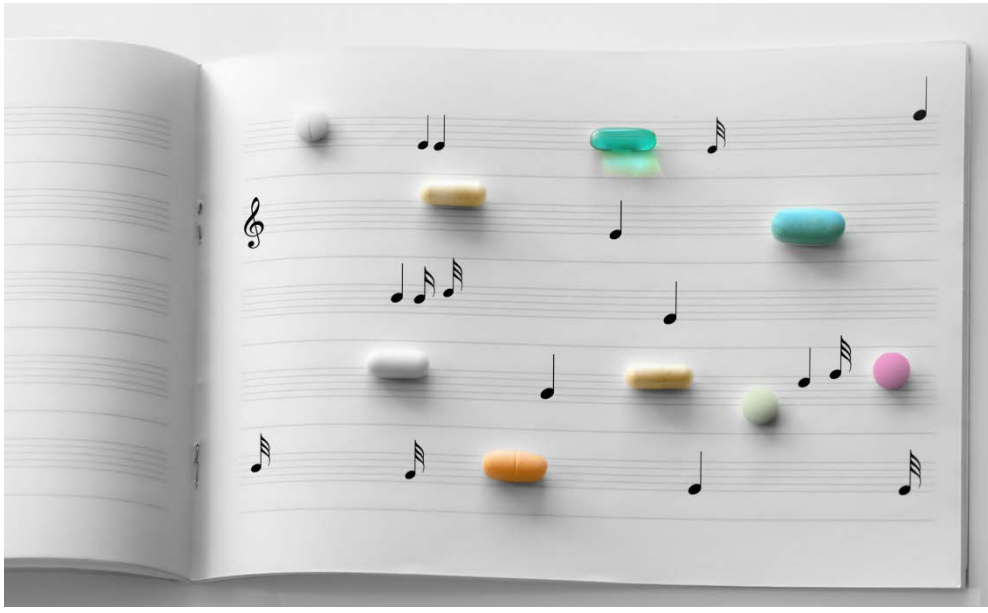
HTA Austria
Austrian Institute for
Health Technology Assessment
GmbH

Music therapy: Clinical and socioeconomic outcomes and long-term measures in substance use disorder



A systematic review and service user interviews

Final report

AIHTA Project Report No.: 178 | ISSN: 1993-0488 | ISSN-online: 1993-0496

HTA Austria
Austrian Institute for
Health Technology Assessment
GmbH

Music therapy: Clinical and socioeconomic outcomes and long-term measures in substance use disorder

A systematic review and service user interviews

Vienna, March 2026

Project Team

Project leader: Mag. Dr. Lucia Gassner

Authors: Mag. Dr. Lucia Gassner (Austrian Institute for Health Technology Assessment GmbH, Austria),
Daria Makurat, MA, Ilona Bidzan-Bluma, PhD and Dr. habil. Assoc. Prof. Łucja Bieleninik
(Institute of Psychology, University of Gdańsk, Poland),
Thomas Solgård Svendsen, PhD MSW, Dr. Assoc. Prof. Bianka Karshikoff and Prof. Aleksander Hagen Erga, PhD
(Centre for alcohol and drug research (KORFOR), Stavanger University Hospital, Norway),
Franziska Magel, MSc. MA (Anton Proksch Institute, Austria),

Project Support

Systematic literature search: Tarquin Mittermayr, MA

Hand search: Mag. Dr. Lucia Gassner; Daria Makurat, MA

Internal review: Dr. rer. soc. oec. Ingrid Zechmeister-Koss, MA

External review: Prof. Dr. Josephine Geipel, Head of the Music Therapy Department, Leopold Mozart College of Music, Faculty of Philosophy and Social Sciences, University of Augsburg, Germany

Correspondence: Mag. Dr. Lucia Gassner, lucia.gassner@gmail.com

Cover photo: @ CURAphotography – stock.adobe.com

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement number 101155881.

This report should be referenced as follows:

Gassner L, Makurat D, Svendsen TS, Bidzan-Bluma I, Magel F, Karshikoff B, Erga AH, Bieleninik L. Music therapy: Clinical and socioeconomic outcomes and long-term measures in substance use disorder. A systematic review and service user interviews. AIHTA Project Report No.: 178 2026. Vienna: HTA Austria – Austrian Institute for Health Technology Assessment GmbH.

Conflict of interest

All authors and the reviewers involved in the production of this report have declared they have no conflicts of interest in relation to the technology assessed according to the Uniform Requirements of Manuscripts Statement of Medical Journal Editors (www.icmje.org).

Disclaimer

The external reviewers did not co-author the scientific report and do not necessarily all agree with its content. Only the AIHTA is responsible for errors or omissions that could persist. The final version and the policy recommendations are under the full responsibility of the AIHTA.

IMPRINT

Publisher:
HTA Austria – Austrian Institute for Health Technology Assessment GmbH
Josefstädter Straße 39 | 1080 Vienna – Austria
https://www.aihta.at/

Responsible for content:
Dr. rer. soc. oec. Ingrid Zechmeister-Koss, MA, managing director

AIHTA Project Reports do not appear on a regular basis and serve to publicise the research results of the Austrian Institute for Health Technology Assessment.

AIHTA Project Reports are only available to the public via the Internet at http://eprints.aihta.at/view/types/hta_report.html.

AIHTA Project Report No.: 178

ISSN 1993-0488

ISSN online 1993-0496

Content

List of figures

List of tables

List of abbreviations

A-COPE .....Adolescent Coping Orientation to Problems Experienced

ACQ-SF-R .....Alcohol craving questionnaire Revised

AIHTA .....Austrian Institute for Health Technology Assessment GmbH

Alcohol-E .....Alcohol Use Disorders Identification Test Extended Version

ALQoL .....Alcohol Quality of Life

ANCI.....American Indian/Alaska Native Cultural Identity Scale

AQoLS.....Alcohol Quality of Life Scale

ARSW .....Adjective Rating Scale for Withdrawal

ASI.....Addiction Severity Index

ATQ.....Automatic Thoughts Questionnaire

AUDIT.....Alcohol Use Disorders Identification Test

AUDIT-C.....Alcohol Use Disorders Identification Test – Consumption

AUQ .....Alcohol urge questionnaire

AWMF .....Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften

BDI.....Beck Depression Inventory

BSCS .....Brief Substance Craving Scale

BSI.....Brief Symptom Inventory

CAPSES.....Composite SocioEconomic Status scale

CESI.....Client Evaluation of Self at Intake

CGI.....Clinical Global Impression Severity Scale

CMR.....Circumstances, Motivation, and Readiness Scales

COMET .....Core Outcome Measures in Effectiveness Trials

COREQ.....COnsolidated criteria for REporting Qualitative research

COS .....Core outcome sets

COSMIN.....COnsensus-based Standards for the selection of health Measurement INstruments

CSRI.....Client service receipt inventory

CSS.....Commitment to Sobriety Scale

CSSRI .....Client Socio-Demographic and Service Receipt Inventory

DAQ.....Desires for Alcohol Questionnaire

DAS.....Disability Assessment Schedule

DASES.....Drug Avoidance Self-Efficacy Scale

DMQ-R.....Drinking Motives Questionnaire – Revised

DRIE.....Drinking-Related Internal-External Locus of Control Scale

DUDIT .....Drug Use Disorders Identification Test

DUDIT-E .....Drug Use Disorders Identification Test Extended

DUQoL .....Drug User Quality of Life Scale

ESDS.....European Socio-demographic Schedule

EuroADAD.....European Adolescent Assessment Dialogue

EuropASI.....European Addiction Severity Index

FACIT.....Functional Assessment of Chronic Illness Therapy

FACIT-F.....Functional Assessment of Chronic Illness Therapy: Fatigue

FACT-Cog.....Functional Assessment of Cancer Therapy – Cognitive Functions

FALCO .....Fighting Addictions, improving Lives: COmprehensive drug rehabilitation with music

FGPS.....Forgiveness Grief Perspectives Scale

FU .....Follow-up

GAATOR.....General Alcoholics Anonymous Tools of Recovery

GAF.....Global Assessment of Function Scale

GAF.....Goal attainment form

GEMS.....Geneva Emotions in Music Scale

GH.....Global health

HADEA .....Health and Digital Executive Agency

HAQ.....Helping Alliance Questionnaire for therapist and client

HER .....Electronic health records

HIS .....Heaviness of Smoking Index

HPQ .....Health and Work Performance Questionnaire
HRSD.....Hamilton Rating Scale for Depression
HTA .....Health technology assessment
ICHOM.....International Consortium for Health Outcomes Measurement
ICR.....Importance, Confidence, Readiness (motivational) Ruler
I-E.....Internal vs. external locus of control scale
IIP-SC .....Inventory of Interpersonal Problems: Short Circumplex form
INAHTA.....International Network of Agencies for Health Technology Assessment
LSD .....Lysergic acid diethylamide
MCID.....Minimum clinically important difference
MDMQ.....Multidimensional Mood Questionnaire
MSPSS .....Multidimensional Scale of Perceived Social Support
MT.....Music therapy
NAI.....Novaco Anger Inventory
NAV.....Native American Version
NHS.....National Health Service
NR .....Not reported
NS-SEC.....National Statistics Socioeconomic classification
NTA.....National Treatment Agency for Substance Misuse
OCDUS.....Obsessive Compulsive Drug Use Scale
OSS.....Occupational Status Score
PANAS.....Positive and negative affect scale
PECUNIA.....ProgrammE in Costing, resource use measurement and outcome evaluatioN for Use in multi-sectoral National and InternAtional health economic evaluations
PIL .....Purpose in Life Test
PROMIS .....Patient-Reported Outcomes Measurement Information System
PSAS .....Perceived Stigma of Addiction Scale
Q-LESQ .....Quality of Life Enjoyment and Satisfaction Questionnaire
QLI.....Quality of Life Interview
QMAD .....Questionnaire of Motivation for Abstaining from Drugs
QoL .....Quality of life
RCT.....Randomised controlled trial
RTCQ-TV.....Readiness to Change Questionnaire – Treatment Version
RUM .....Resource Use Measurement
SAS.....Self-Rating Anxiety Scale
SC.....Symptom Checklist
SCQ.....Modified Self-Administered Comorbidity Questionnaire
SCS.....Self-Consciousness Scale
SDS .....Severity of Dependence Scale
SEI.....Socioeconomic Index
SES .....Socioeconomic status
SF .....Short Form
S-IgA .....Secretory Immunglobulin A
SIgAD .....Selective Immunglobulin A Deficiency
SOC.....Sense of Coherence Scale
SOCRATES.....Stages of Change Readiness and Treatment Eagerness Scale
SSGS .....State Shame and Guilt Scale
STAI.....State-Trait Anxiety Inventory
SUD .....Substance use disorder
SURE .....Substance Use Recovery Evaluator
TOP.....Treatment Outcomes Profile for Substance Misuse
TRIP .....Turning Research Into Practice
URICA.....University of Rhode Island Change Assessment Scale
VAMS.....Visual Analog Mood Scale
VAS .....Visual Analog Scale
WAI-S(R).....Working Alliance Inventory – Short (Revised)
WFPTS .....Wake Forest Physician Trust Scale
WHO.....World Health Organization
WHOQOL .....World Health Organization Quality of Life

Executive Summary

Background

Substance use disorder (SUD) carries a substantial global burden of disease, accounting for 1.3% of disability-adjusted life years attributable to illicit drugs and 4.2% to alcohol. Music therapy (MT) has demonstrated promise as a supplementary intervention: music engages the brain's reward system analogously to addictive substances while fostering positive reinforcement, stress reduction, and emotional regulation. However, music may also trigger negative memories or substance craving in individuals with SUD, necessitating specifically trained music therapists. Given documented high relapse rates in the first year post-treatment, investigating MT's long-term effectiveness and identifying appropriate outcome measurement instruments are of critical importance.

Objectives

Conducted within the framework of the European Horizon Europe project FALCO (Fighting Addictions, improving Lives: COmprehensive drug rehabilitation with music), this report pursued two primary research questions: first, to identify clinical and socioeconomic outcome parameters and tools assessing (serious) adverse events in existing MT evidence for SUD, including how outcomes are measured and whether instruments have been tested for validity and reliability; and second, to characterise the most frequently used, validated, and reliable measuring instruments, with particular attention to their appropriateness, feasibility, and suitability for long-term measurement in patients with SUD. Service user interviews were conducted to complement findings from a patient-relevant perspective.

Methods

A systematic literature search was conducted in February 2025 across four databases (The Cochrane Library, Embase, INAHTA, MEDLINE), limited to reviews published from inception through 2024. Following deduplication and independent dual screening of 451 citations, seven reviews were included for data extraction. Based on predefined selection criteria – including frequency of use, validated and reliable psychometric properties, and coverage of all identified outcome parameters – 14 measuring instruments were selected for detailed characterisation. A complementary manual search identified instruments capturing socioeconomic outcomes. Nine semi-structured interviews with service users from Austria, Norway, and Poland were analysed using AI-supported qualitative content analysis (deductive-inductive approach, manifest and latent levels) to incorporate patient-relevant perspectives into instrument evaluation.

Results

Seven included reviews yielded more than 60 outcomes, categorised into 11 parameters: substance use, recovery, craving, motivation, quality of life, mood, psychological/psychiatric/physiological/cognitive, cultural/spiritual/locus of control, biomedical, music-/therapist-/treatment-related, and socioeconomic and social parameters. No tools for assessing (serious) adverse events were reported. The manual search identified 13 additional tools measuring approximately 20 socioeconomic outcomes.

A Minimal Clinically Important Difference (MCID) was established only for the Beck's Depression Inventory (BDI), and not specifically for SUD populations. Ten instruments required ten minutes or fewer to administer; 11 were self-reporting tools, and four were available in more than ten languages. Key limitations included self-report bias, challenges with cultural adaptation, training requirements for clinician-administered tools, and limited empirical evidence for long-term use – with the longest documented treatment duration being two years and the most recent follow-up assessments reaching only three months. Instruments measuring stable constructs – depression (BDI), problem severity (ASI), self-efficacy (DASES), psychiatric symptoms (SCL-90-R), sense of coherence (SOC), locus of control (DRIE), and interpersonal problems (IIP-SC) – were identified as most appropriate for long-term measurement.

Service user interviews (4 women, 5 men, aged 25-63), all in outpatient therapy without MT and with at least one year of abstinence, revealed social parameters as by far the most frequently mentioned aspects, cautiously suggesting particular user-relevance, followed by socioeconomic and treatment-related parameters. Family, social environment, and motivation for treatment were most salient. Qualitative findings demonstrated that substance use progressively disrupted multiple life domains, recovery required sustained and often repeated support, and social connection and meaningful activities were central to maintaining abstinence. Latent content analysis revealed a fundamental mismatch between standardised instruments and the lived experience of social recovery: service users articulated social destruction in vivid detail but struggled to describe social gains in comparable terms. This asymmetry suggests that standardised instruments – which rely on respondents' ability to recognise and verbalise positive change – may systematically underrepresent progress in social recovery. Four recurring narrative patterns were identified: social exclusion framed as personal moral failure, substance use as an inescapable cycle, a desire for a protected social fresh start, and life as a constant struggle. Notably, children functioned as an internal disciplinary force, pets as a bridge out of isolation, and paid employment as a prerequisite for social belonging.

Conclusions

Validated and reliable instruments exist for measuring MT outcomes in SUD; however, several critical gaps remain. MCIDs are absent for all but one instrument and have not been established specifically for SUD populations. Empirical evidence for long-term application is limited. No standardised protocols for adverse event monitoring were identified, representing a significant methodological deficit. Qualitative data indicate that conventional measurement approaches underrepresent dimensions central to recovery, particularly social reintegration, meaningful activity engagement, and vocational participation.

Future research should prioritise population-specific MCIDs, extended study durations exceeding six months with a minimum 12-month follow-up, validated adverse event monitoring tools, and consensus-based core outcome sets to improve cross-study comparability. Given the nascent state of MT research methodology, this review deliberately extended its scope beyond MT-specific instruments to include validated tools from mental health research; accordingly, outcome domains should be expanded to encompass socioeconomic and broader social parameters – including community belonging and friendship formation – not currently captured by available instruments but central to people with lived experience of SUD, complemented by the development of culturally adapted instruments for diverse SUD populations. A participatory, mixed-methods research design is recommended to align measurement practices with service users' priorities and recovery trajectories.

Zusammenfassung

Hintergrund und Zielsetzung

Problematischer Substanzkonsum (SUD, substance use disorder), zählt weltweit zu den größten Gesundheitsproblemen. In der EU sind allein eine Million Menschen mit problematischem Opioidkonsum registriert. Diese Suchterkrankung verursacht erhebliches individuelles Leid und belastet zugleich das Gesundheitssystem sowie die Gesellschaft insgesamt. SUD tritt häufig in Kombination mit anderen psychischen oder körperlichen Erkrankungen auf und ist eng mit sozialen Problemlagen wie Armut, Arbeitslosigkeit und Kriminalität verknüpft. Viele Betroffene brechen ihre Behandlung ab oder sprechen nicht ausreichend auf gängige Therapien an.

Musiktherapie (MT) gewinnt als ergänzende Behandlungsmethode zunehmend an Bedeutung. Der Grund liegt unter anderem darin, dass Musik ähnliche Hirnareale aktiviert wie Suchtmittel. Sie fördert positive Gefühle, hilft beim Stressabbau und stärkt die Motivation zur Verhaltensänderung. Gleichzeitig kann Musik bei Personen mit SUD auch negative Erinnerungen oder Substanzverlangen auslösen, weshalb speziell ausgebildete Musiktherapeut:innen für den Einsatz notwendig sind.

Da die Rückfallrate im ersten Jahr nach einer SUD-Behandlung generell besonders hoch ist und längere Behandlungsdauern mit besseren Langzeitergebnissen verbunden sind, ist es wichtig zu verstehen, welche Messinstrumente geeignet sind, die Effekte von MT langfristig zu erfassen – also über ein Jahr nach Behandlungsbeginn hinaus. Bislang fehlt jedoch empirische Evidenz dafür, dass die vorhandenen Instrumente auch für diesen Langzeiteinsatz systematisch erprobt wurden.

Dieser Bericht ist Teil des europäischen Horizon-Europe-Projekts FALCO (Fighting Addictions, improving Lives: COmprehensive drug rehabilitation with music), an dem sieben Länder beteiligt sind. FALCO ist die erste große randomisierte, kontrollierte Studie, die die Langzeitwirkungen von MT bei SUD untersucht. Bei dieser klinischen Untersuchung werden Teilnehmende zufällig zu Gruppen zugeteilt, wobei nur eine Gruppe MT erhält, um zu vergleichen, welche Wirkung MT hat.

Ziel dieses Berichts ist es, geeignete klinische und sozioökonomische Messinstrumente für die Langzeitforschung zu identifizieren und die Perspektiven von Betroffenen einzubeziehen. Die Studie entstand in Zusammenarbeit mit der Universität Gdańsk (Polen), dem Zentrum für Alkohol- und Drogenforschung (KORFOR) am Universitätskrankenhaus Stavanger (Norwegen) und dem Anton-Proksch-Institut (Österreich).

Methoden

Der Bericht umfasst drei methodische Zugänge. Zunächst wurde im Februar 2025 eine *systematische Literatursuche* in vier internationalen Datenbanken (Cochrane Library, Embase, INAHTA, MEDLINE) durchgeführt. Eingeschlossen wurden systematische Übersichtsarbeiten, also Studien, die viele Einzelstudien zusammenfassen, zu musikbasierten Interventionen bei Erwachsenen mit SUD (ausgenommen reine Nikotinabhängigkeit), ohne sprachliche Einschränkung, bis einschließlich 2024. Die Studien dienten als Basis für die Identifikation von Messinstrumenten. Da die systematische Suche nur wenige sozioökonomische Messinstrumente lieferte, wurde ergänzend eine *manuelle Suche* in weiteren Quellen (z. B. PubMed, Google Scholar) im Bereich der psychischen Gesundheit durchgeführt.

Parallel dazu wurden neun *halbstrukturierte Interviews* (offene, aber gelenkte Fragen) mit Betroffenen in drei Ländern (Österreich, Norwegen, Polen) durchgeführt. Die Teilnehmenden – vier Frauen und fünf Männer im Alter von 25 bis 63 Jahren – befanden sich zum Zeitpunkt des Interviews in ambulanter Therapie (keine MT) und hatten mindestens ein Jahr Abstinenz. Die Interviews wurden von vertrauten Fachkräften wie Psychotherapeut:innen oder Sozialarbeiter:innen geführt, um eine offene und sichere Gesprächssituation zu gewährleisten. Die Auswertung erfolgte mittels qualitativer KI-gestützter Inhaltsanalyse, einer systematischen Auswertung der Interviewinhalte. Alle Daten wurden anonymisiert und datenschutzkonform verarbeitet.

Ergebnisse

Teil 1: Klinische und sozioökonomische Messinstrumente

Aus der systematischen Suche wurden sieben Übersichtsarbeiten eingeschlossen. Insgesamt wurden mehr als 60 messbare Endpunkte identifiziert und in elf Kategorien eingeteilt wie z. B. Substanzkonsum, Genesung, Substanzverlangen, Motivation, Lebensqualität, Stimmung sowie psychologische, soziale und sozioökonomische Themenbereiche. Mehr als die Hälfte der gefundenen Messinstrumente wurde von den jeweiligen Autor:innen auf Zuverlässigkeit und Gültigkeit geprüft. Es wurde kein Instrument zur Erfassung unerwünschter Ereignisse (z. B. schwerwiegende Rückfälle) berichtet.

Sozioökonomische Aspekte wie Arbeit, Einkommen, Wohnsituation oder Bildung waren in der vorhandenen Forschungsliteratur kaum vertreten. Die ergänzende Handsuche identifizierte jedoch weitere 13 Instrumente, die etwa 20 sozioökonomische Endpunkte messen.

Teil 2: Charakteristika der Messinstrumente

Aus der Gesamtheit der gefundenen Messinstrumente wurden 14 häufig verwendete, validierte und zuverlässige (sprich: wissenschaftlich geprüfte) Messinstrumente für eine detaillierte Analyse ausgewählt. Diese decken ein breites Spektrum ab: von Selbstwirksamkeit (DASES), Problemschwere (ASI) und Substanzverlangen (BSCS) über Motivation (URICA, ICR), Lebenssinn (SOC) und Depression (BDI) bis hin zu psychiatrischen Symptomen (SCL-90-R), zwischenmenschlichen Problemen (IIP-SC), Überzeugung, das eigene Leben selbst steuern zu können (DRIE), therapeutischer Beziehung (HAQ-II), sozioökonomischen Daten (PECUNIA-RUM, CSRI) und einem biologischen Maß der Immunfunktion (Speichel-Immunglobulin-A-Test).

Die Eignung dieser 14 Instrumente für den Langzeiteinsatz wurde anhand von drei Kriterien bewertet: Messqualität und klinische Relevanz, praktische Handhabbarkeit sowie Übertragbarkeit auf verschiedene Kontexte.

