein, M., O'Rourke, C. G., Stavros, G. S., Moon, S. H., & Kehoe, L. E. (2019). Trainee psychotherapy effectiveness at a psychodynamic training clinic: A practice-based study. *Psychoanalytic Psychotherapy*, *33*(1), 20–33. <https://doi.org/10.1080/02668734.2019.1582084>
- Peikert, G., Baum, A., Barchmann, D., Schröder, D., & Kropp, P. (2014). Wie effektiv sind Ausbildungstherapien? Vergleichsstudie zur Effektivität von Ausbildungs- und Regeltherapien. *Verhaltenstherapie*, *24*, 272–279. <https://doi.org/10.1159/000369270>

- Persons, J., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology*, 66, 126–135. <https://doi.org/10.1037//0022-006X.66.1.126>
- Pfammatter, M., & Tschacher, W. (2012). Wirkfaktoren der Psychotherapie – eine Übersicht und Standortbestimmung. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 60(1), 67–76. <https://doi.org/10.1024/1661-4747/a000099>
- Reitinger, C., & Bauer, E. J. (2019). Logotherapy and Existential Analysis. In E. van Deurzen, E. Craig, A. Längle, K. Schneider, D. Tantam, & S. du Plock (Hrsg.), *The Wiley world handbook of Existential Therapy* (S. 324–340). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119167198.ch20>
- Renn, D., Höfer, S., Schüßler, G., Rumpold, G., Smrekar, U., Janecke, N., & Doering, S. (2008). Dimensionale Diagnostik mit dem Fragebogen zur Erfassung von DSM-IV-Persönlichkeitsstörungen (ADP-IV). *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 54(3), 212–226. <https://doi.org/10.13109/zptm.2008.54.3.212>
- Revenstorff, D. (2005). Das Kuckucksei. Über das pharmakologische Modell in der Psychotherapieforschung. *Psychotherapie in Psychiatrie, Psychotherapeutischer Medizin und Klinischer Psychologie*, 10, 22–31. <https://sbt-in-berlin.de/cip-medien/2005-1-04-Revenstorff.pdf>
- Rieß, G., & Löffler-Stastka, H. (2022). VersorgungsNOT – Psychotherapie als zentrale, aber marginalisierte Versorgungsleistung im Gesundheitssystem. *Psychotherapie Forum*, 26, 136–143. <https://doi.org/10.1007/s00729-022-00210-y>
- Riffer, F., Sprung, M., Streibl, L., & Kaiser, E. (2018). Psychotherapeutische Versorgungslage: Aktuelle Daten. *psychopraxis neuropraxis*, 21(3), 128–132. <https://doi.org/10.1007/s00739-018-0468-0>
- Rudolf, G. (1998). Taxonomie der Psychotherapieforschung. *Psychotherapie-Wissenschaft*, 2, Article 2. <https://www.psychotherapie-wissenschaft.info/index.php/psywis/article/view/583>
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57, 1069–1081. <https://doi.org/10.1037/0022-3514.57.6.1069>
- Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69(4), 719–727. <https://doi.org/10.1037/0022-3514.69.4.719>
- Sagerschnig, S., & Valady, S. (2021). *Ausbildungsstatistik 2020. Daten zum Ausbildungsgeschehen in Psychotherapie, Klinischer Psychologie und Gesundheitspsychologie in Österreich*. Gesundheit Österreich GmbH. <https://jasmin.goeg.at/1824/>
- Saß, H. (2003). *Diagnostisches und statistisches Manual psychischer Störungen: Textrevision - DSM-IV-TR: übersetzt nach der Textrevision der vierten Auflage des Diagnostic and statistical manual of mental disorders der American Psychiatric Association*. Hogrefe Verlag für Psychologie.
- Schneider, K. J., & Längle, A. (2012). The renewal of humanism in psychotherapy: Summary and conclusion. *Psychotherapy*, 49(4), 480–481. <https://doi.org/10.1037/a0028026>

- Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist*, 50(12), 965–974. <https://doi.org/10.1037/0003-066X.50.12.965>
- Senft, B., Fischer-Hansal, D., & Schosser, A. (2020). Was heißt signifikant – geheilt oder nur etwas besser? Ein Vergleich verschiedener Berechnungsmethoden zur Bewertung von Veränderungen der depressiven Symptomatik bei ambulanten Rehabilitanden. *Neuropsychiatrie*, 34(3), 130–139. <https://doi.org/10.1007/s40211-020-00343-z>
- Shadish, W., Matt, G., Navarro, A., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, 126. <https://doi.org/10.1037//0033-2909.126.4.512>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin and Company.
- Shean, G. (2014). Limitations of randomized control designs in psychotherapy research. *Advances in Psychiatry*, 2014, e561452. <https://doi.org/10.1155/2014/561452>
- Steffanowski, A., Kramer, D., Fembacher, A., Glahn, E. M., Bruckmayer, E., von Heymann, F., Kriz, D., Pfaffinger, I., & Wittmann, W. W. (2011). Praxisübergreifende Dokumentation der Ergebnisqualität ambulanter Psychotherapie in Bayern. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 40(4), 267–282. <https://doi.org/10.1026/1616-3443/a000124>
- Stieglitz, R.-D., & Hiller, W. (2014). Strategien und Instrumente der Veränderungsmessung. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 62(2), 101–111. <https://doi.org/10.1024/1661-4747/a000187>
- Strauss, B. (2019). Innovative Psychotherapieforschung: Wo stehen wir und wo wollen wir hin? *Psychotherapeuten Journal* 18, 4–10. [https://www.psychotherapeutenjournal.de/ptk/web.nsf/gfx/BC8300C371127969C12583BE0038A8B4/\\$file/PTJ\\_2019-1\\_online.pdf](https://www.psychotherapeutenjournal.de/ptk/web.nsf/gfx/BC8300C371127969C12583BE0038A8B4/$file/PTJ_2019-1_online.pdf)
- Strauss, B. M., Lutz, W., Steffanowski, A., Wittmann, W. W., Boehnke, J. R., Rubel, J., Scheidt, C. E., Caspar, F., Vogel, H., Altmann, U., Steyer, R., Zimmermann, A., Bruckmayer, E., Heymann, F. von, Kramer, D., & Kirchmann, H. (2015). Benefits and challenges in practice-oriented psychotherapy research in Germany: The TK and the QS-PSY-BAY projects of quality assurance in outpatient psychotherapy. *Psychotherapy Research*, 25(1), 32–51. <https://doi.org/10.1080/10503307.2013.856046>
- Stumm, G. (2017). *Geschichte, Paradigmen und Methoden der Psychotherapie* In Psychotherapie - eine Einführung ; Thomas Slunecko (Hg.), facultas. Wien, 2017, S. 33-90
- Tanios, A., Valady, S., & Grabenhofer-Eggerth, A. (2020). *Analyse der Versorgungswirksamkeit von Psychotherapeutinnen und Psychotherapeuten in freier Praxis*. Gesundheit Österreich GmbH. <https://jasmin.goeg.at/1499/>
- Taubner, S., Klasen, J., & Munder, T. (2016). Why do psychotherapists participate in psychotherapy research and why not? Results of the Attitudes to Psychotherapy Research Questionnaire with a