Sieben Instrumente (DASES, ASI, SOC, BDI, SCL-90-R, DRIE, IIP-SC) messen stabile, langfristig relevante Merkmale wie Depression, Problemschwere oder Selbstwirksamkeit und eignen sich daher gut für Langzeitmessungen. Zehn der 14 Instrumente benötigen höchstens zehn Minuten Ausfüllzeit und sind damit auch für wiederholte Anwendungen praktikabel. Elf Instrumente können von den Betroffenen selbst ausgefüllt werden. Vier Instrumente (ASI, ICR, SOC, BDI) stehen in mehr als zehn Sprachen zur Verfügung, was den internationalen Einsatz erleichtert. Die empirische Evidenz, also tatsächliche Belege aus der Forschungspraxis, für den Langzeiteinsatz ist allerdings begrenzt: Die längste dokumentierte Behandlungsdauer beträgt zwei Jahre mit zehn MT-Sitzungen, die längste Nachbeobachtungszeit beträgt drei Monate. Nur für den BDI existiert ein klinisch bedeutsamer Schwellenwert (MCID, Minimal Clinically Important Difference), der angibt, ab welcher Veränderung eine Verbesserung als klinisch bedeutsam gilt – und auch dieser ist nicht spezifisch für SUD validiert.

Teil 3: Perspektiven der Betroffenen

Die Auswertung der neun Interviews zeigte, dass soziale Themenbereiche mit 118 Nennungen bei Weitem am häufigsten erwähnt wurden, gefolgt von behandlungsbezogenen Aspekten (55 Nennungen) und sozioökonomischen Themen (52 Nennungen). Familie, Beruf und Behandlungsmotivation waren die am häufigsten angesprochenen Einzelthemen.

Die Betroffenen beschrieben, wie der Substanzkonsum schrittweise alle Lebensbereiche zerstörte – Beziehungen, Arbeit sowie die psychische und körperliche Gesundheit. Der Weg zur Abstinenz dauerte oft Jahre oder Jahrzehnte und erforderte vielfach mehrere Behandlungsversuche. Rückfälle waren häufig – manche Personen erlebten mehr als 20 Behandlungsepisoden. Wichtige Motivationsfaktoren für den Ausstieg waren zerstörte Familienbeziehungen, Gesundheitskrisen, Jobverlust sowie die Angst um das eigene Leben.

Viele Betroffene berichteten von schweren psychischen Belastungen wie Panikattacken, Depressionen, Psychosen und Suizidgedanken, die sich jedoch nach anhaltender Abstinenz häufig besserten. Soziale Isolation – selbst auferlegt oder von anderen erzwungen – war ein zentrales Thema. Die Nüchternheit wurde schließlich als „Freiheit" und „Wiedergeburt" erlebt.

Eine vertiefende Analyse der sozialen Aspekte zeigte, dass die Betroffenen über das Soziale fast ausschließlich als Verlust sprachen: „zerbrochene Beziehungen", „verlorenes Vertrauen", „erzwungene Isolation". Soziale Zerstörung wurde dabei lebendiger und emotionaler beschrieben als die spätere soziale Erholung. Eine Person sagte etwa: „Ich habe die wenigen sozialen Kontakte, die ich noch hatte, zerstört – immer mit der Ausrede, ich will das nicht auf mich und andere laden. In Wirklichkeit wollte ich einfach nach Hause und wieder trinken." Solche Momente wurden plastisch geschildert. Die nach der Genesung neu gewonnenen sozialen Beziehungen hingegen blieben in den Erzählungen oft blass und knapp – als wäre das neue Leben noch nicht wirklich angekommen. Auffällig war auch, was völlig fehlte: Niemand sprach über Nachbarschaft, Gemeinschaft oder gesellschaftliche Teilhabe. „Das Soziale" bedeutete für die Betroffenen vor allem Familie und Arbeit – alles darüber hinaus blieb unerwähnt.

Vier wiederkehrende Bilder prägten die Erzählungen. Erstens wurde soziale Ausgrenzung als persönlicher Abstieg beschrieben – mit Wörtern wie „sinken", „fallen" oder „sich wieder aufrappeln", als wäre der Verlust sozialer Zugehörigkeit ein moralisches Versagen und keine gesellschaftliche Benachteiligung. Zweitens erlebten viele Betroffene die Zeit des Konsums als ereignislose Wiederholung – ein „Hamsterrad", aus dem es keinen Ausweg zu geben schien. Drittens tauchte der Wunsch nach einem sozialen Neubeginn in einem geschützten, neutralen Raum auf – so, als müsse die soziale Welt erst vollständig zurückgesetzt werden, bevor man neu beginnen könne. Viertens wurde das Leben selbst als Kampf beschrieben, wobei der „Feind" oft schwer zu benennen war – manchmal war es die Substanz, manchmal man selbst, und teilweise die Umwelt.

Darüber hinaus zeigten sich bemerkenswerte, unbewusste Verknüpfungen in den Erzählungen. Kinder wurden nicht nur als Motivationsgrund genannt, sondern fungierten wie eine innere Kontrollinstanz: Die Angst, dass einem das Kind beim Trinken erwischt oder enttäuscht wird, wirkte disziplinierend – ähnlich dem Gefühl, ständig beobachtet zu werden. Haustiere spielten eine besondere Rolle als erster Schritt zurück in soziale Beziehungen: Sie gaben Struktur und emotionale Nähe, ohne die Risiken menschlicher Beziehungen zu tragen – eine Art Brücke aus der Isolation. Schließlich wurde finanzielle Teilhabe – also Arbeit haben, Steuern zahlen – unbewusst als Eintrittskarte in die Gesellschaft beschrieben: Wer nicht arbeitet, gehört nicht dazu.

Diskussion

Die Ergebnisse zeigen eine auffällige Diskrepanz zwischen dem, was die vorhandenen Messinstrumente erfassen, und dem, was Betroffene als zentral erleben. Überraschenderweise stimmte von allen in den klinischen Leitlinien empfohlenen Standardinstrumenten lediglich der ASI mit den in der MT-Forschung tatsächlich verwendeten Instrumenten überein. Dies deutet auf eine geringe Verbreitung standardisierter Messmethoden in der Praxis hin.

Substanzverlangen wurde in fünf von sieben Studien gemessen, Lebensqualität hingegen nur in drei – obwohl sie für die Betroffenen eine zentrale Rolle spielt. Die Interviews zeigten, dass Genesung weit mehr bedeutet als das bloße Aufhören mit dem Konsum: Sie umfasst die Wiedergewinnung von Selbstbestimmung, sozialen Beziehungen, Beruf und Alltagsstruktur.

Besonders relevant ist, dass soziale Aspekte durch die identifizierten Messinstrumente systematisch unzureichend erfasst werden. Bestehende Skalen messen soziales Funktionieren anhand fester, unveränderlicher Momentaufnahmen, etwa von Beziehungsproblemen oder von wahrgenommener Unterstützung. Die tatsächliche Komplexität – z. B. soziale Kontakte als gleichzeitig gewünscht und belastend, der Verlust der Behandlungsgemeinschaft als unterschätztes Risiko beim Therapieende – bleiben damit unerfasst. Auch Freundschaftsaufbau, Nachbarschaft und gesellschaftliche Teilhabe werden von keinem der identifizierten Instrumente systematisch erfasst.

Genesung ist kein geradliniger Prozess, doch die meisten Messinstrumente setzen einen solchen voraus. Forschungsbefunde zeigen, dass langfristige Abstinenz (über ein Jahr) mit den stärksten Verbesserungen bei der Lebenszufriedenheit, den kognitiven Funktionen und dem psychischen Wohlbefinden verbunden ist. Die Interviews bestätigen, dass Aktivitäten, die dem Leben Bedeutung verleihen – Arbeit, Sport, Beziehungspflege, strukturierte Tagesabläufe – dabei eine entscheidende Rolle spielen.

Limitationen

Der Bericht fokussiert sich auf Erwachsene; ein lebenslagenübergreifender Ansatz, der auch Jugendliche einbezieht, wäre jedoch sinnvoll. Da ausschließlich systematische Übersichtsarbeiten herangezogen wurden, bleiben Erkenntnisse aus Einzelstudien unberücksichtigt – was jedoch die Zuverlässigkeit der Befunde stärkt. Nicht alle relevanten Eigenschaften, die zeigen, ob ein Messinstrument auch über längere Zeit zuverlässig misst, konnten vollständig untersucht werden. Die qualitative Interviewstudie erlaubt keine Kausalschlüsse, liefert jedoch wertvolle Einblicke in die gelebte Erfahrung der Betroffenen.

Schlussfolgerungen und Empfehlungen für zukünftige Forschung

Alle 14 analysierten Messinstrumente sind etabliert, validiert und zuverlässig. Ihr Einsatz in der MT bei SUD erfordert jedoch eine sorgfältige Abwägung der Stärken und Schwächen im jeweiligen Anwendungskontext. Für eine aussagekräftige Langzeitforschung empfiehlt sich eine Kombination aus selbstberichteten Instrumenten, klinischen Messungen und biologischen Markern, ergänzt mit qualitativen Erhebungen.

Aus den Ergebnissen lassen sich drei zentrale Empfehlungen für die künftige Forschung ableiten: Erstens sollten MCIDs spezifisch für SUD-Populationen entwickelt werden, damit Verbesserungen besser interpretiert werden können. Zweitens braucht es Langzeitstudien mit Behandlungsdauern von über sechs Monaten und Nachbeobachtungszeiträumen von mindestens zwölf Monaten. Drittens sollten standardisierte Messprotokolle auf Basis international vereinbarter Kernendpunkte entwickelt werden, um die Vergleichbarkeit zwischen Studien zu verbessern. Da die MT-Forschungsmethodologie im Vergleich zur etablierten Mental-Health-Forschung noch wenig ausgereift ist, wurde die Suche bewusst über MT-spezifische Instrumente hinaus auf validierte Instrumente aus der psychischen Gesundheitsforschung ausgeweitet – insbesondere solche, die sozioökonomische Dimensionen abdecken. Darüber hinaus sind eine bessere Erfassung sozioökonomischer Aspekte (z. B. Arbeit, Wohnsituation, Bildung), ein umfassendes Monitoring unerwünschter Ereignisse sowie die Entwicklung kulturell sensibler Instrumente dringend erforderlich. Die Perspektiven von Betroffenen sollten von Anfang an in die Auswahl von Messinstrumenten und Studiendesigns einfließen.

1 Background

internationales Horizon-Europe Projekt "FALCO": Fighting Addictions, improving Lives: COmprehensive drug rehabilitation with music

7 Länder beteiligt

The European Health and Digital Executive Agency (HADEA) funds the Horizon Europe project (10115588) FALCO (Fighting Addictions, improving Lives: COmprehensive drug rehabilitation with music) [1]. Involving seven countries, FALCO is the first large, parallel, pragmatic, randomised controlled trial (RCT) to address the long-term effects of music therapy (MT) for substance use disorders (SUD), aiming to reduce the disease burden through an innovative, effective, and affordable MT treatment while strengthening research expertise. Its recommendations will inform intervention delivery across Europe and beyond, aiming to improve safety, effectiveness, cost-effectiveness, and quality of life (QoL) for people with SUD.

Ziel des Berichts: Überblick zu patient:innen- und politikrelevanten Endpunkten

This report concludes the work on deliverable "D7.1 Review of patient- and policy-relevant outcomes", which is part of WP7 Long-term clinical outcomes: preparations. The objective of this work has been to perform a literature review to ensure the quality of the long-term clinical outcomes. This includes definitions and scales recommended for measuring outcomes of critical (e.g., recovery, substance use severity and QoL) and secondary importance (e.g., mental health, social functioning). As the FALCO project includes several German-speaking partner institutions and this report also aims to disseminate findings to national stakeholders, a detailed German-language summary has been included.

1.1 Public health dimension of substance use disorders

Suchterkrankungen (SUD, substance use disorder): haben hohe globale Krankheitslast

oftmals Behandlungsabbruch

People with SUD carry a high global disease burden, with 1.3% of disability-adjusted life years lost to illicit drugs and 4.2% to alcohol [2], and one million high-risk opioid users in the EU [3]. Multimorbidity is highly prevalent, including polysubstance use, co-occurring mental health conditions, and various preventable sequelae such as infectious and non-communicable diseases [2]. Bidirectional links between SUD, crime, and poverty further underscore the societal importance of improving rehabilitation [2]. Many patients drop out or do not benefit from existing pharmacological and non-pharmacological interventions [4].

Inanspruchnahme von Gesundheitsdienstleistungen: Gründe und Charakterisierung

SUD is also linked to socioeconomic factors such as income and unemployment, and is frequently associated with high health service use. Healthcare utilisation – characterised by outpatient visits, hospital admissions, prolonged length of stay, high re-hospitalisation rates, emergency department visits, prescription drug use, and associated costs – is therefore of particular importance in this population [5, 6].

1.2 Music therapy for treating substance use

Musiktherapie (MT): klinischer/evidenzbasierter Einsatz, um individuelle Behandlungsziele zu erreichen

MT is defined as the clinical, evidence-based use of music interventions to achieve individualised goals within a therapeutic relationship. A credentialed music therapist employs music and musical activities – such as music listening and active music-making – tailored to the individual's cultural context and needs. While music is used across all human cultures for self-expression and social connection, its application varies widely [7-10]. Throughout this chapter, we distinguish between *music* (referring to general effects of music as described in the broader literature) and *MT* (referring specifically to clinical interventions delivered by credentialed music therapists).

MT: wirksam für Personen mit SUD

hohe Rückfallquote im 1. Jahr nach der Behandlung

Various forms of MT have shown benefits for people with SUD (21 RCTs; n=1984) [11], offering unique advantages over treatment as usual. For instance, receptive interventions such as lyric analysis can help address harmful behaviours associated with substance use, while active interventions – such as improvisational MT and songwriting – have shown benefits for motivation, readiness to change, and depressive symptoms [12]. By analysing lyrics as a third object rather than personal history, perceived treatment stigma can be reduced [13], making individuals with low motivation less likely to oppose treatment [12].

MT-Effekte: Abstinenz und alternative Belohnungsstrategien

Previous studies suggest that MT can help patients achieve higher abstinence levels and better treatment outcomes, potentially by helping them acquire alternative reward strategies. MT may therefore provide an effective and affordable addition to SUD treatment [11].

Leitlinien: unterschiedliche Empfehlungen

MT kann als Teil von postakuten Behandlungen angeboten werden und einen positiven Beitrag zur Genesung beitragen

Clinical guidelines for SUD differ considerably in their recommendations, partly because evidence for SUD interventions is limited. In many guidelines, MT is not explicitly mentioned. According to two identified guidelines, MT can be offered as part of post-acute treatment (grade of recommendation: 0, level of evidence: 2b, overall agreement: 96.2%). The Dutch Akwa GGZ guideline further states that arts therapies, such as MT, can be used for problematic alcohol consumption and alcohol dependence. Even with limited evidence of effectiveness, MT as an additional clinical treatment option can support motivation for change and may influence secondary symptoms, potentially contributing to recovery. Little is known about harmful side effects [14].

Musik kann aufgrund von früheren Musikerlebnissen negativ wirken

daher Expertise von Musiktherapeut*innen notwendig

In people with SUD, music can be applied both beneficially and harmfully [7-9]; without therapeutic guidance, complex links between substance use and music experiences [15-17] – potentially explained by underlying traits such as curiosity, sensation-seeking, and risk-taking [18] – may reinforce rather than reduce harmful behaviour. Specialised MT expertise is therefore needed to minimise harmful and maximise health-promoting uses of music in SUD [15, 19]. As detailed in the following sections, substance abuse is a disorder of the brain's reward system, and MT addresses this by stimulating the same brain pathways involved in substance abuse.

MT bietet Möglichkeiten für Veränderungen und Entwicklung

anerkannte Ausbildung in MT notwendig

MT is suggested for people with SUD, offering opportunities for change and development through inter-relational and musical cooperation between therapist and patient. Within this therapeutic framework, MT can facilitate personal expression, communication, and work with anxiety, self-understanding, motivation, abstinence, and personal growth. For instance, lyric analysis, songwriting, and Guided Imagery and Music – each show specific therapeutic potential. Lyric analysis has been associated with enhanced treatment motivation and problem recognition, while songwriting has been associated with

improved treatment readiness and emotional expression, with findings for Guided Imagery and Music remaining mixed [20, 21]. MT can be delivered across settings – specialised health services, community, and prisons – as individual or group interventions, and must be delivered by professionals with recognised MT training [22].

Musik als Grundlage: neurophysiologische, psychologische und soziale Effekte

At the neurophysiological level, music engages reward pathways and dopaminergic systems [23-26], and is consistently rated as highly intrinsically rewarding [27], sharing neurochemical and neurophysiological reward pathways with recreational substances [28]. Dopamine is central to both substance use [26] and musical experiences [23, 26]. At the psychological level, music modulates attention, emotion, and cognition, facilitates nonverbal symbolic expression, and supports self-efficacy [29]. At the social level, shared musical experiences foster interpersonal communication and group cohesion [30, 31]. However, in people with SUD, a discrepancy between expected and actual experience leads to diminished responses in reward regions [32], limiting the naturally occurring benefits of music.

MT: Wirkmechanismen, klinische Wirksamkeit und Langzeitperspektive

MT harnesses these properties within a structured clinical framework. The mechanisms through which MT influences SUD are multifaceted and cannot be reduced to a single pathway. MT functions as an umbrella term encompassing diverse interventions grounded in different theoretical rationales, techniques, and intended effects [30, 33], with different approaches likely operating through distinct but overlapping mechanisms. Therapeutic factors in MT can be distinguished into common factors shared across all psychotherapies, joint factors shared across the creative arts therapies, and specific factors unique to MT [31]. Unlike addictive substances, MT fosters positive reinforcement, stress reduction, and emotional regulation without harmful effects, suggesting its value as an affordable supplementary treatment [11]. Nevertheless, it remains insufficiently understood how, when, and for whom specific MT interventions produce change, underscoring the need for transdisciplinary research on mechanisms and moderators of change [31, 33]. While MT has demonstrated clinical efficacy in SUD treatment through clinically guided engagement of the reward system, craving reduction, and motivation enhancement, a comprehensive evaluation of its long-term value requires an understanding of recovery as a multidimensional process. MT's capacity to provide alternative reward strategies and foster emotional regulation over extended periods suggests its impact extends beyond immediate symptom reduction.

1.3 The concept of substance use disorder recovery

4 Kernmerkmale von Genesung

Recovery is a dynamic, individualised, relational, and intentional process involving sustained efforts to improve wellbeing [34]. Its core characteristics include: it is continuous, extends beyond managing substance use to encompass broader life aspects, involves overall life improvements and well-being, and is person-centred – recognising that each individual's journey and definition of recovery are unique [35].

Weg zur Genesung ist einzigartig und eine nachhaltige professionelle und soziale Unterstützung ist entscheidend

As each recovery path is unique, conceptions of success and well-being differ. Key elements fostering a better life include growing self-awareness of one's substance use, a sense of personal safety, understanding how substance use affects relationships, and feeling involved in the community. For individuals with SUD, sustained professional and social support is considered essential for ongoing recovery [36].

zentrale Elemente der Genesung z. B. negative Emotionen ohne Substanzkonsum bewältigen

The literature identifies four central elements of recovery shared among individuals: personal growth, self-honesty, taking responsibility for what one can change, and reacting to situations more balanced. Additionally, prevalent elements include the ability to enjoy life and manage negative emotions without substance use, achieving abstinence or non-problematic use, and living a life that contributes positively [37].

nachhaltige Genesung erfordert eine erhebliche psychologische Veränderung …

Sustaining recovery requires a significant psychological shift, including a strong commitment to a new way of life – often supported by accepting substance abuse as a 'disease' or finding a new faith-based identity. Strong ongoing support is crucial, including participation in substance abuse support

… und kontinuierliche soziale Unterstützung

groups, close relationships with family, spouses, and sponsors, and a supportive work environment. Actively helping others also plays a significant role in maintaining one's own recovery [38].

biologische, psychologische, soziale, politische und kulturelle Dimensionen integrieren

Research advocates a comprehensive approach to SUD and recovery that integrates biological, psychological, social, political, and cultural dimensions. Acknowledging SUD as a biopsychosocial disorder, recovery is a multifaceted, complex, and often long-term process towards abstinence or improved well-being [36, 39].

Genesung als multidimensionaler Prozess: Aktivität, Soziales und Identität als Schlüsseldimensionen

Qualitative research [40] emphasises that meaningful, structured activities such as music engagement are central to sustained recovery, enabling individuals to develop mastery, repair social connections, and rebuild identity. Close relationships and social support networks further stabilise recovery, with the quality of interpersonal connections significantly influencing sustained behavioural change and well-being [40]. These perspectives indicate that measuring MT's effectiveness requires tools capturing improvements across multiple life domains. The following section, therefore, outlines the multidimensional nature of recovery from substance use.

Schlossbergs Transitionsmodell von Sucht zur Abstinenz ist kategorisiert nach Selbst, Situation, Strategien und Unterstützung …

… welches sich mit MT deckt

Summarising these aspects, we refer to Schlossberg's Transition Process Model [38], with which MT aligns across all four dimensions. Within the 'self', MT fosters psychological resources through emotional expression, self-awareness, and identity reconstruction. In 'situation', music provides a structured, non-stigmatising context for recovery. Regarding 'strategies', MT encompasses diverse interventions with distinct mechanisms: receptive approaches, such as lyric analysis, enhance motivation and problem recognition, while active interventions, such as songwriting and improvisation, support emotional expression and identity work [12, 20, 21]. In 'support', group and community music engagement strengthens social connections and support networks. MT thus functions as a comprehensive intervention activating all four components simultaneously, making it particularly valuable for long-term recovery.

1.4 Research on substance use disorder

1.4.1 Core outcome sets in substance use disorder research

empfohlene Messinstrumente der Kernendpunkte

Three guidelines on outcome measurement in SUD were identified. The 'National Professional Guideline for Treatment and Rehabilitation of Substance Use Problems and Addiction' by the Norwegian Directorate of Health recommends the following core outcome set (COS) measuring instruments [22]:

- Alcohol Use Disorders Identification Test (AUDIT)
- Alcohol Use Disorders Identification Test – Consumption (AUDIT-C)
- Drug Use Disorders Identification Test (DUDIT)
- European Addiction Severity Index (EuropASI)
- European Adolescent Assessment Dialogue (EuroADAD)
- Drug Use Disorders Identification Test Extended (DUDIT-E)
- Alcohol Use Disorders Identification Test Extended Version (Alcohol-E)

das Internationale Konsortium für die Messung von Gesundheitsergebnissen einigt sich weltweit auf Kernendpunkte

The International Consortium for Health Outcomes Measurement (ICHOM) develops patient-centred outcome measure sets, prioritising clinical and QoL outcomes most important to patients and recommending them to all providers. This international, multi-disciplinary working group reached consensus on a globally applicable minimum set of outcome measures [41, 42].

empfohlene Messinstrumente der Kernendpunkte

This consensus guideline, the 'Addiction Data Collection Reference Guide,' recommends the following tools for measuring outcomes most relevant to patients with SUD [41, 42]:

- Symptom burden → Patient-Reported Outcomes Measurement Information System (PROMIS) Short Form v1.0 – Alcohol Use 7a (PROMIS-Alcohol)
- Symptom burden → PROMIS Short Form v1.0 – Severity of Substance Use (Past 30 Days) 7a (PROMIS-Substance)
- Symptom burden → PROMIS Short Form v1.0 – Smoking Nicotine Dependence for All Smokers 8a (PROMIS-Smoking)
- Symptom burden → Heaviness of Smoking Index (HSI)
- Frequency and quantity, health-related QoL, global/psychosocial functioning → National Treatment Agency for Substance Misuse (NHS) Treatment Outcomes Profile for Substance Misuse – Section 1 (TOP-S1)
- Comorbidities → Modified Self-Administered Comorbidity Questionnaire (SCQ)
- Health-related QoL, global/psychosocial functioning → World Health Organization (WHO) Disability Assessment Schedule 2.0 – 12-item version (WHODAS12)
- Overall physical/mental health and wellbeing → PROMIS Scale v1.2 – Global Health (PROMIS-GH-10)
- Health-related QoL, global/psychosocial functioning → KIDscreen-10 Index (KIDSCREEN10)
- Health-related QoL, global/psychosocial functioning → Substance Use Recovery Evaluator (SURE)

empfohlene Messinstrumente der Kernendpunkte der (gesundheitsbezogenen) Lebensqualität

A 'Systematic review of quality of life and health-related quality of life as outcomes in substance and behavioural addictions' identified the following (health-related) QoL instruments [43]:

- Alcohol Quality of Life Scale (AQoLS)
- World Health Organization Quality of Life-26 items (WHOQOL-BREF)
- Alcohol Quality of Life-9 item (ALQoL-9)
- Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form (Q-LESQ-SF)
- Short Form-12 Item (SF-12)
- Drug User Quality of Life Scale (DUQoL)
- Quality of Life Interview (QLI)
- Short Form-36 Item (SF-36)

sozioökonomische Endpunkte kaum gemessen

Socioeconomic aspects are poorly covered by these core outcome sets (e.g., EuropASI, EuropADAD, TOP-S1), but they are relevant to decision-makers regarding public reimbursement.

1.5 Project aims and research questions

hohe Rückfallquote im 1. Jahr nach der Behandlung

daher längere (Nachsorge)behandlungen nötig

Given that drug dependence is a chronic medical illness with relapse rates of 40 to 60%, comparable to those of other chronic conditions such as diabetes and hypertension [44], evaluating substance use patterns and recovery outcomes beyond the immediate treatment period is essential to assess MT's long-term effectiveness. Drug dependence has generally been treated as if it were an acute illness, yet long-term care strategies produce lasting benefits [45]. Longer treatment durations are associated with lower substance use at long-term follow-ups and better aftercare engagement [46, 47], suggesting that extended music interventions may be needed to impact long-term substance use behaviours. Investigating MT's impact on treatment retention and its long-term effects on people with SUD is therefore crucial [11].

1.5.1 Aims of the report

Ziel: patientenrelevante Messinstrumente für Langzeitforschung identifizieren

This report is part of FALCO's work package 7, focusing on long-term MT outcomes. We conducted a systematic review to identify patient- and policy-relevant long-term outcomes, complemented by interviews with people with SUD. The aim is to establish the relative importance and clinical relevance of selected outcomes for FALCO, ensuring high-quality long-term clinical data and reliable, efficient data collection across sites. This includes scales recommended for measuring outcomes of critical (e.g., addiction severity, recovery) and secondary importance (e.g., mental health, social functioning). To better understand which outcomes are most relevant to patients with SUD, we conducted interviews with service users. This shared understanding of perspectives helps identify patient-relevant outcomes and select appropriate measuring instruments.

Part 1 aims to provide an overview of

- **clinical** (e.g. addiction severity, recovery, substance use, craving) and
- **socioeconomic** (i.e. payer-relevant criteria such as work integration, unemployment, health service use, treatment retention)

outcome parameters identified in existing MT evidence for SUD, and whether the **measuring instruments** used have been tested for **validity** and **reliability**. Additionally, this part examines how **(serious) adverse events** are assessed (RQ1).

Teil 1: Übersicht über klinische und sozioökonomische Ergebnisparameter und Messinstrumente

Part 2 describes the **characteristics** of the measuring instruments identified in Part 1, prioritising those that are most frequently used, validated, reliable, appropriate, feasible and user-relevant, with particular focus on their suitability for **long-term** measurement in patients with SUD (RQ2).

Teil 2: Merkmale der Messinstrumente

Patient relevance of identified outcomes is analysed in ***Part 3*** by contrasting findings with perspectives from nine SUD service users across three countries (Austria, Norway, and Poland), informing the selection of service-user-relevant measuring instruments for subsequent FALCO work packages and beyond.

Teil 3: Patient*innenperspektiven aus 9 Interviews

This systematic review was conducted in collaboration with the Institute of Psychology at the University of Gdańsk, Poland, the Centre for Alcohol and Drug Research (KORFOR), Stavanger University Hospital, Norway, and the Anton Proksch Institute, Austria.

Kooperationen

The three parts of this report are closely interrelated and build upon one another sequentially. ***Part 1*** (RQ1) provides the empirical foundation: by systematically identifying all outcome parameters and measuring instruments reported in the MT-SUD literature, it generates a pool from which further selection proceeds. ***Part 2*** (RQ2) narrows this pool by applying predefined selection criteria to identify measuring instruments for detailed characterisation, including their suitability for long-term measurement. The data developed for Part 1 also structure the interview guide used in ***Part 3*** (RQ2), ensuring that service user perspectives are systematically mapped onto the same outcome categories identified in the literature. Triangulating all three parts then enables an assessment of how well current measurement practices reflect what matters to people with SUD – ultimately informing instrument selection for further research.

dreiteiliger Aufbau: sequenzielle Verknüpfung der Forschungsteile

1.5.2 Research questions

In the current report, we respond to the following two research questions (RQs):

RQ1: Which **measuring instruments** assessing **clinical and socioeconomic outcome parameters** and **(serious) adverse events** can be identified in the existing evidence of MT in people with SUD, and are these tools tested for validity and reliability? (→ *review of relevant clinical and socioeconomic outcome parameters and how they are measured; tools assessing (serious) adverse events (Part 1)*)

Wie werden klinische und sozioökonomische Ergebnisparameter und unerwünschte Ereignisse gemessen?

RQ2: What are the **characteristics** of the most often used, validated and reliable measuring instruments in the context of MT in SUD, and are they appropriate, feasible and user-relevant, particularly for **long-term** measures in patients with SUD? (→ *review of measuring instruments and their characteristics and service user interviews (Parts 2 & 3)*)

Was sind die Merkmale dieser Messinstrumente? … geeignet für Langzeitmessungen?

2 Methods

Methodik: Überblick über den dreiteiligen Ansatz

The methodology of this report is structured in three interconnected parts, mirroring the research questions outlined in Chapter 1.5.2. *Parts 1* and *2* are based on a systematic literature search supplemented by a manual search, identifying clinical and socioeconomic outcome parameters and measuring instruments used in MT research for SUD (RQ1), and characterising the most frequently used, validated, and reliable instruments in detail (RQ2). *Part 3* complements these findings through a qualitative interview study with nine service users across three countries, providing lived-experience perspectives on outcome relevance (RQ2).

2.1 Protocol registration

Studienregistrierung

The protocol for the current systematic review has been registered on PROSPERO (https://www.crd.york.ac.uk/PROSPERO/view/CRD420250648836, registration number: CRD42022330778).

ursprüngliche Ein- und Ausschlusskriterien

The *initial* criteria of the population taken from the FALCO project proposal [1] for the systematic search were:

Inclusion criteria:

- Patients with SUD seeking or receiving treatment at the recruitment site for an existing substance use disorder (SUD) based on ICD-10 criteria
- ≥18 years old
- Not currently undergoing detoxification (detoxification has been completed or is not currently planned at the time of recruitment)
- Any type of substance use, including polysubstance and alcohol use

Exclusion criteria:

- Exclusively nicotine dependence
- Psychotic episode in the last three months
- Insufficient language skills to participate in treatment without the use of a translator
- Hearing impairment that considerably impairs hearing of music played at a moderate volume (not relevant if hearing is sufficiently compensated by a hearing aid)
- Existing diagnosis of dementia
- Is currently receiving MT or has received regular MT (i.e. at planned and reoccurring intervals, not counting single random occurrences) during the past year

Ausweitung der Ein- und Ausschlusskriterien für die systematische Suche

As these eligibility criteria were taken from the FALCO project proposal [1], they were too detailed for this current systematic review. More general criteria ensure a comprehensive review, reduce the risk of excluding relevant research, and capture a broader range of studies. Specifically, age restrictions and exclusions such as psychotic episodes and hearing impairment were removed, as these are not always consistently reported. Articles were not exclud-

ed based solely on their use of ICD-10 criteria for SUD diagnosis, given that diagnostic approaches, reporting practices, and criteria vary across clinical and research settings.

aktuelle Ein- und Ausschlusskriterien (siehe PICOS)

We therefore adapted the initial eligibility criteria into more general criteria for study selection, detailed in the PICOS question (see 2.2.2). These changes were also documented in PROSPERO.

Anpassung der Intervention gemäß FALCO-Protokoll

A minor change was made to the intervention in the PICOS: 'Active music groups (AMG) and music listening groups (MLG)' was revised to 'Music-based interventions, incl. active (music making), receptive (music listening) and mixed forms', in accordance with the FALCO project protocol. The FALCO trial compares two non-pharmacological, group-based music interventions against treatment as usual. An active music-making group involves participants actively playing musical instruments, while a music-listening group involves receptive music listening using a stereo system and recordings. Both interventions are delivered by qualified music therapists in weekly group sessions of up to 90 minutes, with approximately six participants per group, and are provided for at least 12 months [1].

2.2 Parts 1 & 2: Review of relevant clinical and socioeconomic outcome parameters, tools assessing (serious) adverse events, and measuring instruments and their characteristics

2.2.1 Search process

systematische Suche in 4 Datenbanken

The systematic literature search was conducted on the 21st of February 2025 in the following databases:

- The Cochrane Library
- Embase
- MEDLINE
- International Network of Agencies for Health Technology Assessment (INAHTA) database

Sucheinschränkung

The systematic search was limited to reviews published in all languages from inception until 2024. An example (Medline via Ovid) of the specific search strategy employed is provided in the Appendix.

2.2.2 Inclusion and exclusion criteria (PICOS question)

Einschlusskriterien für relevante Studien

Inclusion and exclusion criteria for relevant studies are summarised in Table 2-1.

Table 2-1: PICOS question for research questions 1 & 2

PICOS for RQ1 & RQ2 for systematic literature search	
Population	Inclusion criteria: ■ individuals with existing substance use disorder (SUD) ■ any type of substance use, including polysubstance and alcohol use Exclusion criteria: ■ exclusively nicotine dependence
Intervention	Music-based interventions, incl. active (music making), receptive (music listening) and mixed forms
Control	-
Outcomes	**RQ1: Clinical[1] and socioeconomic[2] outcome parameters and tools assessing (serious) adverse events[3]** **RQ2: Measuring instruments and their characteristics** *General characteristics:* ■ Outcome measured ■ Mode of assessment ■ Validation (validity/reliability) ■ Post-assessment in treatment and latest follow-up where measuring instrument was applied ■ Number of languages in which instruments are available ■ Types of languages in which instruments are available ■ Type of MT intervention in which the instrument was applied ■ Setting of assessment where instrument was applied (e.g. in-/outpatient treatment, detox, rehabilitation) ■ Limitations *Application characteristics:* ■ Number of items in each instrument ■ Total score range ■ Interpretation of score and minimum clinically important difference (MCID) ■ Testing time ■ Instrument user (e.g. self-assessment by service user, clinician-, performance reported) ■ Costs and regulations for use
Study design	Reviews
Publication period	From inception to 2024
Databases	The Cochrane Library, Centre for Reviews and Dissemination (CRD) database, Embase, MEDLINE, PsycINFO, Web of Science
Languages	All

[1] E.g., addiction severity: Addiction Severity Index (ASI), psychological symptoms: Beck's Depression Inventory (BDI)

[2] Definition of socioeconomic outcomes is based on the WHO Social Determinants of Health [48], e.g., income, education, occupation, employment, social and material capital, material circumstances such as living and working conditions

[3] E.g., relapse requiring hospitalisation, suicide (attempts)

2.2.3 Flow chart of study selection

Literaturauswahl aus 451 Quellen

After deduplication, overall, 451 citations were identified through the systematic literature search. Two researchers (LG, DM) independently screened the references, and in case of disagreement, a third researcher (LB) was involved in solving the differences. The selection process is displayed in Figure 2-1.

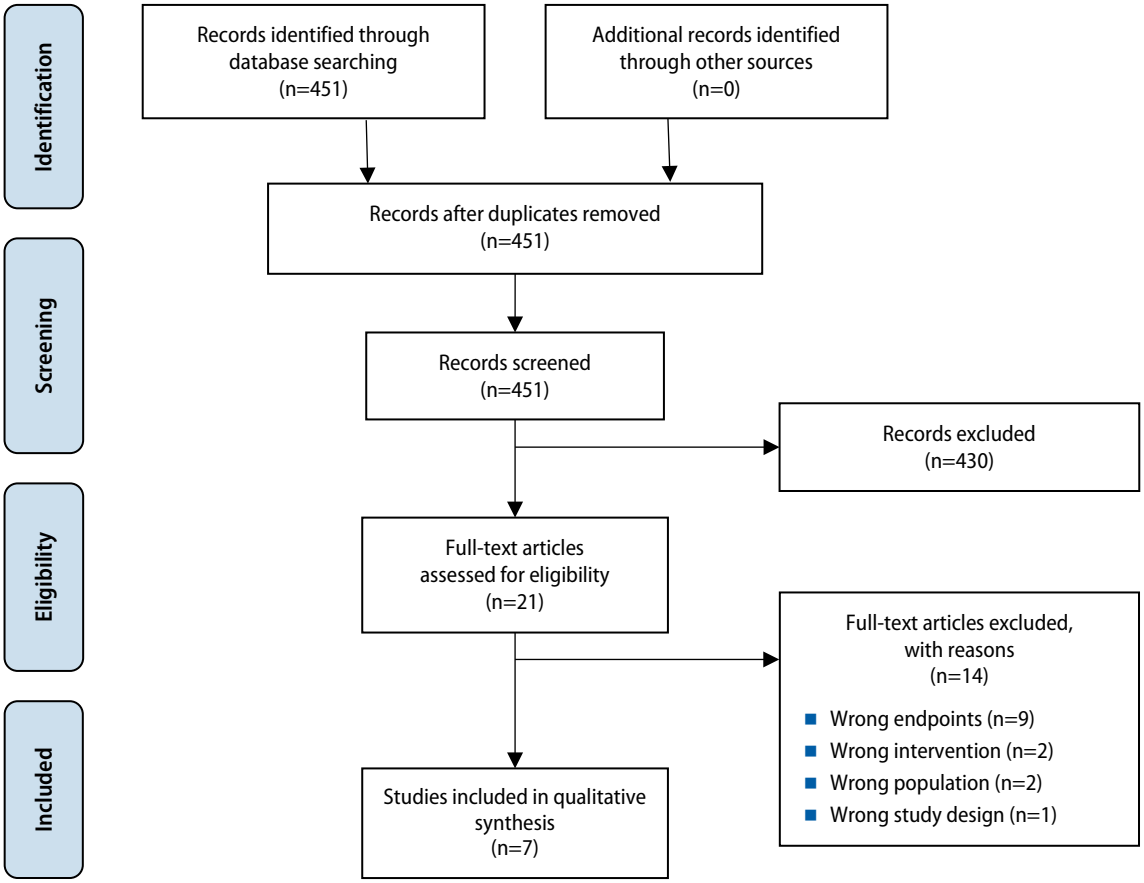


Figure 2-1: Flow chart of study selection (PRISMA flow diagram)

7 Übersichtsarbeiten eingeschlossen

From the systematic search, we identified seven reviews investigating MT in patients with existing SUD or any type of substance use, including polysubstance and alcohol use.

nur Daten hinsichtlich SUD eingeschlossen

In two of these seven studies, we did not extract data on *exclusively* nicotine dependence (n=3 studies extracted [49]; n=12 studies extracted [50]) according to the PICOS question.

Handsuche für Forschungsfrage 2

For RQ2, we conducted manual searches to obtain detailed information and characteristics of the measuring instruments selected from the pool identified in RQ1. The selection was based on predefined criteria established by the authors (see 2.2.6).

Validierungsverfahren

Workflows are organised according to the principle of dual control; the results are subject to internal and external review. The data extraction tables were piloted (IBB) on the 22nd of March 2025. There were no major changes in the extraction tables.

2.2.4 Manual search on socioeconomic data collection tools

Handsuche nach Instrumenten, die sozioökonomische Kriterien messen …

Given that instruments measuring socioeconomic criteria (e.g., payer-relevant outcomes such as integration into work life, unemployment days, health service use, and retention in treatment) were expected to be underrepresented in the core music therapy literature, we supplemented the systematic search with an additional manual search. Indeed, the systematic search yielded few results on socioeconomic aspects, confirming the need for this additional step.

… im Bereich der psychischen Gesundheit …

We therefore focused on instruments assessing socioeconomic factors in mental health services research more broadly. To this end, two researchers (LG, DM) conducted a manual search for socioeconomic outcomes in the mental health field and how they are measured across the following sources:

… in verschiedenen Quellen

- PubMed,
- ResearchGate,
- Google Scholar,
- Research institutions,
- Statistical institutions, and
- General searches on Google.

in die Gesamtergebnisse integriert

The results of the manual search were subsequently integrated into the overall results and processed alongside the findings from the systematic search using the same procedures.

2.2.5 Quality appraisal

keine Bewertung von Studienqualität und Verzerrungsrisiko

No risk of bias assessment (e.g., AMSTAR – Assessing the Methodological Quality of Systematic Reviews) was conducted as this review focuses on outcome parameters and measuring instruments rather than the effectiveness/efficacy of MT.

2.2.6 Data extraction and data analysis

Datenextraktion und -kontrolle

The data from the selected studies were extracted into data extraction tables and can be found in the Appendix (see Table A-1, Table A-3, and Table A-4). The single-data extraction method with verification by another researcher was used: One researcher (LG) extracted the data, and the other researchers (DM, IBB, AHE, BAK) verified it. We categorised the parameters identified that were derived from the literature results. To better represent metric tools, qualitative and general measuring instruments (i.e., telephone calls, unspecific reports/surveys/interviews/lists/questionnaires, point scales, Likert-type scales) were excluded. For RQ1, after data extraction, all clustered measuring instruments were checked for validity and reliability. For RQ2, we decided to select the measuring instruments of RQ1 for detailed analysis according to the following criteria:

Forschungsfrage 2: Auswahlkriterien für Messinstrumente zur detaillierten Analyse …

- Selecting *at least one* measuring instrument of each parameter (e.g. substance use, QoL parameters, biomedical parameters)
- Selecting the *most often used* measuring instruments (deduced from Table A-1 and Table A-2)
- Selecting *validated and reliable* measuring instruments (deduced from Table A-1 and Table A-2)

- When the frequency of identified measuring instruments according to the criteria was similar, an additional literature search on the specific instrument was conducted, or experts from the MT field were consulted

Detailed information was deduced from different sources to answer the following details of RQ2:

- Appropriateness and feasibility (deduced from Table A-3 and Table A-4)
- Appropriateness for long-term measures (deduced from Table A-3 and Table A-4)
- User-relevant measuring instruments (deduced from service user interviews, see 3.3): User-relevance of the measuring instruments was assessed by examining whether the categories and outcomes derived from the systematic review (Part 1) appeared in service users' lived experiences, and how they described them. Open-ended prompts additionally allowed participants to identify relevant outcomes or dimensions not captured by the identified instruments. Triangulating the review findings with service users' perspectives enabled an assessment of how well current measurement practices reflect users' experiences.

… und weitere Informationen zu allgemeinen Merkmalen und Anwendungsmerkmalen

For analysing the *appropriateness, feasibility and user relevance* of the 14 measuring tools, Table 3-4 and Table 3-5 were created. We then examined the positive and negative aspects of the tools and specific criteria, such as testing time, number of items, and translations.

Kriterien für die Bewertung geeigneter, praktikabler und nutzerrelevanter Messinstrumente …

For analysing the *appropriateness for long-term measures,* an inductive approach was used based on Table A-3 and Table A-4). For that purpose, three analytical dimensions were derived to systematically evaluate the long-term suitability of the 14 tools: (1) Core measurement quality and clinical relevance, (2) Practicality and burden of application, (3) Generalisability and context. These three dimensions are grounded in established guidelines and frameworks in health outcome research [51-53], e.g., the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) guideline, providing a theoretically robust foundation for the evaluation. Together, these three dimensions provide a comprehensive framework for evaluating the suitability of measuring instruments for long-term research.

… und die Eignung für Langzeitmessungen

For the *outcome construct stability* of the tools, a classification system ("high", "moderate", "low") was carried out based on the outcomes measured in Table A-3 and Table A-4 assessing the following: behavioural patterns (stable vs. volatile), external factor dependence, HIGH stability (longitudinal consistency and predictable treatment responsiveness), MODERATE (significantly influenced by external factors), LOW (acutely volatile or state-dependent).

Klassifizierungssystem z. B. geringe Stabilität eines Instrumentes, wenn Endpunkt akut schwankend oder zustandsabhängig ist

The *empirical evidence* was assessed by examining the post-assessment (i.e., the last measurement time point within the treatment period) and the latest FU at which the measuring instrument was applied in the included reviews of Part 1, with the following cut-off thresholds: duration of treatment >6 months and latest FU ≥3 months.

weitere Kriterien zur Bewertung der z. B. empirischen Evidenz, …

The following pragmatic cut-off thresholds were considered based on the evidence data in Table 3-4 and Table 3-5 for the *risk assessment* (assessed by limitations of tools; i.e., <4 limitations), *compliance and costs* (assessed by testing time; i.e., ≤10 minutes testing time), *resolution and sensitivity* (assessed by number of items in each instrument (i.e., >15 items) and total score range (i.e., >50 total score range)), and *generalisability* (assessed by number of available languages of the tools; i.e., >10 languages).

… Risikobewertung und Sensitivität des Messinstruments, Therapieadhärenz und Kosten

The identified *limitations* of the measuring instruments (see outcomes in the PICOS question in Chapter 2.2.2) were further analysed to assess their appropriateness and feasibility for RQ2 (see 1.5.2).

Limitationen für weitere Analyse der Instrumente

2.2.7 Data synthesis

A qualitative synthesis of the evidence was conducted. After extracting data in data extraction tables, the data were interpreted and reported in a narrative form. All data from the extraction tables were synthesised.

qualitative Evidenzsynthese

2.3 Part 3: Service user involvement – interview study

Service user involvement generates knowledge that would otherwise not be produced, helping to uncover, for example, restrictions or suffering related to SUD, or aspects expected to improve through MT. To elicit this knowledge, interviews with people with lived experience were conducted in three countries (Austria, Norway, Poland), aiming to improve understanding of patient-relevant outcomes and the selection of appropriate measuring instruments. Reporting followed the COnsolidated criteria for REporting Qualitative research (COREQ) checklist for qualitative research [54].