- sample of experienced German psychotherapists. *Psychotherapy Research*, 26(3), 318–331.  
<https://doi.org/10.1080/10503307.2014.938256>
- Thir, M., & Batthyány, A. (2016). The state of empirical research on Logotherapy and Existential Analysis. In A. Batthyány (Hrsg.), *Logotherapy and Existential Analysis* (S. 53–74). Springer International Publishing. [https://doi.org/10.1007/978-3-319-29424-7\\_7](https://doi.org/10.1007/978-3-319-29424-7_7)
- Tickle-Degnen, L. (2013). Nuts and bolts of conducting feasibility studies. *The American Journal of Occupational Therapy : Official Publication of the American Occupational Therapy Association*, 67(2), 171–176. <https://doi.org/10.5014/ajot.2013.006270>
- Tschuschke, V., Cramer, A., Koemeda, M., Schulthess, P., & von Wyl, A. (2015). *Abschlussbericht. Praxisstudie ambulante Psychotherapie Schweiz (PAP-S)*. Schweizer Charta für Psychotherapie. <https://psychotherapie.ch/wsp/site/assets/files/9723/bd-4-2015-gesamt-abschlussbericht-final.pdf>
- Tschuschke, V., & Freyberger, H. J. (2015). Zur aktuellen Situation der Psychotherapiewissenschaft und ihrer Auswirkungen-Eine kritische Analyse der Lage? *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 61(2), 122–138. <https://doi.org/10.13109/zptm.2015.61.2.122>
- Vaughan, S., Marshall, R., MacKinnon, R., & Roose, S. (1994). Current psychotherapy research methodology applied to psychoanalysis: A feasibility study. *The Journal of Psychotherapy Practice and Research*, 3, 334–340.  
[https://www.researchgate.net/publication/225375007\\_Current\\_pschotherapy\\_research\\_methodology\\_applied\\_to\\_pschoanalysis\\_a\\_feasibility\\_study](https://www.researchgate.net/publication/225375007_Current_pschotherapy_research_methodology_applied_to_pschoanalysis_a_feasibility_study)
- Victor, P., Dresenkamp, A., Haag, E., Merod, R., Ruggaber, G., Sauer, K., Schneider, W., Ruppert, J., Veith, A., & Willutzki, U. (2018). Ausbildungsforschung in Ausbildungsinstituten der Deutschen Gesellschaft für Verhaltenstherapie. *Psychotherapeut*, 63(1), 62–67. <https://doi.org/10.1007/s00278-017-0225-5>
- Vos, J., Cooper, M., Correia, E., & Craig, M. (2015). Existential therapies: A review of their scientific foundations and efficacy. *Existential Analysis*, 26(1), 49–69. <https://psycnet-apa-org.uaccess.univie.ac.at/record/2015-05162-004>
- Vos, J., & Vitali, D. (2018). The effects of psychological meaning-centered therapies on quality of life and psychological stress: A metaanalysis. *Palliative & Supportive Care*, 16(5), 608–632.  
<https://doi.org/10.1017/S1478951517000931>
- Wampold, B. (2021). Evaluation: Methodenpapier des Wissenschaftlichen Beirats Psychotherapie. [https://bvvp.de/wp-content/uploads/2021/07/Psychotherapie\\_Wampold.-Kritik-am-Methodenp.-Uebersetzg\\_public.pdf](https://bvvp.de/wp-content/uploads/2021/07/Psychotherapie_Wampold.-Kritik-am-Methodenp.-Uebersetzg_public.pdf)
- Whitehead, A. L., Sully, B. G. O., & Campbell, M. J. (2014). Pilot and feasibility studies: Is there a difference from each other and from a randomised controlled trial? *Contemporary Clinical Trials*, 38(1), 130–133. <https://doi.org/10.1016/j.cct.2014.04.001>
- Wilmers, F., Munder, T., Leonhart, R., Herzog, T. N., Plassmann, R., Barth, J., & Linster, H. W. (2008). Die deutschsprachige Version des Working Alliance Inventory-Short revised (WAI-SR)-Ein

schulübergreifendes, ökonomisches und empirisch validiertes Instrument zur Erfassung der therapeutischen Allianz. *Klinische Diagnostik und Evaluation*, 1(3), 343-58.

<https://doi.org/10.7892/BORIS.27962>

Wittmann, W., Lutz, W., Steffanowski, A., Kriz, D., Glahn, E., Völkle, M., Böhnke, J., Köck, K., Bittermann, A., & Ruprecht, T. (2011). Qualitätsmonitoring in der ambulanten Psychotherapie: Modellprojekt der Techniker Krankenkasse-Abschlussbericht

Techniker Krankenkasse. Hamburg. [https://www.bptk.de/wp-content/uploads/2019/01/TK-Abschlussbericht\\_Qualitaetsmonitoring-in-der-ambulanten-Psychotherapie.pdf](https://www.bptk.de/wp-content/uploads/2019/01/TK-Abschlussbericht_Qualitaetsmonitoring-in-der-ambulanten-Psychotherapie.pdf)

Wittmann, W., & Steffanowski, A. (2011). Qualitätsmonitoring in der ambulanten Psychotherapie: Ergebnisse des TK-Modellprojektes. *Psychotherapie Aktuell*, 3(3), 6-12.