Interviews: Erfahrungswissen von Suchterkrankten soll patientenrelevante Ergebnisse bringen

2.3.1 Ethical approval

The interview study had been approved in all three participating countries (*Austria*: Ludwig Boltzmann Gesellschaft, Ref: 019_2025; *Norway*: Regional Ethics Committee in Norway, no. 904119; *Poland*: Research Ethics Committee at University of Gdansk, no. 18/2025/WNS). As the interview study included three service users from Norway, three from Poland, and three from Austria, the interview guidelines, participant information, and consent forms were translated into their native languages.

positiver Ethikbescheid in 3 Ländern

2.3.2 Description of service users and the nature of their involvement

The interviews were conducted in outpatient clinic settings using a semi-structured approach, exploring service users' understanding of patient-relevant outcomes and the measurement instruments used to assess them. Face-to-face interviews lasted approximately 30-60 minutes, were audio-recorded via smartphone at each site, transcribed verbatim (using Microsoft Word with manual checking), and translated. No field notes were made, and transcripts were not returned to participants for review. The transcripts served as primary material and were analysed using qualitative content analysis in MAXQDA. Data saturation was not discussed.

Interviews mit Personen mit SUD: Erfahrungen aus erster Hand einbezogen

Einschlusskriterien für interviewte Personen

Interviews in klinischem Setting durch Fachpersonen

At the time of the interview, participants were in therapy with at least one year of sobriety but not currently receiving MT, nor had they received regular MT at planned and recurring intervals during the past year. This criterion was adopted from the FALCO trial eligibility criteria [1], which serves to prevent confounding: participants who are already receiving MT could show effects attributable to that prior or ongoing exposure rather than to the intervention under study. Interview selection criteria aligned with the PICOS defined for the literature search (see Table 2-1). As service users might have had additional physical illnesses, disabilities, or mental illnesses and constituted a vulnerable group, interviews were conducted in clinical settings by psychotherapists, clinical/health psychologists, or social workers (FM, DM, TS), ensuring a safe environment and appropriate support.

Anonymität sichergestellt

Interview data were de-identified to ensure that views expressed could not be attributed to any identifiable person. Interested participants may request a copy of the published report or paper by contacting any of the researchers.

Interviews von Bezugspersonen durchgeführt

Interviews were conducted by the service users' key workers rather than an external researcher, building on the pre-existing therapeutic relationship and its foundation of trust and familiarity. This approach was chosen because service users were expected to speak more openly with a familiar person, particularly given the sensitive nature of the topics. Introducing an unfamiliar researcher could have created barriers to disclosure and compromised the depth and authenticity of responses.

2.3.3 Service user acquisition

Interviews mit 9 Personen aus 3 Ländern

Nine service users participated in the interview study – three from each site (Austria, Norway, Poland). They were consecutively recruited face-to-face within the outpatient clinics by their practitioners (clinical health personnel, (clinical) psychologists, or research assistants). Service users received an information sheet, consent form, and the researchers' contact details. Interviewers explained the study, answered questions, and arranged interview appointments at the clinic.

2.3.4 Informed consent procedures

schriftliche Einwilligung der interviewten Personen

All service users gave written informed consent before participating. Prior to consent, they were fully informed via a participant information sheet about the study's character, importance, relevance, and potential consequences. The consent form was dated and signed by the participant. In Austria and Poland, participants received a copy of the consent and information forms; in Norway, copies were provided upon request. The original signed documents were securely stored by the investigator.

2.3.5 Potential risks to the service users

ev. unangenehme Erinnerungen …

The only risk associated with participation was the inconvenience of responding to a semi-structured interview regarding SUD. All participants were informed that the study sought their opinions, and that no responses were considered "right" or "wrong". Differences in opinion were expected and could reflect various factors, including individual experiences with SUD support services.

As interviewees were actively in therapeutic or psychosocial care, the risk was estimated to be low. Any memories arising during interviews could be addressed in subsequent group or individual therapy. Additionally, debriefing was provided, as the interviewers were social workers or (clinical) psychologists. Participants were free to end the interview or withdraw at any time, decline to answer any question, or stop the recording if they wished to share something off the record.

… welche von klinischen Psycholog:innen/ Sozialarbeiter:innen abgefangen wurden

This study investigated illegal behaviour, such as the use of illegal drugs, among patients with SUD across all genders, socioeconomic statuses, and cultural backgrounds. No legal duty to report to authorities arose from this investigation.

trotz illegaler Thematiken, keine gesetzliche Meldepflicht

The benefits likely outweigh the risks, as improved understanding of patient-relevant outcomes may inform the selection of outcome parameters relevant for research, policy, and people with lived experience, ultimately aiding in the selection of appropriate measurement instruments.

gemeinsames Verständnis der Perspektiven

Additionally, participation may have allowed people with lived experience to be heard and have their experiences taken seriously, while contributing to professional development through knowledge sharing and discussions of best practices for measuring outcomes (e.g., addiction severity, recovery, substance use, craving) in patients with SUD.

Erfahrungsberichte werteten Wissenschaftsbeitrag auf

2.3.6 Data management

Data were stored electronically with password protection and were accessible only to the study's investigators and researchers. Personal data, such as date of birth or address, was not collected; names were replaced with codes, leaving the consent form as the only document containing identifiable data. Each transcript received a unique identifier, and participant identities were never shared in publications or communications.

Datenverwaltung: personenbezogene Daten geschützt

Data were stored on password-protected computers that complied with the General Data Protection Regulation, securely retained for five years after the project's completion, and then destroyed to ensure confidentiality and regulatory compliance.

Datenschutz-Grundverordnung war erfüllt

Research findings will be made publicly available through project reports/ manuscripts, conference presentations, and journal publications. De-identified findings were also shared via the public dissemination strategy, including the FALCO website and social media channels.

Forschungsergebnisse werden durch Publikationen veröffentlicht

Service users' anonymity was maintained throughout. All local data protection requirements were adhered to, and study findings and documents were treated as confidential. No patient-related data were disclosed to any other organisation.

Anonymität war gewahrt

keine Datenweitergabe

2.3.7 Interview guide

The interview guide was systematically derived from the findings of the systematic review (Part 1). The categories derived from the identified parameters across the included studies were used to structure the interview into thematic sections, covering all key clinical and socioeconomic outcome parameters, with each question mapping onto a specific category. This approach en-

halbstrukturierter Interviewleitfaden aus den Kategorien der systematischen Übersichtsarbeit abgeleitet

sured a systematic analysis of outcome parameters while directly addressing the categories captured by the measuring instruments and allowing to assess whether these constructs appeared in service users' lived experience. Importantly, the interviews were conceived as an add-on to the systematic review rather than as a validation instrument. The following semi-structured interview guide was used in all three countries:

Time before starting a therapy

- How did problematic substance use affect your life before you decided to go into treatment?
- How did your problematic substance use affect your surroundings?
- How has problematic substance use affected your mental health?
- What were the effects of problematic substance use on
  a) familial relationships/couple relationship/children,
  b) social exclusion,
  c) job situation (e.g. loss of job),
  d) financial troubles (e.g. healthcare costs), and
  e) other socioeconomic aspects (e.g. crime, homelessness)?
- How do you think participating in music therapy might influence your ability to return to work or maintain a stable daily routine?[4]
- What were your own efforts when managing to quit problematic substance use?
- What role(s) does music play in your life?

While therapy

- Why did you decide to quit substance use?
- Can you tell me about your experiences with substance use treatments/follow-up services? What were the most helpful and the most challenging aspects of these experiences? How has the period without using substances affected your life?

Long-term aspects

- What have been the most important aspects for you to stay abstinent over time?
- Has your social life changed over time? How?
- Has your position/participation in society changed over time? How?

2.3.8 Quality assurance

Pilotierung des Interviewleitfadens

The interview guide was piloted (FM) on the 25th of March 2025, interviewing one Austrian service user. Some minor changes were made, mainly to use simplified language and add examples.

[4] Participants may have varying degrees of familiarity with MT, e.g., prior personal experience, knowledge from peers, or existing opinions about MT. To address this, the interview guide was piloted prior to data collection to ensure accessibility of MT. To ensure reflexivity, interviewers were trained to probe for clarification where needed while avoiding leading responses.

2.3.9 Qualitative content analysis

qualitative Inhaltsanalyse:

This chapter outlines the methodological approach for analysing qualitative data from nine semi-structured interviews across three countries. Transcripts were produced verbatim in the original language of each interview and subsequently translated into English prior to analysis (see 2.3.2). Qualitative content analysis was chosen as the primary method for systematically analysing the textual data and deriving meaningful insights.

ein systematischer und methodengeleiteter Ansatz zur Analyse von Texten …

Qualitative content analysis is a systematic approach to analysing qualitative data, reducing complex information into manageable categories while maintaining the essential meaning and context of original statements. This method enhances transparency, reliability, and validity, making it well-suited to exploring complex social phenomena using interview data [55, 56].

… folgt einem systematischen und deduktiven-induktiven Verfahren

The analysis follows a systematic deductive-inductive procedure, drawing on prior knowledge from the systematic review (Part 1) while remaining open to emergent phenomena [55, 56]:

- Material selection and preparation
- Initial reading of the transcripts, familiarisation, taking notes/memos
- Category development: categories were deductively developed (LG) based on the systematic literature review (Part 1) and checked by another researcher (FM) regarding the inter-coder reliability

Kategorienentwicklung

- During systematic transcript review, text segments were assigned to these categories following an inductive coding process: within each category, subcategories were inductively derived from the data to capture emergent themes and patterns (LG)
- Category definitions and coding rules were reviewed and refined iteratively during the analysis process
- A qualitative data analysis software (MAXQDA) was used to facilitate the coding process (e.g., creating summary grids, summary tables) and to enable efficient organisation, retrieval, and management of coded segments. AI-assisted tools within MAXQDA were used solely to generate a few initial summaries of coded segments, which were thoroughly reviewed, validated, and refined by a researcher (LG). All coding, categorisation, and interpretation were conducted independently without AI assistance. The coding tree is available upon request.

qualitative Inhaltsanalyse in MAXQDA

- Evaluation and refinement, including reviewing categories, highlighting prominent themes (and less emphasised areas) by examining frequency and density of codes within categories
- Interpretation of categorised data to identify overarching themes and patterns
- Presentation of findings: The analysis followed a two-step approach. *First*, a manifest content analysis was conducted, staying close to the explicit interview content to ensure transparency and traceability. Original quotes were included to preserve interviewees' voices and lend authenticity. This lower level of interpretation reduces vulnerability to criticism regarding researcher subjectivity – particularly relevant as the interviews were not conducted by the analysing researcher. Following a quantitative frequency analysis (see 3.3.13), the most frequently mentioned category was identified. In the *second* step, this category underwent a latent content analysis – including metaphor analysis and analysis of absences – to explore underlying meanings, pat-

zweistufige Analyse: manifeste und latente Inhaltsanalyse

terns, and connections beyond what was explicitly stated. This interpretive approach yields deeper analytical insights, revealing relationships and contradictions that service users may not have been consciously aware of. The resulting findings offer a more condensed understanding of the most salient theme from service users' perspectives.

Kontingenzanalyse

- A contingency analysis was conducted in MAXQDA using summary tables generated for each interview to systematically examine co-occurrences among the developed codes. This analysis identified content-related connections between subcategories across the nine interviews.

Häufigkeitsanalyse

- A frequency analysis was performed based on the same summary tables. The relative frequency of each category was visualised using a code cloud (main categories) and a bar chart (top 10 subcategories). Frequency was operationalised as the number of coded segments per category across the nine interviews of the summary table. It should be noted that the frequency reflected the overall prominence of a theme in the summarised material.

latente Inhaltsanalyse

- The latent content analysis of the most prominent main category was conducted in a structured, researcher-led process. The methodological literature on latent analysis techniques was first reviewed to identify suitable analytical dimensions, such as metaphor analysis, absence analysis, and examination of linguistic registers. The researcher's deep familiarity with the data, developed through prolonged engagement across all prior analytical stages, informed the selection of these dimensions. On this methodological basis, the researcher used Claude Pro as an AI-assisted tools[5] to help identify latent patterns, implicit linkages, and metaphorical structures across the summary tables. All AI-generated outputs were subsequently validated against the original raw data, and all interpretive judgments remained with the researcher throughout.
- Participant feedback was not integrated into the analysis to maintain analytical independence and prevent post hoc rationalisation, consistent with the study design, which was a descriptive analysis with no planned participant involvement in interpreting the results.

[5] Data security was ensured through multiple measures: data were fully anonymised prior to analysis; only anonymised summary tables, rather than raw transcripts, were used as input.

3 Results

The results are presented in three parts: Part 1 provides an overview of relevant clinical and socioeconomic outcome parameters and tools assessing (serious) adverse events (RQ1). Part 2 focuses on the characteristics of measuring instruments identified through the selection process (see 2.2.6) (RQ2), and Part 3 summarises the interview study results, i.e., the service user involvement (RQ2).

klinische und sozioökonomische Ergebnisparameter, Merkmale von Messinstrumenten und Interviews

3.1 Results of Part 1: Instruments measuring clinical and socioeconomic outcome parameters and tools assessing (serious) adverse events

From the systematic search, seven articles were included for data extraction. We identified more than 60 outcomes in total and categorised them into 11 parameters. The manual search identified 13 additional tools that measure approximately 20 socioeconomic outcomes. Most outcomes were found for the 'music-, therapist-, and treatment-related parameters' (see Table A-1).

7 Übersichtsarbeiten aus systematischer Suche eingeschlossen

The clinical and socioeconomic outcomes were divided into the following 11 parameters:

in klinische und sozioökonomische Ergebnisparameter unterteilt

- Substance use parameters
- Recovery parameters
- Craving parameters
- Motivational parameters
- Quality of life parameters
- Mood-related parameters
- Psychological, psychiatric, physiological and cognitive parameters
- Cultural, spiritual, and locus of control parameters
- Biomedical parameters
- Music-, therapist-, and treatment-related parameters
- Socioeconomic and social parameters

From the manual search, we identified 14 tools that measure approximately ten socioeconomic outcomes (see Table A-2).

In several parameter categories, multiple instruments met the predefined selection criteria with comparable frequency, necessitating additional case-by-case decisions. These are documented here for transparency:

zusätzliche Auswahlentscheidungen bei mehreren geeigneten Instrumenten pro Parameter …

For 'psychological, psychiatric, physiological, and cognitive parameters', four instruments (SCL-90-R, BSI, CGI, FACT-Cog) met the criteria. The CGI was excluded due to discrepancies in validity and reliability [57-59]. The FACT-Cog was excluded as it was originally developed for oncology populations and focuses specifically on cancer-related cognitive impairment, making it unsuitable for the broader psychological and psychiatric symptom profile relevant to SUD populations. As the BSI is a short version of the SCL-90-R [60], only the SCL-90-R is described in detail.

… für psychologische, psychiatrische, physiologische und kognitive Parameter …

… und Motivationsparamter

Six instruments (CESI, CMR, ICR, RTCQ-TV, SOCRATES, URICA) were used with comparable frequency to measure motivational parameters. To ensure a manageable analysis, two were selected: the URICA and the ICR. The URICA is among the most frequently used continuous measures of motivation and readiness for change in SUD [61, 62] and, according to MT field experts, has been the standard for decades – originating in the substance abuse field ('stages of change') and underpinning many therapy models later extended to other areas. The ICR offers a quick, flexible, and practical way to assess and build core motivations within a single conversation.

Table 3-1 provides an overview of all identified tools that measure clinical and socioeconomic outcomes for each parameter.

Table 3-1: Overview of all identified tools measuring clinical and socioeconomic outcomes per parameter and (serious) adverse events

Clinical and socioeconomic parameters and outcomes	Measuring instruments
Substance use parameters	
(Drug avoidance) self-efficacy	**DASES**, **VAS**
Recovery parameters	
Problem severity	**ASI**
Adoption of principles and practices	**GAATOR 2.1**
Craving parameters	
(Intensity of) substance/alcohol craving	**BSCS**, **OCDUS**, **ACQ-SF-R(/-NOW)**, **AUQ**, **VAS**, **DMQ-R**, **DAQ**
Coping skills/Knowledge of triggers and coping skills	**SOC**, A-COPE[6]
Motivational parameters	
Motivation/readiness/eagerness for treatment/change	**CESI**, **CMR**, **ICR**, **RTCQ-TV**, **SOCRATES**, **URICA**
Motivation to reach/maintain sobriety	**CSS**, QMAD
Immediate goal attainment	GAF (goal)
Quality of life parameters	
Manageability, comprehensibility and meaningfulness of life/Orientation to life/Sense of coherence	**SOC**
Well-being	**FACIT-F**
Mood-related parameters	
Mood	**MDMQ**, **VAS**, **VAMS**
Depression	**BDI**, HRSD, **SDS**, Likert scale
Depressogenic thought frequency	**ATQ**
Positive and negative affect	**PANAS**
Sadness	**VAS**
Anxiety	**SAS**, **VAS**, **STAI**
Shame, pride and guilt	**SSGS**
Emotional valence and arousal	**GEMS-9**
Anger	**VAS**, NAI
Forgiveness and grief	**FGPS**
Psychological, psychiatric, physiological and cognitive parameters	
Self-consciousness	SCS
Constructive self-awareness and perceived purpose in life	PIL

[6] Tested, but not recommended.

Clinical and socioeconomic parameters and outcomes	Measuring instruments
Psychiatric symptoms	**SCL-90-R**
Physical and psychiatric symptoms	**BSI**
General functioning	**CGI**, GAF
Withdrawal (symptoms)	ARSW
Cognitive functioning	**FACT-Cog**
Cultural, spiritual, and locus of control parameters	
Cultural identity	ANCI
Comfort and strength derived from spirituality	**FACIT**
Locus of control	**I-E**, **DRIE**
Biomedical parameters	
Immune function	**Salivary Immunoglobulin A Test**
Music-, therapist-, and treatment-related parameters	
Attitudes towards MT/therapist/treatment groups	**VAS**
Trust in therapist	**WFPTS**
(Therapeutic) working alliance	**HAQ**, **WAI-S(R)**
Quality of therapist-client relationship	**HAQ**
Importance of MT	**VAMS**
Perceived effectiveness and enjoyment	Analogue scales
Intervention assessment compared to other groups	Analogue scales
Socioeconomic and social parameters	
Employment, occupation and income	**ASI**, **CSRI**, CSSRI, SES questionnaire, Duncan SEI, NS-SEC, CAPSES, Nam-Powers OSS
Productivity loss	**PECUNIA-RUM**
Occupational mental health	**WHO-HPQ**
Service use/receipt, use of care, health and social care	ESDS, CSI, **CSRI**, CSSRI, **UAC-I**, **PECUNIA-RUM**
Material capital (wealth, income, trust funds, etc.), human capital (skills, credentials, abilities, etc.), and social capital (instrumental relationships such as being friends with doctors and lawyers)	CAPSES, **SURE**
Education	SES questionnaire, **PECUNIA-RUM**, Duncan SEI, Nam-Powers OSS
(Criminal) justice, legal problem areas	**ASI**, **PECUNIA-RUM**
Relationship, family/social problem areas, family history, and informal care	**ASI**, **SURE**, SES questionnaire, **PECUNIA-RUM**
Usual living situation	**CSRI**, CSSRI
Retention in treatment (and completion)	Number of participants remaining at the end of treatment, number of days attending MT sessions
Attendance	Percent of attendance at any given time, number of sessions
Medication adherence	Number of sessions
Perception of social support	**MSPSS**
Perception of stigma	**PSAS**
Interpersonal (relationship) problems	**IIP-SC**
(Serious) adverse events	
(Serious) adverse events	No tools reported[7]

Bold: *Tested for validity/reliability*.

[7] One of the included seven studies reported that no (serious) adverse events occurred, and no specified measurement instruments for adverse event assessment were stated.

Abbreviations: A-COPE … Adolescent Coping Orientation to Problems Experienced. ACQ-SF-R(/-NOW) … Alcohol Craving Questionnaire-Short-Form-Revised (/-Now). ANCI … American Indian/Alaska Native Cultural Identity Scale. ARSW … Adjective Rating Scale for Withdrawal. ASI(-NAV) … Addiction Severity Index (Native American Version). ATQ … Automatic Thoughts Questionnaire. AUQ … Alcohol urge questionnaire. BDI … Beck Depression Inventory. BSCS … Brief Substance Craving Scale. BSI … Brief Symptom Inventory. CAPSES … Composite SocioEconomic Status scale. CESI … (Texas Christian University Treatment Motivation Scale) – Client Evaluation of Self at Intake. CGI … Clinical Global Impression Severity Scale. CMR … Circumstances, Motivation, and Readiness Scales for Substance Abuse Treatment. CSI … Cornell Services Index. CSRI … Client service receipt inventory. CSSRI … Client Socio-Demographic and Service Receipt Inventory. CSS … Commitment to Sobriety Scale. DAQ … (Multidimensional) Desires for Alcohol Questionnaire. DASES … Drug Avoidance Self-Efficacy Scale. DMQ-R … Drinking Motives Questionnaire – Revised. DRIE … Drinking-Related Internal-External Locus of Control Scale. ESDS … European Socio-demographic Schedule. FACIT … Functional Assessment of Chronic Illness Therapy: Spiritual Questions. FACIT-F … Functional Assessment of Chronic Illness Therapy: Fatigue. FACT-Cog … Functional Assessment of Cancer Therapy – Cognitive Functions. FGPS … Forgiveness Grief Perspectives Scale. GAATOR 2.1 … General Alcoholics Anonymous Tools of Recovery. GAF (goal) … Goal attainment form. GAF … Global Assessment of Function Scale. GEMS-9 … Geneva Emotions in Music Scale. HAQ(-II) … Revised Helping Alliance Questionnaire for therapist and client. HPQ … Health and Work Performance Questionnaire. HRSD … Hamilton Rating Scale for Depression. ICR … Importance, Confidence, Readiness (motivational) Ruler. I-E … Internal vs. external locus of control scale. IIP-SC … Inventory of Interpersonal Problems: Short Circumplex form. MDMQ … Multidimensional Mood Questionnaire. MSPSS … Multidimensional Scale of Perceived Social Support. MT … music therapy. NAI … Novaco Anger Inventory Short Form. NS-SEC … National Statistics Socioeconomic classification. OCDUS … Obsessive Compulsive Drug Use Scale. OSS … Occupational Status Score. PANAS … Positive and negative affect scale. PECUNIA-RUM … ProgrammE in Costing, resource use measurement and outcome evaluatioN for Use in multi-sectoral National and InternAtional health economic evaluations – Resource Use Measurement. PIL … Purpose in Life Test. PSAS … Perceived Stigma of Addiction Scale. QMAD … Questionnaire of Motivation for Abstaining from Drugs. RTCQ-TV … Readiness to Change Questionnaire – Treatment Version. SAS … Self-Rating Anxiety Scale. SCL-90-R … Symptom Checklist 90-R. SCS … Self-Consciousness Scale. SDS … Severity of Dependence Scale. SES … socioeconomic status. SEI … Socioeconomic Index. SOC … Sense of Coherence Scale. SOCRATES … Stages of Change Readiness and Treatment Eagerness Scale. SURE … Substance Use Recovery Evaluator. SSGS … State Shame and Guilt Scale. STAI … State-Trait Anxiety Inventory. UAC-I … Utilisation and Cost Inventory. URICA … University of Rhode Island Change Assessment Scale. VAMS … Visual Analogue Mood Scale. VAS … Visual Analogue Scale. WAI-S(R) … Working Alliance Inventory – Short (Revised). WHO … World Health Organisation. WFPTS … Wake Forest Physician Trust Scale.*

Wie werden Endpunkte gemessen?

Each of the following subchapters provides information on how the outcomes are measured and whether the measuring instruments have been tested for validity and reliability.

To improve the readability of the following results, Table 3-2 provides an overview of all measuring instruments in both their abbreviated and full forms.

Table 3-2: Acronym and full name of tools

Acronym	Full name of tool
A-COPE	Adolescent Coping Orientation to Problems Experienced
ACQ-SF-R(/-NOW)	Alcohol Craving Questionnaire-Short-Form-Revised (/ Now)
ANCI	American Indian/Alaska Native Cultural Identity Scale
ARSW	Adjective Rating Scale for Withdrawal
ASI(-NAV)	Addiction Severity Index (Native American Version)
ATQ	Automatic Thoughts Questionnaire
AUQ	Alcohol urge questionnaire
BDI	Beck Depression Inventory
BSCS	Brief Substance Craving Scale
BSI	Brief Symptom Inventory
CAPSES	Composite SocioEconomic Status scale
CESI	Client Evaluation of Self at Intake

Acronym	Full name of tool
CGI	Clinical Global Impression Severity Scale
CMR	Circumstances, Motivation, and Readiness Scales for Substance Abuse Treatment
CSI	Cornell Services Index
CSRI	Client service receipt inventory
CSS	Commitment to Sobriety Scale
CSSRI	Client Socio-Demographic and Service Receipt Inventory
DAQ	Desires for Alcohol Questionnaire
DASES	Drug Avoidance Self-Efficacy Scale.
DMQ-R	Drinking Motives Questionnaire – Revised
DRIE	Drinking-Related Internal-External Locus of Control Scale
Duncan SEI	Socioeconomic Index
ESDS	European Socio-demographic Schedule
FACIT	Functional Assessment of Chronic Illness Therapy: Spiritual Questions
FACIT-F	Functional Assessment of Chronic Illness Therapy: Fatigue
FACT-Cog	Functional Assessment of Cancer Therapy – Cognitive Functions
FGPS	Forgiveness Grief Perspectives Scale
GAATOR	General Alcoholics Anonymous Tools of Recovery
GAF	Global Assessment of Function Scale
GAF (goal)	Goal attainment form
GEMS	Geneva Emotions in Music Scale
HAQ	Revised Helping Alliance Questionnaire for therapist and client
HRSD	Hamilton Rating Scale for Depression
ICR	Importance, Confidence, Readiness Ruler
I-E	Internal vs. external locus of control scale
IIP-SC	Inventory of Interpersonal Problems: Short Circumplex form
MDMQ	Multidimensional Mood Questionnaire
MSPSS	Multidimensional Scale of Perceived Social Support
NAI	Novaco Anger Inventory Short Form
Nam-Power OSS	Occupational Status Score
NS-SEC	National Statistics Socioeconomic classification
OCDUS	Obsessive Compulsive Drug Use Scale
PANAS	Positive and negative affect scale
PECUNIA-RUM	ProgrammE in Costing, resource use measurement and outcome evaluatioN for Use in multi-sectoral National and InternAtional health economic evaluations – Resource Use Measurement
PIL	Purpose in Life Test
PSAS	Perceived Stigma of Addiction Scale
QMAD	Questionnaire of Motivation for Abstaining from Drugs
RTCQ-TV	Readiness to Change Questionnaire – Treatment Version
SAS	Self-Rating Anxiety Scale
SCL-90-R	Symptom Checklist 90-R
SCS	Self-Consciousness Scale
SDS	Severity of Dependence Scale
SES questionnaire	Socioeconomic status questionnaire

Acronym	Full name of tool
SOC	Sense of Coherence Scale
SOCRATES	Stages of Change Readiness and Treatment Eagerness Scale
SSGS	State Shame and Guilt Scale
STAI	State-Trait Anxiety Inventory
SURE	Substance Use Recovery Evaluator
UAC-I	Utilisation and Cost Inventory
URICA	University of Rhode Island Change Assessment Scale
VAMS	Visual Analogue Mood Scale
VAS	Visual Analogue Scale
WAI-S(R)	Working Alliance Inventory – Short (Revised)
WFPTS	Wake Forest Physician Trust Scale
WHO-HPQ	World Health Organization – Health and Work Performance Questionnaire

Substance use parameters

DASES*, VAS*
(* = auf Zuverlässigkeit und Gültigkeit geprüft)

(Drug avoidance) self-efficacy was tested by the DASES [12, 20, 21] and VAS [21], which are both tested for reliability and validity (DASES [20], VAS [63]).

Recovery parameters

ASI*, GAATOR*

Recovery parameters could be found in one article [21]. **Problem severity** was assessed using the ASI-(NAV), and **adoption of principles and practices** by the GAATOR 2.1. Both are validated and reliable instruments (ASI [64], GAATOR 2.1 [65]).

Craving parameters

BSCS*, OCDUS*, ACQ*, AUQ*, VAS*, DMQ*, DAQ*, SOC*, A-COPE

We found craving parameters in five [11, 12, 20, 21, 50] of the seven included articles. **(Intensity of) substance/alcohol craving** was determined using the following instruments: BSCS [11, 12, 20, 21], OCDUS [11], ACQ-SF(-R/-NOW) [11, 12, 50], AUQ [21, 50], VAS [21, 50], DMQ-R [50], and DAQ [50]. **Coping skills/Knowledge of triggers and coping skills** were measured using the SOC and the A-COPE [21]. The following tests are tested for validity and reliability: BSCS [66, 67], OCDUS [68-70], ACQ-R/-NOW [71, 72], VAS [63], DMQ-R [73], AUQ [74, 75], DAQ [76-78], and SOC [20].

Motivational parameters

CESI*, CMR*, ICR*, RTCQ-TV*, SOCRATES*, URICA*, CSS*, QMAD, GAF

Five studies [11, 12, 20, 21, 49] reported motivational parameters. Measurements of **motivation/readiness for treatment/change** were taken using the CESI [11, 12, 20, 21], CMR [11, 12, 20, 21], ICR [11, 20, 21, 49], RTCQ-TV [11, 12, 20, 21], SOCRATES [11, 12, 20, 21], and URICA [11, 12, 20, 21]. **Motivation to reach/maintain sobriety** was assessed using the CSS [11] and QMAD [11]. The GAF (goal) served as the measure for **immediate goal attainment**. The following tests are tested for validity and reliability: CESI [20], CMR [20], ICR [20, 79], RTCQ-TV [20], SOCRATES [20], URICA [20, 80-82], and CSS [83].

Quality of life parameters

QoL was assessed in three articles [20, 21, 49]. The authors used the SOC [20, 49] to assess **manageability, comprehensibility and meaningfulness of life/ Orientation to life/Sense of coherence**, and the FACIT-F [21] to measure **well-being**. The SOC [20] and FACIT-F [73] are both tested for validity and reliability.

SOC*, FACIT-F*

Mood-related parameters

All studies except one [50] assessed mood-related parameters. **Mood** was evaluated by the MDMQ and VAS [21] as well as by the VAMS [21, 84]. The BDI(-II) [11, 12, 20, 21, 49], HRSD [11, 12, 21], and SDS [11] served as the measuring tool for **depression**; and the ATQ for assessing **depressogenic thought frequency** [21]. **Positive and negative affect** were assessed using the PANAS [59], and sadness using the 7-point VAS [21]. Measurements of **anxiety** were taken using the SAS [11], 7-point VAS [21], and the STAI [20, 21]. The SSGS [12] measured **shame, pride and guilt**. One study [21] assessed further outcomes: **emotional valence and arousal** using the GEMS-9; **anger** by the 7-point VAS and NAI; and **forgiveness and grief** with the FGPS [21]. The following tools are tested for reliability and validity: BDI [20], SDS [85], SAS [86, 87], MDMQ [88, 89], VAS [63], VAMS [90], ATQ [91, 92], PANAS [93, 94], STAI [20], GEMS-9 [95, 96], FGPS [97], and SSGS [98, 99].

MDMQ*, VAS*, VAMS*, BDI*, HRSD, SDS*, ATQ*, PANAS*, SAS*, STAI*, SSGS*, GEMS-9*, NAI, FGPS*

Psychological, psychiatric, physiological and cognitive parameters

We identified psychological, psychiatric, physiological, and cognitive parameters across three studies [12, 20, 21]. The SCS and PIL were used to determine **self-consciousness** and **awareness and perceived purpose in life**, respectively [12]. **Withdrawal (symptoms)** were tested with the ARSW [12, 20, 21]. The SCL-90-R measures **psychiatric symptoms**, and the BSI measures **physical and psychiatric symptoms** [21]. **General functioning** was assessed using the CGI and GAF; and **cognitive functioning** by the FACT-Cog [21]. The following instruments are tested for reliability and validity: SCL-90-R [100, 101], BSI [102-104], CGI [57-59], and FACT-Cog [105-107].

SCS, PIL, ARSW, SCL-90-R*, BSI*, CGI*, GAF, FACT-Cog*

Cultural, spiritual, and locus of control parameters

Three studies [12, 20, 21] evaluated cultural, spiritual, and locus-of-control parameters. The ANCI served as a measure of **cultural identity** [21]. **Comfort and strength derived from spirituality** were measured by the FACIT [21], and **locus of control** by the DRIE [12, 20, 21] and I-E [21]. The FACIT [108], I-E [109], and DRIE [20] are tested for reliability and validity.

ANCI, FACIT*, DRIE*, I-E*

Biomedical parameters

Two articles evaluated biomedical parameters [20, 49]. Salivary Immunoglobulin A Tests were used [20, 49], a measure of **immune function**.

Speichel-Immunglobulin-A-Test

Music-, therapist-, and treatment-related parameters

Four articles determined music-, therapist-, and treatment-related parameters [12, 20, 21, 84]. **Attitudes towards MT/therapist/treatment groups** were measured using the VAS [84]. **(Therapeutic) working alliance** and the **quality of therapist-client relationship** were measured by the HAQ(-II) [12, 20,

VAS*, HAQ*, WAI-S(R)*, WFPTS*, VAMS*, 25-Punkte-Analogskala

21]; **alliance** was also assessed using the WAI-SR [12] and WAI-S [20]. The WFPTS was used to measure **trust in therapists** [12, 20]. One study [21] also evaluated the **importance of MT** (by VAMS), **perceived effectiveness and enjoyment** (by 25-point analogue scales), and **intervention assessment compared to other groups** (by 25-point analogue scales) [21]. The VAS [63], HAQ(-II) [20], WAI-S [20], WFPTS [20], and VAMS [90] are tested for reliability and validity.

Tools assessing (serious) adverse events

No tools measuring (serious) adverse events were reported. In one [11] of the seven included studies, no (serious) adverse events occurred; however, no specified measurement instruments for adverse event assessment were stated.

1/7 Studien: keine Messinstrumente berichtet

Socioeconomic and social parameters

Employment, occupation and income were assessed using the ASI, CSRI, CSSRI, SES questionnaire, Duncan SEI, NS-SEC, CAPSES, and Nam-Powers OSS. The PECUNIA-RUM evaluated **productivity loss** and the WHO-HPQ **occupational mental health**. **Service use/receipt, use of care, health and social care** were measured using the ESDS, CSI, CSRI, CSSRI, UAC-I, and PECUNIA-RUM. **Material capital (wealth, income, trust funds, etc.), human capital (skills, credentials, abilities, etc.), and social capital (instrumental relationships, such as friendships with doctors and lawyers)** were assessed by the CAPSES and SURE. The SES questionnaire, PECUNIA-RUM, Duncan SEI, and Nam-Powers OSS measure **education**. **(Criminal) justice and legal problem areas** were evaluated by the ASI and PECUNIA-RUM. **Relationship, family/social problem areas, family history, and informal care** were assessed using the ASI, SURE, SES questionnaire, PECUNIA-RUM and **usual living situation** using the CSRI and CSSRI. The UAC-I, SURE, CSRI, WHO-HPQ, and PECUNIA-RUM are tested for reliability and validity (see Table A-2).

ASI*, CAPSES, CSI, CSRI*, CSSRI, Duncan SEI, ESDS, Nam-Powers OSS, NS-SEC, PECUNIA-RUM*, SES questionnaire, SURE*, UAC-I*, WHO-HPQ*

All except one study [50] reported socioeconomic and social parameters. **Retention in treatment (and completion)** was measured by the number of participants remaining at the end of treatment [11] and the number of days attending MT sessions [21]. The number of sessions [84] and the percentage of attendance at any given time [54] were also used to assess the **attendance**. The number of sessions [21] also measured **medication adherence**. The MSPSS and PSAS served as the measures for **perception of social support** and **perception of stigma**, respectively [12]. **Interpersonal (relationship) problems** were determined using the IIP-SC [20, 49]. The MSPSS [110-113], PSAS [114, 115], and IIP-SC [20] are tested for reliability and validity.

Anzahl der Teilnehmer am Ende der Behandlung, Anwesenheitsquote in Prozent, Anzahl der Sitzungen, MSPSS*, PSAS*, IIP-SC*

3.2 Results of Part 2: Review of measuring instruments and their characteristics

Analyse der Messinstrumente und ihre Merkmale

In this chapter, *firstly*, 14 measuring instruments were selected and then described in detail with respect to their general and application characteristics. *Secondly*, we focused on the appropriateness, feasibility, and user relevance of these frequently used, validated and reliable 14 instruments. *Finally*, we assessed whether these 14 tools are suitable for long-term use in patients with SUD.

3.2.1 Measuring instruments and their general and application characteristics

Messinstrumente nach Kriterien für Detailanalyse ausgewählt

For the first approach, the following 14 measuring instruments were chosen according to the selection criteria in Chapter 2.2.6:

Substance use parameters

- DASES (measuring drug avoidance self-efficacy/self-efficacy)

Recovery parameters

- ASI (measuring problem severity)

Craving parameters

- BSCS (measuring (intensity of) substance/alcohol craving)

Motivational parameters

- ICR (measuring motivation/readiness for treatment/change)
- URICA (measuring motivation/readiness for treatment/change)

Quality of life parameters

- SOC (measuring manageability, comprehensibility and meaningfulness of life/Orientation to life/Sense of coherence)

Mood-related parameters

- BDI (measuring depression)

Psychological, psychiatric, physiological and cognitive parameters

- SCL-90-R (measuring psychiatric symptoms)

Cultural, spiritual, and locus of control parameters

- DRIE (measuring locus of control)

Biomedical parameters

- Salivary Immunoglobulin A Test (measuring immune function)

Music-, therapist-, and treatment-related parameters

- HAQ(-II) (measuring (therapeutic) working alliance and quality of therapist-client relationship)

Socioeconomic and social parameters

- IIP-SC (measuring interpersonal (relationship) problems)
- PECUNIA-RUM (measuring resource use in all relevant sectors for costing from a societal perspective in the adult population: health and social care, education, (criminal) justice, productivity)
- CSRI (measuring economic data: sociodemographic information, usual living situation, employment and income)

The *general characteristics* (i.e., outcome measured, mode of assessment, validation, post-assessment in treatment and latest FU where measuring instrument was applied, number and types of languages in which instruments are available, type of MT intervention and setting of assessment where instrument was applied, and limitations) and *application characteristics* (i.e., number of items in each instrument, total score range, interpretation of score and minimum clinically important difference (MCID), testing time, instrument user, and costs and regulations for use) of the selected measuring instruments are extracted and described in the following sections.

allgemeine Merkmale und Anwendungsmerkmale extrahiert

General characteristics

Outcome measured

The following outcomes were measured by the selected instruments: (drug avoidance) self-efficacy (DASES [12, 20, 21]), problem severity (ASI [21]), (intensity of) substance/alcohol craving (BSCS [11, 12, 20, 21]), motivation/readiness for treatment/change (URICA [11, 12, 20, 21], ICR [11, 20, 21, 49]), manageability, comprehensibility and meaningfulness of life/orientation to life/sense of coherence (SOC [20, 49]), depression (BDI [11, 12, 20, 21, 49]), psychiatric symptoms (SCL-90-R [21]), locus of control (DRIE [12, 20, 21]), (therapeutic) working alliance and quality of therapist-client relationship (HAQ(-II) [12, 20, 21]), interpersonal (relationship) problems (IIP-SC (IIP-32) [20, 49]), immune function (Salivary Immunoglobulin A Test [20, 49]), resource use (PECUNIA-RUM [116]), and economic data (CSRI).

z. B. Selbstwirksamkeit, Problemschwere, Verlangen, Behandlungsbereitschaft, Lebenssinn, Depression, Beziehungsqualität, Immunfunktion, Ressourcennutzung, ökonomische Daten

Mode of assessment

All used tools were questionnaires, except for two: the visual analogue scale, a short rating instrument (ICR [13, 45, 59, 60], and salivary samples (salivary Immunoglobulin A Test [20, 49]).

Fragebögen, visuelle Analogskala, Speichelproben

Post-assessment in treatment and latest FU where measuring instrument was applied (of included reviews)

In the included studies from the systematic literature search, the longest duration over which the measuring instrument was applied was two years (10 sessions) [45], and the latest FU where the measuring instrument was applied was three months [11].

Behandlungsdauer: bis zu 2 Jahre (10 Sitzungen), Nachbeobachtungszeit: bis zu 3 Monate

Number of languages in which instruments are available

The SOC is translated into a minimum of 33 languages [85, 86], the BDI into a minimum of 18-20 languages [109], and the ASI into a minimum of 13 languages. The other tools are translated into between two (BSCS and DRIE) and eight languages (SCL-90-R, CSRI). The PECUNIA-RUM will be available in multiple languages in the future [117]. As the ICR is a visual analogue scale, no translation is needed. For the Salivary Immuno-globulin A Test, a translation is not applicable.

Übersetzungen in bis zu 33 Sprachen (SOC)

Types of languages in which instruments are available

The following Table 3-3 presents the languages in which instruments are available according to the geographical location.

Übersetzungen in vielen europäischen Sprachen und darüber hinaus

Table 3-3: Types of languages in which instruments are available

	Tools	Languages
European languages	DASES	English, Turkish
	ASI	English, Czech, Danish, Dutch, French, Hungarian, Italian, Lithuanian, Polish, Portuguese, Russian, Spanish, Swedish, German
	BSCS	English, Bulgarian
	URICA	English, German, Norwegian, Polish, Portuguese, French, Swedish
	ICR	All
	SOC	Croatian, Czech, English, German, Hungarian, Icelandic, Lithuanian, Norwegian, Polish, Russian, Slovenian, Spanish, Swedish, Turkish, Welsh
	BDI	English, Spanish, Dutch, Finnish, German, Polish, Swedish, Turkish
	SCL-90-R	English, French, Spanish, German, Russian, Italian, Portuguese, Dutch, Swiss German, Hebrew
	DRIE	English
	HAQ(-II)	English, Spanish, French
	IIP-SC (IIP-32)	English, Dutch, Spanish, German, Italian, Swedish, Polish, Finnish, French, Greek, Slovenian
	PECUNIA-RUM	English, Dutch, German
	CSRI	English, German, Danish, Spanish, Italian, Dutch
Non-European languages	DASES	Farsi/Persian, Malay, Turkish, Chinese
	ASI	Arabic, Russian
	BSCS	Arabic
	URICA	Chinese, Portuguese (Brazil), Arabic
	ICR	All
	SOC	Indonesian, Japanese, Hebrew, Portuguese (Brazil), Russian, Turkish
	BDI	French (Canadian), Korean, Chinese, Arabic, Hindi, Marathi, Tamil, Malayalam, Japanese, Persian, Kannada, Xhosa, Turkish
	SCL-90-R	Hindi, Kannada, Tamil, Marathi, Malayalam, Japanese, Chinese, Korean, Vietnamese, Arabic, Russian, Hebrew
	DRIE	Chinese
	IIP-SC (IIP-32)	Chinese, Arabic, Korean, Malay
	CSRI	Portuguese (Brazil)

Type of MT intervention in which the instrument was applied (active/receptive/both)[8]

Messinstrumente für aktive und rezeptive musiktherapeutische Interventionen

All measuring instruments were used during both active and receptive MT interventions. One study [44] additionally used the following measuring instruments only during *receptive* MT interventions: ICR, SOC, BDI, Salivary Immunoglobulin A Test, and IIP-SC.

Setting of assessment where instrument was applied

häufigsten Messinstrumente nach Setting

The following measuring instruments were most often applied in inpatient detoxification settings [11, 12, 20, 21]: BSCS, SOCRATES, URICA, RTCQ-TV, CMR, CESI, BDI(-II), HRSD, ARSW, and DRIE. In the setting of inpa-

[8] It should be noted that the included reviews did not consistently distinguish between MT subtypes (improvisational, compositional, recreative, receptive [241]), reporting primarily at the level of active versus receptive delivery.

tient abuse treatment and rehabilitation centres as well as therapeutic communities for SUD, tools with a broad, holistic approach were used, including addiction-specific tools and measuring instruments for recording psychosocial factors, therapeutic processes and QoL [12]. In residential substance use treatment facilities and chemical dependency treatment programmes, the emphasis was on tools measuring addiction-specific and behaviour-oriented parameters [11].

Limitations

Self-report bias, cross-cultural applicability/sensitivity, and validation needs of specific groups

Verzerrung des Selbstberichts, kulturell angepasste Versionen mit psychometrischen Einschränkungen, kontinuierliche Validierung nötig, Empfindlichkeit gegenüber sprachlichen und kulturellen Faktoren, Berücksichtigung der klinischen Bedürfnisse spezieller Gruppen

Nevertheless, several limitations were identified for the selected measuring instruments. One limitation is self-report bias, in which individuals may underestimate or exaggerate symptoms (DASES, BSCS, URICA, SOC, BDI, SCL-90-R, DRIE, HAQ(-II), IIP-SC (IIP-32), PECUNIA-RUM). Culturally adapted versions also exhibit psychometric limitations, for example, lower consistency (ASI [118]). The BDI is not specific to any one culture, so its accuracy may vary across cultural contexts without specific validation [119]. Also, for the PECUNIA-RUM, some categories might be too broad to capture country-specific service nuances, as the tool is designed for international use [117]. Measuring instruments require ongoing validation, particularly confirmatory factor analysis in diverse populations to ensure cross-cultural applicability and generalisability (e.g., SOC [120]). Findings from culturally adapted versions should be viewed as preliminary and require future research to determine their impact on long-term treatment outcomes (e.g., DASES [121]). The ASI is sensitive to linguistic and cultural factors, necessitating replication and revalidation across settings, languages, and cultures [122]. For the ASI, challenges in adequately representing the clinical needs of special groups (e.g., women, ethnic minorities, mental/physical health problems) are reported [123]. Some tests are restricted to a specific group; for example, the DASES is restricted for use in young multiple drug users (aged 16-30) [124]. The ICR may not resonate with everyone as a natural way to conceptualise mixed feelings about change [125].