[https://www.researchgate.net/publication/254560582\\_Wittmann\\_W\\_W\\_Steffanowski\\_A\\_2011\\_Qualitaetsmonitoring\\_in\\_der\\_ambulanten\\_Psychotherapie\\_Ergebnisse\\_des\\_TK-Modellprojektes\\_Psychotherapie\\_Aktuell\\_33\\_6-12](https://www.researchgate.net/publication/254560582_Wittmann_W_W_Steffanowski_A_2011_Qualitaetsmonitoring_in_der_ambulanten_Psychotherapie_Ergebnisse_des_TK-Modellprojektes_Psychotherapie_Aktuell_33_6-12)

Wong, P. T. P. (2017). Meaning-centered approach to research and therapy, second wave positive psychology, and the future of humanistic psychology. *The Humanistic Psychologist*, 45(3), 207–216.

<https://doi.org/10.1037/hum0000062>

World Medical Association. (2013). World Medical Association declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 310(20), 2191–2194.

<https://doi.org/10.1001/jama.2013.281053>

von Wyl, A. von, Cramer, A., Koemeda, M., Tschuschke, V., & Schulthess, P. (2013). Praxisstudie ambulante Psychotherapie Schweiz (PAP-S): Studiendesign und Machbarkeit. *Psychotherapie-Wissenschaft*, 3(1). <https://psychotherapie-wissenschaft.info/article/view/179>

## Anhang

### Abkürzungsverzeichnis

RCT .....	Randomisiert-kontrollierte Studie (randomized-controlled trial)
GLE-Ö .....	Gesellschaft für Logotherapie und Existenzanalyse
SCL .....	Symptom-Checklist-90-Standard (SCL-90-S)
GSI .....	Global Severity Index (Globalmaß der Symptombelastung der SCL-90-S)
DEPR .....	Depressionsskala der SCL-90-S
ANGS .....	Angstskala der SCL-90
IIP .....	Inventar zur Erfassung interpersonaler Probleme (IIP-D)
ESK .....	Gesamtskala der Existenz-Skala
RYFF .....	Gesamtskala der Ryff-Skala
GAF .....	Global Assessment of Functioning
ADP-IV .....	Assessment of DSM-IV Personality Disorders
ICD .....	International Classification of Diseases
IDCL .....	Internationale Diagnosen Checklisten für ICD-10
i.A.u.S. ....	in Ausbildung unter Supervision (Psychotherapeut*innen)
LOCF .....	Last Observation Carried Forward (Imputationsmethode)
RCI .....	Reliable Change Index

### Abbildungsverzeichnis

<b>Abbildung 1</b> Studiendesign vorliegender Arbeit .....	40
<b>Abbildung 2</b> Gezogene Teilstichproben für die Analyse .....	48
<b>Abbildung 3</b> Studienteilnahmegrund der Klient*innen zu Therapiebeginn .....	60
<b>Abbildung 4</b> Beurteilung des Therapieerfolges durch die Klient*innen und deren Therapeut*innen zu Therapieende .....	72
<b>Abbildung 5</b> Übereinstimmung der Behandlungspaare in den Beurteilungen zur Therapiezufriedenheit .....	73
<b>Abbildung 6</b> Beurteilung der Zielerreichung durch die Klient*innen und deren Therapeut*innen zu Therapieende .....	74
<b>Abbildung 7</b> Übereinstimmung der Behandlungspaare in den Beurteilungen zum Erreichen der Therapieziele .....	78
<b>Abbildung 8</b> Effektstärken in den einzelnen Verfahren über die Testzeitpunkte hinweg	
<b>Abbildung 7</b> Zusammenfassung der Ergebnisse .....	103