Practical limitations: training intensity and skills

hoher Schulungsaufwand (ASI), Wirksamkeit abhängig von Fähigkeiten von Interviewer:innen (ICR)

The ASI requires intensive training (2 full days) [123]. Reliability decreases significantly when performed by interviewers who are less intensively trained and monitored [123]. Similarly, the effectiveness of the ICR is highly dependent on the helper's proficiency in Motivational Interviewing skills [125], as trained interviewers can best elicit accurate and comprehensive information from the CSRI [126].

Structural aspects with psychometric consequences

strukturelle Limitationen z. B. Mix aus Items (DRIE) oder schwacher Zusammenhang von Items (SOC) können zu psychometrischen Konsequenzen führen …

…wie z. B. begrenzte Darstellung von Items (BSC)

For the ASI, issues with composite scores commonly used in treatment outcome studies are reported as well as problems in specific domains (e.g., medical domain, employment and self-support, family and social relationships, psychiatric problems) [123]. The DRIE is criticised for its mixture of both personalised and general statements [127]; changes in scores during treatment may be confounded by the fact that personalised items may not be sensitive when alcohol is no longer consumed, limiting their responsiveness in later recovery stages [127]. The BSCS may offer a limited representation of the multidimensional nature of craving (e.g., reward, relief, and obsessive craving) [128]. Alliance levels, as measured by the HAQ(-II), are not associated with

pretreatment psychiatric severity or level of depression [129]. The relationship between SOC and physical health is not strong, potentially due to mental and emotional factors being intertwined in the scale [120]. For the PECUNIA-RUM, the development was guided by existing RUM instruments only to a limited extent [117]. This might reduce content validity and comparability with other studies.

fragwürdige konvergente Validität (URICA), mangelnde Korrelation (BSCS), eingeschränkte interne Konsistenz und strukturelle Inkonsistenz (IIP-SC)

Regarding the psychometric properties of the tools, for example, the URICA exhibits questionable convergent validity [80]. The BSCS has a lack of significant correlation with objective behavioural outcomes, such as actual cocaine use [130]. For the IIP-SC (IIP-32), it is stated that brevity comes with an inherent trade-off in internal consistency, although it is still described as adequate [131]; a larger gap would be theoretically anticipated in the warm-dominant quadrant, suggesting a potential structural inconsistency in its circumplex representation [131].

Diagnostic limitations, scope of assessment, and economic constraints

Einschränkungen des BDI: z. B. schwankende Grenzwerte, Fehlnutzung als Diagnostikum statt Schweregradskala, begrenzte Erfassung (nur Symptome, keine Anamnese), statische Momentaufnahme

PECUNIA-RUM: generisches (nicht spezifisches) Design

Importantly, the optimal cutoff points (threshold values) that maximise the diagnostic utility of the BDI vary considerably depending on the sample studied [132]. The BDI is often misused as a diagnostic instrument, despite being designed as a rating scale to measure the severity of depression [133]. This tool has a limited scope, as the BDI only assesses symptoms of depression and does not cover other important factors like family history or genetic predisposition [134]. A single administration using the BDI provides a static measure, capturing a snapshot rather than dynamic symptom fluctuations over time [135]. Furthermore, the BDI, SCL-90-R, and IIP-SC (IIP-32) are copyrighted/licensed [115, 136, 137], requiring a fee for each copy used, which can probably limit accessibility, even if they are free to download. For the PECUNIA-RUM, a generic versus specific design trade-off may arise, as it is designed for international applicability rather than optimised for specific contexts, potentially sacrificing precision [117].

Confounding factors and biological variability in salivary Immunoglobulin A testing

Limitationen wie z. B. störanfällig durch Nahrung/Getränke, bei IgA-Mangel unzuverlässig, multiple biologische/ Lifestyle-Störfaktoren (Tagesrhythmus, Infektionen/Medikamente)

Salivary Immunoglobulin A Test exhibits significant diurnal variation, with levels peaking in the morning and gradually declining throughout the day, necessitating precise timing of the sample collection [138]. Reliability may be affected by immediate consumption of food, drink, and alcohol prior to testing [139]. Selective Immunoglobulin A (IgA) Deficiency (SIgAD), a condition in which individuals lack IgA, can naturally affect test results [140]. The Salivary Immunoglobulin A Test is highly susceptible to numerous confounding biological and lifestyle factors (e.g., recent infection, antibiotic intake, caffeine intake, food or drink, daily rhythm) [141].

Application characteristics

Number of items in each instrument

von Kurzskalen (ICR) bis zu umfassenden (SCL-90-R) oder variablen (PECUNIA-RUM, CSRI) Instrumenten

The selected instruments can be categorised into three groups based on the number of items. *Brief* measuring instruments comprise the first group, ranging from the ultra-short ICR (2 items) to medium-length instruments like the DASES, BSCS (16 items each), and the HAQ-II (19 items). A second group consists of instruments with *moderate* length (21 to 32 items), including the BDI, DRIE, SOC, URICA, and the IIP-SC (IIP-32), which balance psychometric depth with manageable participant burden. In contrast, the SCL-90-R

is a comprehensive, broadband instrument that utilises 90 items to provide a multidimensional psychological profile. Distinct from these fixed-length scales is the ASI, a semi-structured interview where the number of items varies depending on the individual's problem areas. For the PECUNIA-RUM and CSRI, the number of items in each instrument varies because users can select the relevant modules based on their needs. For the Salivary Immunoglobulin A Test, this section is not applicable.

Total score range

von schmalen (URICA) bis hin zu weitgefassten Score-Spannweiten (SCL-90-R)

The scoring systems of the selected instruments vary according to their clinical objectives and measurement depth. *Low-range* measuring instruments include the ICR, BSCS, and URICA, which utilise narrow score ranges (typically between -2 and 30) to provide quick indices of motivation or craving intensity. A second group comprises *moderate-range* instruments such as the DRIE, BDI, and ASI, with total scores typically ranging up to 64; notably, the ASI also includes a 0-9 severity rating for clinical assessment across its seven domains. *High-range* multidimensional tools require more complex scoring, with ranges extending from approximately 100 (DASES, HAQ-II) to over 200 (SOC), peaking with the SCL-90-R, which has a maximum score of 360 to reflect its broad symptomatic coverage. In contrast, the IIP-SC (IIP-32) reaches 128, while the IgA test provides a purely biological measure with a clinical reference range of 4 to 37 mg/dL [142]. As the PECUNIA-RUM and CSRI are not scoring instruments, this section is not applicable.

Interpretation of score and minimum clinically important difference (MCID)

schwerere Symptomatik bedeutet höherer Wert bei z. B. ASI, BSCS, BDI, SCL-90-R

Higher score expresses a *better outcome* for the following measuring instruments:

- DASES (higher self-efficacy in avoiding substance use),
- URICA/ICR (greater motivation for change),
- SOC (greater ability to manage tension, perceive life as comprehensible, manageable, and meaningful), and
- HAQ(-II) (stronger therapeutic alliance);

and a less favourable outcome for the ASI, BSCS, BDI, SCL-90-R, DRIE (stronger external locus of control, meaning that external factors, rather than personal control, are more influential in substance use), IIP-SC (IIP-32) (greater interpersonal problems), and Salivary Immunoglobulin A Test (higher levels of antibodies).

minimal für Patient:in bedeutsamer Unterschied nur für BDI

An MCID was found only for the BDI [11, 12, 20, 21, 49]. No MCID was detected in the SUD population. However, for depression, the literature reports the smallest important change of 17.5% reduction, indicating a meaningful change in depression management. For people with a longer duration of depression who had not responded to antidepressants, an MCID of 32% reduction was reported [143].

Testing time

Dauer der Testung: wenige Minuten bis 1 Stunde

Most testing times range between five and 15 minutes (DASES, BSCS, SOC, BDI, SCL-90-R, URICA, DRIE, HAQ(-II), IIP-SC (IIP-32), Salivary Immunoglobulin A Test). The PECUNIA-RUM and CSRI take between 15 and 30 minutes. The shortest test is the ICR, which lasts a few minutes; the most comprehensive one takes 60 minutes (ASI).

Instrument user

meist selbstberichtete Messinstrumente

Most tests are self-reported tools (DASES, BSCS, URICA, SOC, BDI, SCL-90-R, DRIE, IIP-SC (IIP-32), PECUNIA-RUM). The HAQ(-II) is a self-administered tool; however, the client and therapist complete it independently. All other tests are clinician-rated (ASI, ICR, Salivary Immunoglobulin A Test) or interviewer-administered (CSRI) measuring instruments.

Costs and regulations for use

manche Messinstrumente lizenziert (BDI, SCL-90-R, and IIP-SC (IIP-32))

All the selected measuring tools are available for free download, except the PECUNIA-RUM. However, the BDI, SCL-90-R, and IIP-SC (IIP-32) are protected by copyright and require permission or a license to use them (i.e., copyrighted/licensed) [115, 136, 137]. The PECUNIA-RUM is free to use for non-commercial research, but registration and adherence to the PECUNIA methodology are required (Open Access). For the Salivary Immunoglobulin A Test, this section is not applicable.

3.2.2 Appropriateness, feasibility, and user relevance of the measuring instruments

Angemessenheit, Durchführbarkeit und Nutzerrelevanz der Messinstrumente

For answering, if the selected 14 most often used, validated and reliable measuring instruments are appropriate, feasible and user-relevant, detailed information was deduced from different sources (i.e., Table A-3, Table A-4, and service user interviews, see 3.3) as stated in Chapter 2.2.6. The following Table 3-4 and Table 3-5 give an overview:

Table 3-4: Overview of the appropriateness, feasibility and user-relevance of the 14 measuring instruments (part 1/2)

Parameters:	Substance use parameters	Recovery parameters	Craving parameters	Motivational parameters		Quality of life parameters
Measuring instruments	**DASES**	**ASI**	**BSCS**	**URICA**	**ICR**	**SOC**
Appropriateness and feasibility (deduced from Table A-3 and Table A-4)	– Self-report bias – Restricted use to young multiple drug users – Culturally adapted versions are preliminary + Self-reporting tool + <20 items + Testing time ≤10 minutes	– Culturally adapted versions with lower consistency – Intensive training needed to maintain reliability – Challenges in representing clinical needs of special groups – Possible problems with composite scores and in specific domains + Available in >10 languages	– Self-report bias – Lack of significant correlation with objective behavioural outcomes – Possible limited representation of the multidimensional nature of craving + Self-reporting tool + <20 items + Testing time ≤10 minutes	– Self-report bias – Questionable convergent validity + Self-reporting tool + Testing time ≤10 minutes	– Effectiveness depends on proficiency in Motivational Interviewing skills – Problems with "mixed feelings" + Self-reporting tool + Available in >10 languages + <20 items + Testing time ≤10 minutes	– Self-report bias – Problems with link between SOC and physical health – Requires ongoing validation + Self-reporting tool + Testing time ≤10 minutes + Available in >10 languages
User relevance (deduced from service user interviews)	Relevant according to the interviews (e.g., "self-determined life", "self-confidence", " self-awareness", "self-insights", "self-care", "self-image")	Relevant according to the interviews (e.g., "disturbed/ unworthy life", "physical appearance", "influenced consciousness", "experiences tension", "affected everything, my whole life, for a long period", "life was destroyed", "unable to use or live without drugs", "feeling helpless", "tired of living a lie", "alcohol clearly was the centre of my life")	Relevant according to the interviews (e.g., "craving", "inability to cope", "triggers")	Relevant according to the interviews (e.g., "motivation for change", "feeling ready", "decision to change and live a normal life")		Relevant according to the interviews (e.g., "life problems", "life improvements", "life transformation", "restarting life", "enjoy life", "daily functioning", "lifestyle changes")

– … *negative aspect.* + … *positive aspect.*

Abbreviations: *ASI … Addiction Severity Index. BSCS … Brief Substance Craving Scale. DASES … Drug Avoidance Self-Efficacy Scale. ICR … Importance, Confidence, Readiness (motivational) Ruler. SOC … Sense of Coherence Scale. URICA … University of Rhode Island Change Assessment Scale.*

Table 3-5: Overview of the appropriateness, feasibility and user-relevance of the 14 measuring instruments (part 2/2)

Parameters:	**Mood-related parameters**	**Psychological, psychiatric, physiological and cognitive parameters**	**Cultural, spiritual, and locus of control parameters**	**Music-, therapist-, and treatment-related parameters**	**Socioeconomic and social parameters**			**Biomedical parameters**
Measuring instruments:	**BDI**	**SCL-90-R**	**DRIE**	**HAQ(-II)**	**IIP-SC (IIP-32)**	**PECUNIA-RUM**	**CSRI**	**Salivary Immuno-globulin A Test**
Appropriateness and feasibility (deduced from Table A-3 and Table A-4)	– Self-report bias – Misuse as a diagnostic test (BDI rates severity) – Limited scope (BDI assesses symptoms) – Captures snapshot – Not culturally specific – Optimal cut-off points vary considerably across different samples + Self-reporting tool + Testing time ≤10 minutes + MCID available (not specifically for SUD) + Available in >10 languages	– Self-report bias – Sensitive to linguistic and cultural factors + Self-reporting tool	– Self-report bias – Criticised for mixture of personalised and general statements – Confounding of scores during treatment + Self-reporting tool + Testing time ≤10 minutes	– Self-report bias – Lack of association with pretreatment severity + Self-reporting tool + <20 items + Testing time ≤10 minutes	– Self-report bias – Brevity vs internal consistency – Structural inconsistency in the circumplex model + Self-reporting tool + Testing time ≤10 minutes	– Self-report bias – Generic design: too broad to capture country-specific nuances – Development: limitedly guided by existing RUM instruments + Self-reporting tool + Modular use (items)	– Trained interviewer needed + Modular use (items)	– Critical role of diurnal variation in data accuracy – Impact of ingestion on reliability – Natural confounding due to selective Immunoglobulin A deficiency – High susceptibility to confounding factors + Testing time ≤10 minutes
User relevance (deduced from service user interviews)	Relevant according to the interviews (e.g., "depression", "hopelessness")	Relevant according to the interviews (e.g., "anxiety", "aggression", "psychosis", paranoia", "pessimism", "panic attack", "suicidal thoughts", "mental sobriety/changes", "trauma")	Relevant according to the interviews (e.g., "loss of control", "sense of control")	Relevant according to the interviews (e.g., "long-term, trusting contact with outreach team member was crucial for recovery", "sense of freedom increased drug use when unsupervised", "positive therapist reactions")	Relevant according to the interviews (e.g., "relationship", "withdrawal from social life", "new social life", "parenthood", "family")	Relevant according to the interviews (e.g., "hospital/ emergency visit/ stay", "criminalities", "job/career", "finances", "school/ education")	Relevant according to the interviews (e.g., "homelessness", "hospital/ emergency visit/stay", "job/career", "finances")	Relevant according to the interviews (e.g., "depression", "stress", "infections", "mood", "abnormal blood values")

– … *negative aspect.* + … *positive aspect.*

Abbreviations: BDI … Beck Depression Inventory. DRIE … Drinking-Related Internal-External Locus of Control Scale. HAQ(-II) … Revised Helping Alliance Questionnaire for therapist and client. IIP-SC … Inventory of Interpersonal Problems: Short Circumplex form. MCID … minimum clinically important difference. SCL-90-R … Symptom Checklist 90-R.

Analyse der 14 Messinstrumente nach z. B. Limitationen, Anzahl der Items, Testdauer, Übersetzungen, Nennung der Endpunkte in Interviews

When analysing the appropriateness/feasibility and user relevance from Table 3-4 and Table 3-5, 11 measuring instruments were *self-reporting* tools. The ASI, BDI, and salivary test show the most *limitations* of the 14 assessed tools. The DASES, BSCS, ICR, and HAQ(-II) have fewer than 20 *items*. A *testing time* less than or equal to ten minutes requires the following ten tools: DASES, BSCS, URICA, ICR, SOC, BDI, DRIE, HAQ(-II), IIP-SC (IIP-32), and a salivary test. The ASI, ICR, SOC, and BDI are available in more than ten *languages*. The PECUNIA-RUM and CSRI can be *modularly* used. These aspects speak to the appropriateness and feasibility of the measuring instruments. Only for the BDI, an *MCID* is available, which is important for the clinical interpretation; however, it is not specifically for the SUD population. All measured parameters of the 14 tools were mentioned by the interviewees.

3.2.3 Suitability of measuring instruments for long-term use in patients with SUD

Eignung für Langzeitmessungen hinsichtlich 3 Dimensionen analysiert

Long-term measurements require a stable instrument (reliable over time) that is sensitive to clinically relevant changes and practical for repeated use. In this chapter, we assessed whether these 14 measuring instruments are suitable for long-term use in patients with SUD according to the criteria described in 2.2.6. For this purpose, we assessed the following three dimensions, which are summarised in the Table 3-6 and Table 3-7.

Table 3-6: Suitability for long-term measures (part 1/2)

Parameters:	Substance use parameters	Recovery parameters	Craving parameters	Motivational parameters		Quality of life parameters
Measuring instruments:	**DASES**	**ASI**	**BSCS**	**URICA**	**ICR**	**SOC**
Core measurement quality and clinical relevance						
Minimum clinically important difference (MCID) – essential for interpretation	**X**	**X**	**X**	**X**	**X**	X
Outcome measured – outcome construct stability	High ((drug avoidance) self-efficacy: stable but improves/worsens over time)	High (problem severity: stable, long-term metric)	Low to moderate ((intensity of) substance/alcohol craving: acute/short-term measure)	Moderate to high (motivation/readiness for treatment/change: relatively stable, yet definitely modifiable, construct)		High (manageability, comprehensibility and meaningfulness of life/ Orientation to life/Sense of coherence: stable, modifiable recovery outcome)
Post-assessment in treatment and latest FU where measuring instrument was applied (of included reviews) – empirical evidence	**X** Duration of treatment 6 months **X** Latest FU 1 month	**X** Duration of treatment 6 months **X** Latest FU 1 month	**X** Duration of treatment 6 months ✔ Latest FU 3months	**X** Duration of treatment 6 months ✔ Latest FU 3 months	✔ Duration of treatment 2 years ✔ Latest FU 3 months	✔ Duration of treatment 2 years **X** Latest FU 1 month
Limitations – risk assessment	✔ <4 limitations	**X** ≥4 limitations	✔ <4 limitations	✔ <4 limitations	✔ <4 limitations	✔ <4 limitations
Practicality and burden of application						
Testing time (min) – compliance and cost	✔ Testing time ≤10 minutes	**X** Testing time >10 minutes	✔ Testing time ≤10 minutes	✔ Testing time ≤10 minutes	✔ Testing time ≤10 minutes	✔ Testing time ≤10 minutes
Number of items in each instrument/Total score range – resolution and sensitivity	✔ >15 items ✔ range >50	✔ >15 items ✔ range >50	✔ >15 items **X** range ≤50	✔ >15 items **X** range ≤50	**X** ≤15 items **X** range ≤50	✔ >15 items ✔ range >50
Instrument user – feasibility	✔ (self-reported)	**X**	✔ (self-reported)	✔ (self-reported)	✔ (clinician-administered or self-completion format)	✔ (self-reported)
Mode of assessment – logistical feasibility	✔ Questionnaire	✔ Questionnaire (interview)	✔ Questionnaire	✔ Questionnaire	✔ Visual analogue scale, short rating instrument	✔ Questionnaire
Setting of assessment where instrument was applied – flexibility	✔ Flexible	✔ Flexible	✔ Flexible	✔ Flexible	✔ Flexible	**X** Semi-flexible (only short-term detoxification units [20] and NR [49])

Parameters:	Substance use parameters	Recovery parameters	Craving parameters	Motivational parameters		Quality of life parameters
Measuring instruments:	**DASES**	**ASI**	**BSCS**	**URICA**	**ICR**	**SOC**
Costs and regulations for use – sustainability	✔ Free to download	✔ Free to download	✔ Free to download	✔ Free to download	✔ Free to download	✔ Free
Generalisability and context						
Number of languages in which instruments are available – generalisability	**X** Available in <10 languages	✔ Available in >10 languages	**X** Available in <10 languages	**X** Available in <10 languages	✔ Available in >10 languages	✔ Available in >10 languages
Type of MT intervention in which the instrument was applied – contextual fit	✔ Both	✔ Both	✔ Both	✔ Both	✔ Both	✔ Both

✔ … *yes.* **X** … *no.*

Abbreviations: ASI … Addiction Severity Index. BSCS … Brief Substance Craving Scale. DASES … Drug Avoidance Self-Efficacy Scale. ICR … Importance, Confidence, Readiness (motivational) Ruler. SOC … Sense of Coherence Scale. URICA … University of Rhode Island Change Assessment Scale.

Table 3-7: Suitability for long-term measures (part 2/2)

Parameters:	Mood-related parameters	Psychological, psychiatric, physiological and cognitive parameters	Cultural, spiritual, and locus of control parameters	Music-, therapist-, and treatment-related parameters	Socioeconomic and social parameters			Biomedical parameters
Measuring instruments:	**BDI**	**SCL-90-R**	**DRIE**	**HAQ(-II)**	**IIP-SC (IIP-32)**	**PECUNIA-RUM**	**CSRI**	**Salivary Immuno-globulin A Test**
Core measurement quality and clinical relevance								
Minimum clinically important difference (MCID) – essential for interpretation	✔	**X**	**X**	**X**	**X**	NA	NA	X
Outcome measured – outcome construct stability	High (depression: stable, yet treatable, metric)	High (psychiatric symptoms: overall level of symptoms is a long-term metric)	High (locus of control: fundamental, relatively stable personality trait)	Low ((therapeutic) working alliance and quality of therapist-client relationship: in-session process variable, not stable)	High (interpersonal (relationship) problems: indicators of long-term social functioning and recovery success; problems tend to be chronic)	Low to Moderate (resource usage is highly dependent on the situation and influenced by external factors, e.g., cost changes; while it can be modified in the short term, it lacks long-term stability	Moderate to High (records actual, documented service usage retrospectively, which is objective/ stable; the metric can be used relatively consistently over a longer period of time, even if service usage remains modifiable through interventions.	Mixed (immune function: highly acute/short-term measures; overall immune health: stable, modifiable indicators)
Post-assessment in treatment and latest FU where measuring instrument was applied (of included reviews) – empirical evidence	✔ Duration of treatment 2 years ✔ Latest FU 3 months	**X** Duration of treatment 6 months **X** Latest FU 1 month	**X** Duration of treatment 6 months **X** Latest FU 1 month	**X** Duration of treatment 6 months **X** Latest FU 1 month	✔ Duration of treatment 2 years **X** Latest FU 1 month	NA	NA	✔ Duration of treatment 2 years **X** Latest FU 1 month
Limitations – risk assessment	X ≥4 limitations	✔ <4 limitations	✔ <4 limitations	✔ <4 limitations	✔ <4 limitations	✔ <4 limitations	✔ <4 limitations	**X** ≥4 limitations
Practicality and burden of application								
Testing time (min) – compliance and cost	✔ Testing time ≤10 minutes	**X** Testing time >10 minutes	✔ Testing time ≤10 minutes	✔ Testing time ≤10 minutes	✔ Testing time ≤10 minutes	**X** Testing time >10 minutes	**X** Testing time >10 minutes	✔ Testing time ≤10 minutes
Number of items in each instrument/ Total score range – resolution and sensitivity	✔ >15 items ✔ range >50	✔ >15 items ✔ range >50	✔ >15 items **X** range ≤50	✔ >15 items ✔ range >50	✔ >15 items ✔ range >50	Variable/Modular NA	Variable/Modular NA	NA

Parameters:	Mood-related parameters	Psychological, psychiatric, physiological and cognitive parameters	Cultural, spiritual, and locus of control parameters	Music-, therapist-, and treatment-related parameters	Socioeconomic and social parameters			Biomedical parameters
Measuring instruments:	**BDI**	**SCL-90-R**	**DRIE**	**HAQ(-II)**	**IIP-SC (IIP-32)**	**PECUNIA-RUM**	**CSRI**	**Salivary Immuno-globulin A Test**
Instrument user – feasibility	✔ (self-reported)	✔ (self-reported)	✔ (self-reported)	✔ (self-reported)	✔ (self-reported)	✔ (self-reported)	**X**	**X**
Mode of assessment – logistical feasibility	✔ Questionnaire	✔ Questionnaire	✔ Questionnaire	✔ Questionnaire	✔ Questionnaire	✔ Questionnaire	✔ Questionnaire (interview)	**X** Salivary sample
Setting of assessment where instrument was applied – flexibility	✔ Flexible	✔ Flexible	✔ Flexible	✔ Flexible	**X** Semi-flexible (only short-term detoxification units [20] and NR [49])	NA	NA	**X** Semi-flexible (only short-term detoxification units [20] and NR [49])
Costs and regulations for use – sustainability	✔ Copyrighted [144], free to download	✔ Licensed [145], free to download	✔ Free to download	✔ Free to download	✔ Licensed [137], free to download	✔ Free to use (for non-commercial research, registration and adherence to the PECUNIA methodology required)	✔ Free to download	NA
Generalisability and context								
Number of languages in which instruments are available – generalisability	✔ Available in >10 languages	**X** Available in <10 languages	**X** Available in <10 languages	**X** Available in <10 languages	**X** Available in <10 languages	**X** Available in <10 languages	**X** Available in <10 languages	NA
Type of MT intervention in which the instrument was applied – contextual fit	✔ Both	✔ Both	✔ Both	✔ Both	✔ Both	NA	NA	✔ Both

✔ … *yes.* **X** … *no.*

Abbreviations: BDI … Beck Depression Inventory. DRIE … Drinking-Related Internal-External Locus of Control Scale. HAQ(-II) … Revised Helping Alliance Questionnaire for therapist and client. IIP-SC … Inventory of Interpersonal Problems: Short Circumplex form. NA … not applicable. SCL-90-R … Symptom Checklist 90-R.

Summarising the suitability of measuring instruments for long-term use in individuals with SUD, we examined 14 tools across three key dimensions: 1) core measurement quality and clinical relevance, 2) practicality and burden of application, and 3) generalisability and context.

Bewertung der Langzeiteignung hinsichtlich der 3 Dimensionen

Core measurement quality and clinical relevance

Only the BDI has an established MCID (for depression only, not for SUD), limiting the interpretability of clinically meaningful change for the remaining instruments. For the PECUNIA-RUM, CSRI, and salivary Immunoglobulin A test, an MCID is not applicable.

klinische Interpretation nur für BDI (jedoch für Depression)

Outcome construct stability varied considerably. Instruments measuring stable, long-term constructs demonstrated high suitability, including DASES (self-efficacy in avoiding substance use), ASI (problem severity), SOC (sense of coherence), BDI (depression), SCL-90-R (psychiatric symptoms), DRIE (locus of control), and IIP-SC (interpersonal problems). The URICA, ICR, and CSRI demonstrated moderate-to-high appropriateness for measuring motivation and economic data. Conversely, the BSCS (craving), the salivary Immunoglobulin A test (immune function), and the PECUNIA-RUM (resource use) were deemed to have low-to-moderate or mixed appropriateness, respectively. The HAQ(-II) showed low appropriateness as it measures in-session process variables (therapist-client relationship, working alliance) rather than stable outcomes.

stabile (DASES, ASI, SOC, BDI, SCL-90-R, DRIE, IIP-SC) vs. akut schwankende oder zustandsabhängige (HAQ(-II)) Ergebniskonstrukte

Empirical evidence for long-term use was limited. Only five instruments (SOC, ICR, BDI, IIP-SC, and salivary Immunoglobulin A test) were used in treatments lasting more than six months. Four instruments (BSCS, URICA, ICR, BDI) demonstrated FU assessments of three months.

empirische Evidenz für Langzeitnutzung: begrenzte Datenlage

Regarding limitations, 11 instruments had fewer than four significant limitations, while the ASI, BDI, and salivary Immunoglobulin A test had four or more limitations.

mehrere Limitationen bei 3/14 Messinstrumenten

Practicality and Burden of Application

Ten instruments required ten minutes or less for administration, enhancing feasibility for repeated use. The ASI, SCL-90-R, PECUNIA-RUM, and CSRI exceeded this threshold. All instruments, except the ICR, contained more than fifteen items, providing adequate resolution and sensitivity, though four (BSCS, URICA, ICR, DRIE) had limited scoring ranges (≤50 points). The PECUNIA-RUM and CSRI are variable, where the relevant modules can be selected.

Praktikabilität für wiederholte Anwendung: 10/14 Messinstrumente ≤ 10 Minuten

Four instruments (ASI, ICR, salivary Immunoglobulin A test, CSRI) are employed by clinicians or interviewers, requiring trained personnel; 11 relied on self-report (ICR has a clinician-administered or self-completion format). Most instruments were questionnaire-based, except the ICR (visual analogue scale) and salivary Immunoglobulin A test (biological sample); the ASI and CSRI are interviews. Assessment flexibility varied: nine instruments offered flexible settings, while three (SOC, IIP-SC, salivary Immunoglobulin A test) showed semi-flexibility, limited to specific treatment contexts. For the PECUNIA-RUM and CSRI, this aspect is not applicable. All instruments were freely available/downloadable or free to use, supporting sustainability.

Durchführung und Flexibilität: 10/14 Instrumente durch Selbstauskunft; mehrheitlich Fragebögen;

Setting: 12/12 Messinstrumente flexibel

Generalisability and Context

Four instruments (ASI, ICR, SOC, BDI) were available in more than ten languages, supporting international application. The remaining ten had more limited linguistic availability. All instruments demonstrated contextual fit for both active and receptive MT interventions[9] (not applicable for the PECUNIA-RUM and CSRI).

>10 Sprachen (4/14 Instrumente); für aktive und rezeptive MT geeignet

3.3 Results of Part 3: Service user involvement – interview study

This chapter presents findings from nine interviews conducted in Austria, Norway, and Poland, supported by direct quotations from service users. The interviewees (4 female, 5 male) were aged 25-63. Six used polysubstances – primarily alcohol and cannabis, but also opioids – while three mainly consumed alcohol.

Stichprobe: n=9 (4 weiblich, 5 männlich, 25-63 Jahre)

Category development

To connect the interview study to the previous systematic literature review, the following categories were deductively developed based on Part 1:

deduktive Kategorienbildung basierend auf der systematischen Literaturübersicht

Substance use parameters

- Substance as an issue
- Thoughts about substances
- Reasons for substance use
- Long-term/Progression

Motivational parameters

- Reasons/Motivation for quitting/treatment
- Own efforts to quit

Recovery parameters

- Recovery
- Most important aspects to stay abstinent
- Relapse

Craving parameters

- Inability to cope
- Coping
- Trigger
- Sports

Quality of life parameters

- Quality of life
- (In)stability/Loss of control

[9] Contextual fit refers exclusively to documented empirical application in MT studies employing both active and receptive intervention types. It does not imply theoretical alignment between instrument constructs and specific MT methods, nor does it reflect demonstrated treatment effects.

- Daily life/Functioning
- Lifestyle changes/"New" life/Effects on life without substances/ sobriety

Social parameters

- Friends/Social environment/Relationships/Team/Network
- Family
- Parent-child-relationship
- Romantic relationship
- Animals
- Social exclusion/Isolation/Withdrawal
- Change in social life
- Change in position/Participation in society

Socioeconomic parameters

- Socioeconomic
- Job/Career
- Finances
- School/Education
- Criminalities
- Homelessness

Psychiatric/Psychological and mood-related parameters

- Aggression/Mental health/Psychosis/Paranoia/ Fear/Pessimism/Panic attack
- Suicidal thoughts/Will to live
- Depression/Hopelessness
- Anxiety

Physiological/Biomedical parameters

- Physiological/Physical health issues

Cognitive parameters

- Memory loss

Music-related parameters

- Music in life
- MT

Treatment-related parameters

- Hospital/Emergency visit/stay
- Treatment experiences
- Positive/Helpful aspects of treatments
- Negative/Challenging aspects of treatments
- Prejudices
- Supervision/Aftercare

manifeste Interviewanalyse

The following section presents findings using a manifest approach, providing an overview of interviewees' responses and identifying mentioned outcomes such as motivation, hospital stays, and depression.

3.3.1 Substance use parameters

Substance as an issue: Recognising the substance as the core problem

Substanz als identifiziertes Kernproblem

Service users came to recognise the substance itself as their central problem. Initial perceptions that it did not alter consciousness proved misleading. Substituting one substance for another was understood as merely a different form of changing one's state of consciousness.

Thoughts about substances: Persistent preoccupation and triggers for use

ständige Gedanken an Substanzgebrauch

externe Stressoren erhöhen Konsumrisiko

Service users described a constant preoccupation with substance use. Easy and cheap access reduced motivation to stop, and lack of supervision was associated with increased consumption. Stressful life events, such as grief or family crises, further intensified use. Even after extended periods of sobriety, thoughts about substances persisted, though not all service users relapsed during difficult times.

Reasons for substance use: Progressive disruption and loss of control

schleichende Lebensbeeinträchtigung

Lebensereignisse als Verstärker

Substances disrupted service users' lives gradually and progressively, affecting everyday emotions and leading to the breakdown of relationships. Drinking was often exacerbated by significant life changes, such as job loss or retirement. Over time, substances became increasingly central to daily life, with consumption progressing from occasional to continuous use, ultimately leading to a loss of control.

> "Alcoholism affected everything, my whole life, for a long period." (interview #9)

Long-term/Progression: From temporary relief to long-term recovery

Genesung als jahrelanger Neuanfang

Alcohol provided temporary relief, but problems persisted, and new ones emerged. Long-term use led to visible physical deterioration. Over time, service users recognised the severity of their situation, and that continued use would only worsen it. Recovery required a fundamental reinvention of their lives – a difficult process spanning years or decades, ultimately leading many to pursue inpatient treatment. Long-term, trusting contact with outreach team members was crucial in this process.

3.3.2 Motivational parameters

Reasons for quitting and motivation for treatment

Psychological and emotional factors

emotionale Erschöpfung

Selbsttäuschung belastend

Panikattacken als Wendepunkt

Interviewees experienced profound emotional exhaustion and mental devastation, describing themselves as "tired of themselves" and of living a lie. They expressed feelings of uselessness, a lack of achievements, and an absence of future prospects. The constant deception of relatives, particularly partners, created a significant psychological burden and self-disappointment. Many came to view their substance use as leading an "unworthy life" which – together with the recognition that acting was the only path forward – contributed to their motivation to change. For some, panic attacks served as a critical catalyst, marking a turning point that led to serious commitment to sobriety.

Physical health factors

körperliche Symptome als Ausstiegsantrieb

Angst vor Gesundheitskrisen, v. a. in Verantwortung für Kinder

Severe physical symptoms provided powerful motivation for cessation. Interviewees experienced serious health consequences, including cramps severe enough to require hospitalisation, which created acute fear for their lives. The deteriorating state of their health due to continued substance use became increasingly difficult to ignore. Beyond the immediate physical suffering, there was a particular fear of experiencing health crises in public settings, especially when responsible for caring for children, which added urgency to their decision to seek help.

Social and relational factors

Beziehungsbelastung als Veränderungsmotiv

Schuldgefühle und Elternschaft als Herausforderung

The impact on relationships emerged as a significant motivator for change. Interviewees felt diminished in interactions with non-substance users and experienced deep guilt about deceiving their loved ones. Parenthood posed a particular challenge, often leading to on-and-off substance use patterns. Strained family relationships, including loss of contact with children, highlighted both their vulnerability and the strain substance use placed on family bonds. The desire for a better life was a recurring motivator.

> "Having a kid has made me want to quit." (interview #6)

Practical and external factors

äußere Druckfaktoren wie z. B. Jobverlust, Angst vor rechtlichen Konsequenzen

External circumstances created concrete pressures that motivated cessation. Job loss was described as particularly catastrophic, especially when work had been central to the interviewees' identity. The unsustainable financial burden of substance use created a practical impossibility of continuing. Additionally, fear of legal consequences reinforced the decision to quit. These practical concerns, combined with internal motivations, create comprehensive pressure for change.

Previous attempts and escalation

gescheiterte Ausstiegsversuche

Rückfälle mit schweren Folgen

Previous attempts to quit – including compensating through overwork or earlier therapy – proved insufficient. Relapse often led to severe consequences such as loss of housing or relationships. Substance use was often rooted in earlier problems and gradually became central to the interviewees' lives, to the point where they felt unable to function without it. Some interviewees described a progression from fear of overdose to indifference towards death, ultimately reaching rock bottom.

> "So, either you just want to die, or quit." (interview #4)
>
> "I reached a point where I knew that if I didn't do something about myself, this addiction would stay with me for the rest of my life." (interview #1)
>
> "I was so far down that it couldn't get any worse for me." (interview #5)

The decision to seek treatment

Hilflosigkeit als Auslöser

The decision to seek treatment emerged from a convergence of internal realisations and external pressures. Interviewees reached a critical low point where they recognised their inability to stop independently and felt that their situation "could not get any worse." This recognition of hitting bottom, combined with a desire to "live a normal life" and achieve stability, motivated contact with rehabilitation services. However, internal motivation alone was often insufficient – external pressure from family members, outreach teams, or existing treatment contacts played a crucial role in pushing individuals toward voluntary treatment, particularly when severe illness made intervention urgent. For some, previous treatment experience, such as prior stays in therapeutic communities, made re-entry easier and more conceivable. Importantly, the desire for treatment was not simply about abstinence but about achieving stability while remaining active and engaged in life, as some interviewees found sobriety uncomfortable without the structure and support that treatment provided.

externer Druck (z. B. Familie) oft entscheidend

frühere Behandlungserfahrung erleichtert Wiedereinstieg

Ziel: Stabilität mit Lebensteilhabe, nicht nur Abstinenz

> "I couldn't take drugs anymore, and I couldn't live without drugs." (interview #2)
>
> "I simply gave up, because I didn't have any drugs, I didn't want to take drugs anymore, and I had to sort things out somehow, and I went to that centre." (interview #2)

Own efforts to quit: Treatment experience and sustaining abstinence

Therapie als Befreiung

Through treatment, interviewees achieved a sense of freedom from substance use. Entering therapy, particularly inpatient treatment, was a strong motivator for achieving abstinence. Professional guidance helped interviewees recognise the risks of isolation and boredom, leading them to proactively seek work or activities to prevent relapse. Institutional support and self-help groups were instrumental in sustaining the path to abstinence, alongside personal health concerns as a key motivator.

Isolation/Langeweile als Rückfallrisiko

institutionelle Hilfe stützt

Krisenerkenntnis als Veränderungsantrieb

The accumulation of failed attempts and worsening consequences ultimately led interviewees to a critical turning point: the recognition that substance use had become inseparable from their daily functioning and that independent cessation was no longer feasible. This realisation – often emerging from a prolonged sense of crisis rather than a single event – prompted them to actively seek treatment. For some, returning to familiar therapeutic settings, such as previously attended therapeutic communities, lowered the threshold for re-entering treatment. However, the transition to sobriety presented its own challenges: interviewees described feelings of discomfort, restlessness, and uncertainty in the absence of substances, making structured treatment environments an essential source of stability during this vulnerable phase. Family support frequently played a decisive role in facilitating admission and sustaining the commitment to treatment.

strukturierte Behandlung als Stabilität

3.3.3 Recovery parameters

Recovery and staying abstinent

Vertrauensaufbau: soziale Einbindung als Grundlage der Genesung

The process of recovery was linked to rebuilding trust and social connections, particularly within family dynamics. Recovery was often a repeated process, requiring multiple stays in rehabilitation clinics and psychiatric wards. The

sense of being able to talk freely to people who viewed substance abuse as an illness – rather than a moral failing – substantially encouraged interviewees to stay abstinent. The realisation that substance abuse is not bound to a stereotypical low socioeconomic population but is spread across all social classes encouraged a sense of belonging and reduced feelings of being an outlier.

Most important aspects to stay abstinent

Motivation durch wiedergewonnene Fähigkeiten und Beziehungen

Surrounding oneself with people who were also abstinent, as well as personnel who understood the complexity of substance abuse, helped maintain abstinence. The return of interest in long-lost hobbies and activities that had been abandoned during active substance abuse created a hopeful environment. Motivation was drawn from recognising regained abilities and interests, as well as from reconnecting with family and friends. The personal and relational advantages of staying abstinent were a significant factor in maintaining abstinence.

Relapse

Patterns and triggers

nichtlinearer Genesungsverlauf

Struktur und Monitoring verhindern vollständigen Rückfall

Recovery from substance use disorder follows a non-linear path characterised by periods of progress and relapse, demonstrating that sustained sobriety requires ongoing vigilance. Interviewees described varying lengths of abstinence – from approximately one year to nearly a decade – but vulnerability persisted regardless of duration. Work presented both opportunities and challenges; some interviewees experienced substance use episodes after encountering known individuals or substances in workplace settings, with use escalating quickly despite recent treatment and initial well-being. Structure and monitoring helped prevent complete relapse, catching problems before they spiralled entirely out of control.

> "Then I got a bit of a 'no, I don't want this', and then I haven't continued with it." (interview #5)

Emotional and relational factors

Rückfälle mit Beziehungsstress verknüpft

Schuldgefühle als emotionale Falle

Relapses were often intertwined with significant life stressors, particularly relationship difficulties and family dynamics. Relationships formed during vulnerable periods frequently dissolved following substance use resumption, with the cycle of connection, hope, relapse, and dissolution repeating itself. The impact extended beyond individuals to family members, with interviewees describing painful moments of honesty, such as explaining hospitalisations to their children. Guilt and the desire to "dull the pain" created an emotional trap, leading to repeated withdrawals and relapses.

> "Another withdrawal and another withdrawal. But I couldn't do it. And I felt very guilty. You reach for alcohol again to 'dull the pain'. It's a cycle, and it goes round and round like that." (interview #9)

Persistence and ongoing recovery

Erschöpfung durch zahlreiche Behandlungsversuche

Hilfe zum richtigen Zeitpunkt unterbricht Kreislauf

The exhaustion of repeated attempts was evident – some interviewees reported more than a dozen treatment episodes over many years, reflecting both the disorder's chronicity and the persistence required for recovery. Despite these challenges, sustained progress was achievable. Some interviewees maintained abstinence for several years despite occasional brief slips, demonstrating that recovery remains possible with continued effort and support. Even during therapy, individuals may still relapse but can stop immediately with continued support – evidence that help at the right moment can interrupt the cycle. These experiences reveal that recovery is not about perfection, but persistence – about the continued effort to break free from the cycle, one day at a time.

3.3.4 Craving parameters

Triggers and coping strategies

Substanzgebrauch als Bewältigungsmechanismus bei Stress/Konflikten

Alternativen (z. B. Sport, Natur) als Bewältigungsstrategien

Substance use served as a coping mechanism for agitated and uncomfortable situations, such as arguments or stress. The inability to cope with these emotional states led to continued substance use. Certain everyday activities could still trigger cravings long into recovery, highlighting the deeply ingrained nature of substance-related associations. Finding alternatives to alleviate uncomfortable emotional states was crucial. Interviewees reported using sports, nature, and, particularly, football to address cravings and cope with overstimulation.

3.3.5 Quality of life parameters

Quality of life: Instability and loss of control

Substanzgebrauch mit katastrophalen Folgen

Sozialprogramme hilfreich

Substance use significantly impacted interviewees' quality of life, as evidenced by housing loss and deteriorating health. The effects were described as disastrous and catastrophic, affecting every aspect of life. Extreme instability could also lead to psychotic states. Social programmes for housing and stable treatment helped improve these conditions. Despite the severity of negative consequences, they were rarely sufficient to trigger self-reflection or prompt action against the substance abuse.

Daily life and functioning

Funktionieren nur noch durch Substanz möglich

Interviewees tried to maintain daily life as long as possible, often reporting a false sense of control despite being deep within the substance abuse. Daily functioning was upheld by – and eventually only possible with – substance use. However, such coping strategies eventually proved insufficient, and what had once been an effort to maintain functioning became an inability to do so.

Life in sobriety: Effects on life without substances

Nüchternheit als Freiheit/Wiedergeburt

Being sober was described as a feeling of freedom and rebirth. This regained freedom was not only linked to no longer needing the substance but also to factors such as not having to worry about appearing intoxicated or being able to manage daily life independently.

"You pay more attention to everything and to yourself. You become more aware. It's great to feel free from alcohol. It really is a different life. Much freer." (interview #9)

berufliche/soziale Wiedereingliederung

Interviewees described profound and intense changes in their quality of life due to sobriety – rediscovering connections, sensations, and new life circumstances such as being able to hold a job or take on leadership positions. Job changes, reconnecting with family and friends, and regaining control over one's mind and actions gave them a new perspective on life. However, struggles with staying sober and adjusting to newly gained freedom were also crucial factors to consider.

aber auch Anpassungsschwierigkeiten

"Sober life was very different indeed. It takes a while to realise how much everything changes and how much alcohol influences which areas of life in terms of thoughts, feelings, and sensitivity. It's incredible. Your attitude really changes a lot. You become much freer and more self-confident. Yes, as the saying goes: a self-determined life." (interview #8)

Nachsorge, Kontinuität und Routine entscheidend

Aftercare, ongoing treatment, consistency, and routine were identified as crucial to maintaining sobriety. This concerned not only daily routines and new work structures but also therapy, social networks, and being able to spend time with oneself.

3.3.6 Social parameters

Social environment

Substanz als soziales Zugehörigkeitsmittel und Ursache für Isolation

At the beginning of substance use, interviewees described using substances as a way of feeling at ease and belonging within social groups. Over time, this changed drastically – what once provided the means to participate in social events became the reason; they could no longer maintain a social environment. Through intoxication, missed events, and increasing isolation, interviewees reported a steadily growing disconnection that was partly self-imposed and partly the result of being excluded by friends and family.

wachsende Entfremdung

fehlende soziale Unterstützung hält in Suchtdynamik

The struggle of removing oneself from substance-dominated social environments posed a significant difficulty during both active substance abuse and treatment phases. Some interviewees retreated to solitary substance use when wanting to distance themselves from their substance-praising environment, yet remained within substance use dynamics due to a lack of social support.

Family and partner

Familienbeziehungen fast durchgehend belastet

Experiences were almost exclusively marked by strained relationships with family members. Family dynamics were damaged through repeated failures to stay sober, leading to estrangement or complete loss of contact. Some interviewees reported attempts to mend relationships with varying degrees of success, though many relationships remained strained. In some cases, interviewees pushed family members away; in others, the burden of caring for them led to strained mental states among relatives, ultimately resulting in the loss of the relationship.