## **Tabellenverzeichnis**

<b>Tabelle 1</b> <i>Wirkung von Psychotherapie und Einfluss von kontextuellen bzw. spezifischen Faktoren (Strauss, 2019, S. 6).....</i>	13
<b>Tabelle 2</b> <i>Cut-Off-Werte für klinische Auffälligkeit der symptombezogenen Skalen.....</i>	52
<b>Tabelle 3</b> <i>Rücklauf der Klient*innen .....</i>	57
<b>Tabelle 4</b> <i>Vorhandene Informationen zu Abbruchsfällen .....</i>	59
<b>Tabelle 5</b> <i>Kennwerte und interne Konsistenzen der Skalen .....</i>	62
<b>Tabelle 6</b> <i>Soziodemographische Daten der Klient*innen zu Therapiebeginn .....</i>	63
<b>Tabelle 7</b> <i>Gruppenmittelwerte auf den Skalen zu Therapiebeginn .....</i>	64
<b>Tabelle 8</b> <i>Personen mit klinisch (un-)auffälliger Symptomatik zu Therapiebeginn .....</i>	65
<b>Tabelle 9</b> <i>Hauptdiagnosen zu Therapiebeginn .....</i>	66
<b>Tabelle 10</b> <i>Anzahl an Diagnosen pro Person zu Therapiebeginn .....</i>	66
<b>Tabelle 11</b> <i>Verdacht auf Persönlichkeitsstörungen zu Therapiebeginn .....</i>	67
<b>Tabelle 12</b> <i>Dauer der Psychotherapie in Monaten und Anzahl der Einheiten .....</i>	68
<b>Tabelle 13</b> <i>Ergebnisse des gepaarten t-Tests in der Prä-Post-Stichprobe .....</i>	70
<b>Tabelle 14</b> <i>Ergebnisse des RCI bei den Personen mit zu Therapiebeginn klinisch auffälliger Symptomausprägung .....</i>	71
<b>Tabelle 15</b> <i>Ergebnisse des RCI bei den Personen mit zu Therapiebeginn klinisch unauffälliger Symptomausprägung .....</i>	72
<b>Tabelle 16</b> <i>Ergebnisse des gepaarten t-Tests in der Verlauf-Stichprobe.....</i>	76
<b>Tabelle 17</b> <i>Ergebnisse des gepaarten t-Tests in der Katamnese-Stichprobe.....</i>	78
<b>Tabelle 18</b> <i>Ergebnisse des explorativen gepaarten t-Tests in der Completer- bzw. ITT-Stichprobe .....</i>	79
<b>Tabelle 19</b> <i>Korrelativer Zusammenhang der ESK-Gesamtskala mit den anderen Skalen der Studie .....</i>	80
<b>Tabelle 20</b> <i>Naturalistische Vergleichsstudien .....</i>	90

## Zusammenfassung

Naturalistische Psychotherapiestudien ermöglichen neben einer Überprüfung der praktischen Wirksamkeit von Psychotherapie unter Routinebedingungen auch eine Qualitätssicherung der täglichen Versorgung. Während bisher insbesondere die Wirksamkeit von Behandlungen ausgebildeter Psychotherapeut\*innen erforscht wurde, fanden - trotz deren Versorgungsrelevanz - Untersuchungen bei Ausbildungstherapien in freier Praxis kaum Eingang in die evaluative Wirksamkeitsforschung. Vorliegende Arbeit soll einen Beitrag zur Erforschung der Behandlungsqualität von existenzanalytischen Psychotherapeut\*innen in Ausbildung unter Supervision (i.A.u.S.) in freier Praxis leisten.

Im Rahmen eines prospektiven Ein-Gruppen-Studiendesigns wurden die Symptombelastung, interpersonale Probleme, das psychologische Wohlbefinden und die existenzielle Erfüllung der Klient\*innen zu Therapiebeginn, optional nach 25 Einheiten, zu Therapieende und sechs Monate nach Therapieende erfasst. Die Therapeut\*innen konnten auf Wunsch, regelmäßige Rückmeldungen zu den Fragebogenergebnissen ihrer Klient\*innen erhalten.

Es konnten bisher Daten von 31 existenzanalytischen Psychotherapeut\*innen i.A.u.S. in freier Praxis und 263 Klient\*innen gesammelt werden. Die Studienbedingungen wurden von den teilgenommenen Therapeut\*innen sowie den Klient\*innen gut angenommen. In der Prä-Post-Stichprobe (*available-case*,  $n_{KL} = 79$ ;  $n_{TH} = 23$ ) wurde eine - über alle Instrumente hinweg - durchschnittliche Effektstärke von Cohen's  $d = 0.95$  ermittelt. Von den Personen mit zu Therapiebeginn in den symptombezogenen Verfahren klinisch auffälligen Werten ( $n = 53$ ) wiesen nach dem *RCI* durchschnittlich 75 % eine reliable Verbesserung (*response*) auf. Bei durchschnittlich 60 % konnte zudem eine klinisch signifikante Veränderung (*remission*) festgestellt werden. Rund 80 % der Klient\*innen der Prä-Post-Stichprobe gaben zu Behandlungsende eine sehr hohe Therapiezufriedenheit an.

Das Untersuchungsdesign der Pilotstudie war in den therapeutischen Alltag integrierbar. Sowohl die gruppenstatistischen als auch individuellen Veränderungsmaße und die Zufriedenheitswerte der Klient\*innen sprechen für eine hohe Ergebnisqualität der Existenzanalyse bei Psychotherapeut\*innen i.A.u.S in freier Praxis. Die Befunde sind mit den Ergebnissen anderer praxisorientierter Studien aus dem deutschsprachigen Raum vergleichbar. Das naturalistische Studiendesign ermöglichte eine Evaluierung der täglichen Versorgung in freier Praxis. Zu Therapieende konnten nur wenig Datensätze von Abbruchfällen erhoben werden. In naturalistischen Ein-Gruppen-Studiendesigns könnte diese Gruppe als Vergleichsgruppe herangezogen werden. Zukünftige praxisorientierte Studien zur Wirksamkeit von (Ausbildungs-) Therapien sollten den Einfluss von Rückmeldungen auf das Therapieergebnis untersuchen.

*Schlagerwörter:* naturalistische Psychotherapieforschung, Psychotherapeut\*innen in Ausbildung unter Supervision in freier Praxis, Logotherapie und Existenzanalyse, statistische und klinische Signifikanz, Studienakzeptanz und -machbarkeit

## Abstract

Naturalistic psychotherapy studies enable not only an examination of the practical effectiveness of psychotherapy under routine conditions but also a quality assurance of daily care. Up to now, especially the effectiveness of treatments of trained psychotherapists has been researched, but despite their relevance for health care, studies of trainee-therapists especially in private practice have hardly found their way into evaluative effectiveness research. The present study aims to contribute to research on the quality of treatment of existential analytic psychotherapists in training under supervision in private practice.

Within the framework of a prospective one-group study design, the psychological symptoms, interpersonal problems, psychological well-being, and existential fulfillment of the clients were assessed with questionnaires at the beginning of therapy, optionally after 25 units, at the end of therapy, and six months after the end of therapy. The therapists had the opportunity to receive regular feedback on the questionnaire results of their clients.

So far, data of 31 existential analytic psychotherapists in private practice and 263 clients could be collected. The study conditions were well accepted by the participating therapists as well as the clients. In the pre-post sample (*available-case*,  $n_{KL} = 79$ ;  $n_{TH} = 23$ ), an average effect size of Cohen's  $d = 0.95$  was determined. According to the *RCI*, an average of 75% of the people with clinically abnormal values in the symptom-related instruments at the beginning of therapy ( $n = 53$ ) showed a reliable change (*response*) and an average of 60% also showed a clinically significant change (*remission*). About 80% of clients in the pre-post sample reported a high treatment satisfaction at the end of treatment.

The research design of the pilot study was integrated without problems into everyday therapeutic routine. Both, the group-statistical and individual measures of change and the satisfaction-scores of the clients indicate a high outcome quality of Personal Existential Analysis among psychotherapist trainees in private practice. The findings are comparable with the results of other practice-oriented studies from the German-speaking countries. The naturalistic study design enabled an evaluation of daily care in private practice.

At the end of therapy, only few data sets of dropout-cases could be collected. In naturalistic single-group study designs, this group could be used as a comparison group.

Future practice-based studies on the effectiveness of (trainee's) therapies should investigate the influence of feedback on the therapy outcome.

*Keywords:* naturalistic psychotherapy research, psychotherapists in training under supervision in private practice, logotherapy and existential analysis, statistical and clinical significance, study feasibility and acceptability