Angehörige psychisch mitbelastet

Romantic relationship

Partnerschaften häufig durch Substanzgebrauch zerbrochen

Long-term romantic relationships were frequently reported to have ended due to substance use. Some interviewees managed to rebuild these relationships, though for many the separation was final. Substance use within partnerships also created conflict – in some cases, a partner's substance use contributed to the interviewee's own consumption, while in others, the substance abuse led to the dissolution of the relationship entirely.

Parent-child relationship

Kinder als Abstinenzmotivation

Angst vor Suchtweitergabe

The impact on children was described as severe. Interviewees feared their children would view them negatively and reported their children's disappointment. Having children was a motivator for staying sober and continuing therapy, though this did not always suffice to uphold sobriety, leading to anger and disappointment among the children. The fear of their children also becoming substance users influenced interviewees to stop using and to be better role models.

Animals

Tiere als Stütze in Genesung, gibt Struktur/Sinn

Animals were given high value for their support throughout the recovery process. Animal care provided a daily structure and purpose, helping interviewees stay abstinent. However, strained family relationships sometimes led relatives to take away animals.

Isolation and withdrawal

Isolation als Schutz anderer und gleichzeitig urteilsfreier Konsum

Isolation was also used as a tool to avoid burdening others, though it simultaneously served as a means to continue using without judgment. Interviewees prioritised substance use over social engagement, with daily routines often revolving around acquiring substances.

> "I hurt the little social contacts that I had. Always with the excuse, "I don't want to put this on me and the others", in reality, I just wanted to go home and drink again." (interview #8)

Change in social life

Beziehungen wiedergewonnen

Austausch mit anderen Betroffenen hilfreich

Relationships were regained, and attitudes towards social life changed over time. Some interviewees initially found it helpful to be in contact with others in recovery and to talk about substance abuse openly, before gradually distancing themselves from substance-related topics and learning to accept their new sober identity. Others cut contact with all addiction-related acquaintances immediately to start a new life.

> "I noticed a change in the justification from a diplomatic excuse to an empowered justification. To stand with it and not fear it." (interview #7)

Change in participation and position in society

veränderte gesellschaftliche Teilhabe durch neue Arbeitsethik

Feelings of change regarding societal participation were often attributed to changes in work ethic and responsibilities. Interviewees reported a regained sense of authority and self-efficacy, with some achieving professional milestones they had previously thought impossible.

3.3.7 Socioeconomic parameters

Job and career

Substanzgebrauch gefährdet Arbeitsfähigkeit

nach Abstinenz Zeit für Wiedereinstieg nötig

Problematic substance use over time challenged interviewees' abilities to maintain their jobs and careers, making it difficult to uphold working relationships due to trust issues. Following periods of abstinence, time was needed to regain re-entrance and trust in work settings. Notably, several interviewees pursued self-employment after substance use ceased.

Financial aspects, school, and education

finanzielle Auswirkungen und Bildungserfahrungen …

Experiences with financial issues due to substance use varied. For some, substance use led to comprehensive financial challenges over time, while for others, it was not perceived as leading to financial difficulties. Welfare support was mentioned as helpful in avoiding debt.

> "Yes, I have built up a lot of debt during my [substance use] career. I've made a lot of loans, credit cards and all kinds of smart stuff. So, now I have at least 85,000 Euros in debt." (interview #5)

… unterschiedlich berichtet

Experiences with school and education also varied. For some, substance use led to non-completion of education, while others completed their formal education over time.

Criminality and homelessness

aus Kriminalität/Haft folgt gesellschaftliche Ausgrenzung

Many interviewees had extensive experiences with criminal activities over the years, leading to convictions and prison stays, which in turn led to further exclusion from society. Experiences of housing and homelessness varied – while some maintained a permanent and safe home, others' substance use had led to loss of housing and periods of living on the street.

3.3.8 Psychiatric, psychological, and mood-related parameters

Mental health challenges

schwere psychische Folgen

Isolation durch Verfolgungsängste

Verlust des Lebenswillens

Interviewees had comprehensive experiences with various mental health challenges due to substance use, particularly episodes of panic attacks, psychosis, and paranoia. Mental health was described as completely shattered. For several, fear of being followed and monitored led to social isolation and further deterioration of mental health. For some, losing contact with people in their close networks, combined with destructive substance use patterns, led to a loss of will to live. Substance use was experienced as seriously destructive, leading interviewees to recognise that they would deeply regret it if they did not finally stop.

"And then there's no one who can be around you because you're not on this planet, or you can't have normal conversations. You're just very paranoid. And then … then you cut all social contacts." (interview #5)

Depression, hopelessness, and anxiety

Abstinenz als Voraussetzung für psychische Gesundheit

Many interviewees experienced depression, existential hopelessness, and anxiety during the years of substance use. Following cessation, these mental health symptoms decreased for many over time. Substance use abstinence was experienced as a prerequisite for improvement of mental health.

3.3.9 Biomedical and physiological parameters

körperliche Abhängigkeit und Schmerzkreislauf

Interviewees reported somatic symptoms and health problems associated with substance use. In some cases, multiple hospitalisations due to substance-related illness led to job termination. Constant substance use led to physical dependence, resulting in chronic pain that called for further substance use – creating a vicious cycle. Long-term use also left visible marks on the body, deteriorating physical appearance.

"Even if you don't know I'm a drug addict, you can see that I'm a drug addict." (interview #2)

Arztbesuche aus Scham vermieden

Some interviewees reported that their physicians failed to link health problems to substance use, even when signs were evident. Others avoided medical appointments out of embarrassment regarding their substance abuse.

3.3.10 Cognitive parameters

z. B. eingeschränkte Reaktionsfähigkeit, Gedächtnislücken

Some interviewees reported cognitive effects. These included impaired reactivity and memory lapses when using substances, as well as repeated falls and memory loss attributed to substance use.

3.3.11 Music-related parameters

Music in the interviewees' lives

Musik als positive Ressource, gleichzeitig kann sie Verlangen triggern

Music had varied impacts on interviewees' lives. For many, music, rhythms, and lyrics evoked positive childhood memories and were used to foster states of energy, calm, concentration, and self-reflection. Simultaneously, several interviewees stated that they avoided specific musical genres to prevent being reminded of or triggering substance use cravings.

Music therapy

Interviewees had little or no direct experience with MT, as this was an inclusion criterion. For some, it was considered unnecessary for their recovery, while for others, it was seen as a potential support for recovery.

Musiktherapie: unterschiedliche Einschätzung ob relevant

> "When music gets inside me, that would have helped me a lot. I think it would have been good and that music therapy would have made a difference." (interview #9)

3.3.12 Treatment-related parameters

Emergency visits and hospital stays

For several interviewees, substance use and deterioration of mental health led to multiple emergency visits and hospital stays, including incidents of panic attacks, infections, cramps, and pancreatitis.

Krankenhausaufenthalte aufgrund z. B. Infektionen, Krämpfe

Treatment experiences

Many interviewees had various and repeated treatment experiences, including outpatient individual treatment, opioid substitution treatment, and group treatment. For some, treatment was effective for a period before slips and relapses occurred. For others, treatment led to self-reflection and provided structure.

vielfältige Behandlungserfahrungen wie z. B. ambulant, Gruppe

Positive and helpful aspects of treatment

Strong positive expressions were made concerning treatment experiences. These included self-help skills, a sense of community, a turning point, trauma resolution, strength-building, and life-saving effects.

z. B. Selbsthilfefähigkeit, Gemeinschaftsgefühl, Traumaverarbeitung

> "Hearing how others are coping or what they've been through really helps me. I can learn from their mistakes or draw strength from their successes." (interview #1)

Negative and challenging aspects of treatment and prejudices

Some interviewees found it hard to remain motivated to continue treatment due to long waiting lists. For several, it took time to establish trust and honesty with treatment facilitators, particularly concerning slips and relapses. Treatment was challenging as it required self-reflection on previous life choices and confrontation with personal emotions and past avoidance. The importance of identifying the right treatment to meet personal needs was emphasised. Initial prejudices about other patients or preconceived notions about people struggling with substance abuse were reported but quickly dissipated upon entering treatment.

z. B. lange Wartezeiten, Vertrauensaufbau, Rückfälle

richtige Behandlungsform wichtig

Supervision and aftercare

Many interviewees underlined the importance of long-term supervision and aftercare following treatment. These services included housing support, low-threshold contact, general everyday support, prevention of isolation, and well-timed, frequent contact with services.

Nachsorge wichtig für z. B. Alltagsbegleitung, Isolationsprävention

3.3.13 Quantitative results

Contingency analysis

Kontingenzanalyse: z. B. ,

Regarding the code-code relations, we determined that there are qualitative, content-related connections and, therefore, relations between codes. We found connections between 'family' and 'reason and motivation for quitting substance consumption and starting treatment'. 'Sports' were mentioned as one of the most essential aspects for 'staying abstinent'. We identified correlations between the 'changes in position and participation in society' over time and 'lifestyle changes in a new life without substance use' (sobriety) and 'job and career', respectively.

Familie ↔ Grund für Behandlungsbeginn

Sport ↔ Abstinenz-erhaltung

Frequency analysis

quantitative Häufigkeitsanalyse

Figure 3-1 provides a visual overview of the main categories derived from the qualitative analysis. The code cloud illustrates the relative prominence of each main category in the summarised material. Social parameters were by far the most frequently mentioned category, followed by socioeconomic and treatment-related parameters.

Figure 3-1: Code cloud of main categories derived from the qualitative interview analysis

häufig genannte Subkategorien:

Evaluating the data in more detail, Figure 3-2 displays the ten most frequently coded subcategories across the nine interviews. The length of each bar reflects the relative frequency of coding. The subcategories shown are drawn from various main categories and represent the themes that were most prominently captured in the summarised interview material. The three most frequently coded subcategories were 'family', 'friends/social environment/relationships/team/network', and 'reasons/motivation for quitting/treatment', suggesting that family and social relationships, as well as motivational aspects for entering treatment, were particularly prominent themes across the interviews.

Familie, soziale Beziehungen, Motivation für Behandlung

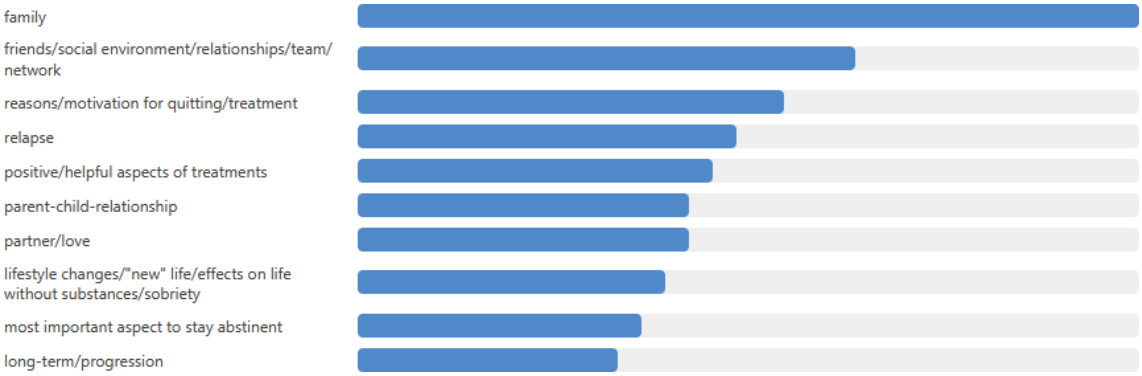


Figure 3-2: Top 10 subcategories by coding frequency

Given that social parameters constituted the most frequently coded main category, this domain was subsequently subjected to a more in-depth latent content analysis.

latente Inhaltsanalyse der sozialen Parameter

3.3.14 "The Social" as the relevant parameter in substance use recovery: in-depth latent content analysis

How the social was constituted as a theme

Service users never thematised "the social" as an abstract category. It tended to surface more often in its negation than in its affirmation, such as "broken relationships", "lost trust", and "enforced isolation". The social functioned as an absent presence and could be seen as a condition becoming perceptible at the point of collapse. The language of social destruction tended to be more vivid and emotional, whereas the language of social reconstruction was more compressed and less affective. Latently, this asymmetry suggested that even where social recovery had factually occurred, service users had not yet made it fully "their own". They could describe in detail what they lost, but struggled to articulate what they have gained, as though the new social life had not yet been sufficiently inhabited to generate its own language.

Soziales häufiger als Verlust denn als Gewinn thematisiert

sprachliche Asymmetrie deutet auf unvollständige Verinnerlichung sozialer Erholung

Equally revealing was what is systematically absent. All interviewees spoke about family and partnerships; almost none addressed neighbourhood, community belonging or civic participation. This convergence revealed a privatised narrative template in which "the social" was reduced to the intimate sphere and the labour market, while meso-level social structures remained invisible.

auffällige Abwesenheit: Nachbarschaft, Gemeinschaft

Metaphorical architecture

Four dominant metaphorical clusters organised social experience latently. *Verticality* ("sinking", "falling", "getting up off one's knees") mapped social position onto a moral axis, constructing exclusion as personal descent rather than structural condition. *Mechanisation* ("cycles", "conveyor belts", "Groundhog Day") framed social life during substance abuse as non-eventful repetition. *Being removed from one's world* ("being lifted out like a crane", "placed somewhere white, clean, neutral") revealed a longing to be placed in a space entirely emptied of social demands – as though the social world must first be reset before it could be rebuilt. *Fight* ("fighting", "surrendering", "pulling emer-

Ausgrenzung als persönlicher Abstieg; Sucht als ereignislose Wiederholung; Wunsch nach sozial entleertem Raum als Voraussetzung für Neuanfang; …

gency brakes") constructed the social world as combative, with the "enemy" remaining ambiguously located between the substance, the self, and the environment.

… soziale Welt als Schlachtfeld

Linguistic form

Three linguistic registers alternated throughout the data: clinical/therapeutic language (from treatment settings), raw colloquial language (moments of unmediated affect), and moralising language (internalised societal judgments). It could be suggested that service users drew on different discursive resources to make sense of their social experience.

unterschiedliche narrative Ressourcen zur Verarbeitung sozialer Erfahrung

Latent linkages and novel findings

Service users connected themes that were conceptually distinct but became intertwined through the way they were narrated. These implicit linkages, never explicitly marked by the speakers themselves, revealed underlying structures that shaped how social reality was experienced and organised in recovery. *Children* functioned as internalised surveillance mechanisms, not only as motivational objects but also as representatives of normative society whose gaze disciplined parental behaviour. The fear that children "discovering" or "judging" positions could be seen as a moral panopticon rather than a subject of care. *Animals* served as transitional social objects, providing relational presence without relational risk, bridging the gap between social isolation and the capacity for human connection. *Financial participation* functioned as a social admission ticket: paying taxes and earning money were narrated not as economic necessities but as prerequisites for moral citizenship, in which belonging was conditional upon productivity.

implizite Verknüpfungen:

Kinder als moralisches Panoptikum, Tiere als relationale Übergangsobjekte, finanzielle Teilhabe als Voraussetzung moralischer Zugehörigkeit

The analysis identified a tension within *therapeutic sociality*. The treatment group was one of the few communities where service users felt genuinely understood and accepted, yet it was defined by the very condition from which they sought to recover. The transition out of treatment, therefore, carried a latent risk of losing a significant source of social belonging, precisely at the moment when "ordinary" social life still felt demanding and unfamiliar. Furthermore, *social life after recovery* was experienced ambivalently. As desired yet threatening, requiring energy that substances previously might have provided. The finding that social contact felt overwhelming without substances challenged the idea that recovery follows a straightforward path.

therapeutische Gemeinschaft als wichtige Zugehörigkeitsquelle

Behandlungsende birgt Verlustrisiko

Positive social outcomes were expressed in two distinct forms: through negation ("I don't want that much more anyway", "no longer tolerates") and through affirmation ("much freer", "more self-confident", "everything has changed"). Finally, institutional and structural barriers were addressed frequently. Service users described exclusion from employment due to criminal records and substance use history, loss of jobs through mandatory drug testing, lack of affordable individual therapy and long waiting times. While narrated as personal experiences, these stories recurred across service users, revealing how deeply individual recovery was entangled with systemic conditions that enable or obstruct social reintegration.

Verzicht vs. Freiheit, Selbstvertrauen

strukturelle Barrieren z. B. Drogentests, Wartezeiten

soziale Reintegration systemabhängig

4 Discussion

Summary of the report

MT zur langfristigen Behandlung von SUD: Potenzial und Forschungsbedarf

SUD carries a significant global burden, leading to substantial disability and often co-occurring with other health issues, crime, and poverty [2]. Music engages the brain's reward system through mechanisms similar to those of addictive substances [23-25] and is intrinsically rewarding [27], providing a neurobiological rationale for its use in clinical contexts. MT harnesses these properties within a structured therapeutic framework, and various forms of MT have demonstrated benefits for people with SUD [11]. However, high relapse rates within the first year post-treatment and evidence that longer treatment durations are associated with better long-term outcomes [46, 47] highlight the need to further investigate MT's impact on long-term treatment retention and its sustained effects on substance use behaviours.

7 systematische Übersichtsarbeiten eingeschlossen

14 Messinstrumente im Detail analysiert

This systematic review identified clinical and socioeconomic outcome parameters and measuring instruments for MT in individuals with SUD. Seven reviews were analysed, identifying more than 60 outcomes, categorised into 11 parameters. Tools assessing (serious) adverse events were not reported. Accordingly, 14 frequently used, validated measuring instruments were selected for detailed analysis to assess their general and application characteristics, appropriateness, feasibility, user relevance, and suitability for long-term use.

10 Instrumente <10 Minuten, überwiegend Selbstberichte, limitierte Langzeitevidenz

Ten instruments required less than or equal to ten minutes to administer; 11 were self-reporting tools; and four were available in more than ten languages. Limitations identified across instruments included self-report bias, challenges with cultural adaptation, and training requirements. Regarding long-term suitability, seven instruments measuring stable constructs such as depression, problem severity, and self-efficacy were highly appropriate. However, empirical evidence for long-term use was limited, with the longest documented treatment duration of two years and the most recent FU assessments at three months.

9 Interviews: soziale Parameter zentral

soziale Bindung und sinnvolle Aktivitäten für Abstinenz entscheidend

Nine semi-structured interviews with service users from Austria, Norway, and Poland revealed that social parameters were most frequently mentioned, followed by socioeconomic and treatment-related parameters. The three most frequently coded subcategories were 'family', 'social environment', and 'reasons for quitting and motivation for treatment'. The interview data demonstrated that substance use progressively disrupted multiple life domains, recovery required sustained support and repeated treatment attempts, and social connection and meaningful activities were critical for maintaining abstinence.

Discussion of the report

nur ASI und SURE stimmen mit COS-Leitlinien überein

The question arises whether the outcomes identified in our systematic literature search are consistent with the existing COS guidelines (see 1.4.1). Surprisingly, among the guidelines described (i.e., [22, 41, 43]), only the European Addiction Severity Index (EuropASI) was also used in the studies meeting our inclusion criteria. The Substance Use Recovery Evaluator (SURE) also appeared in our manual search.

standardisierte Ergebnismaße sind unerlässlich, um Gesundheitsergebnisse zu vergleichen, werden jedoch in der klinischen Praxis selten angewendet

Standardised outcome measures are essential for comparing healthcare outcomes across institutions and patient groups worldwide. Yet their adoption in clinical practice remains uncommon [146], consistent with our findings. This limited uptake may stem from challenges such as difficulties achieving expert consensus, trade-offs in selecting measurement tools, high implementation costs, and duplication of efforts in developing similar standard sets. Greater transparency in the development process could help reduce this redundancy and promote broader adoption [146].

daher ist es wichtig, die COS-Leitlinien zu verbreiten, um die Berichterstattung zu standardisieren und die methodische Stringenz zu verbessern

Furthermore, researchers may not be fully aware of the existence or specific recommendations of relevant COS guidelines, and implementing COS-recommended outcomes may present practical challenges in certain research settings (e.g., resource constraints, specific study populations). The limited use of COS-recommended outcomes hinders effective comparison and synthesis of findings across studies. Therefore, it is crucial to disseminate and promote COS guidelines within the research community to benefit from standardised outcome reporting, improved methodological rigour, and enhanced comparability.

Einbindung von Service Users in SUD-Forschung sichern

Securing service users' engagement in longitudinal SUD recovery research is essential, as high attrition threatens study quality [147]. Longitudinal studies, especially those involving hard-to-reach groups like individuals with SUD, often face the challenge of participant dropout over time. In addition, there is a particular lack of recovery research concerning individuals who have achieved stable abstinence and functional recovery from SUD [4, 148].

Herausforderungen bei der Implementierung von patient:innenberichteten Ergebnisparametern wie z. B. Behandlungsabbruch

Implementing patient-reported outcome measures in SUD treatment faces significant challenges, including patient discontinuation and increased workload for both staff and patients. Researchers in this field often seem reluctant to rely on self-reported data, primarily due to concerns about reliability given the social stigma surrounding substance use and potential negative repercussions of disclosure. However, a substantial body of research demonstrates high concordance between self-reported drug use and biological measures, suggesting that these concerns may be overrated [149].

Kriterien der Langzeitmessung wie z. B. Belastung der Patient:innen und Verlust bei Nachuntersuchungen

Criteria for long-term assessment instruments include burden on patients (time, cognitive load, accessibility, retention) and on researchers and clinicians (resources, workflow integration, data collection, training). Attrition and loss to FU should be mitigated through regular communication, incentives, and flexible data collection. Researchers must also account for intervening events, such as comorbidities or life changes, that may confound results and ensure that measures are sensitive enough to capture both improvements and deteriorations in health status over time.

3 Faktoren welche Service Users an Studienteilnahme binden

To understand what sustains engagement in SUD recovery studies, three key themes were identified: adaptability in response to slips or relapses, the development of strong and flexible relationships with participants, and an emphasis on long-term perspectives on both study participation and treatment – all of which can significantly improve retention [4].

bei Langzeitstudien: Beziehungen und flexible Teilnahme sichern

Tracking-Prozesse und proaktive Strategien integrieren

For long-term FU studies on diverse recovery journeys, strong research relationships and clear agreements allowing continued participation during periods of drug use were considered highly relevant. Incorporating these factors into tracking procedures could strengthen longitudinal research. Similarly, clinical services could improve interventions by recognising long-term recovery as a nonlinear process and adopting proactive strategies to support individuals throughout their journey [4, 150].

Service Users in Genesungskonzept einbeziehen

The concept of recovery continues to spark debate in drug policy and treatment service delivery. To make recovery more meaningful, it is essential to listen to the perspectives of people who use substances and to critically examine the foundational principles and real-world outcomes of current treatment policies and practices [151].

Verlangen nach Substanzen in 5/7 Studien gemessen

Lebensqualität wie z. B. Schlafqualität wurde weniger oft gemessen

Our results show that craving parameters appeared in all but two [49, 84] of the included reviews. However, this contrasts with findings from [148], which reported that craving was present in only 31.1% of short-term and 0% of long-term studies. QoL was similarly underrepresented – present in 7.1% of short-term and 0% of long-term studies in [147] – consistent with our results: few QoL parameters, such as sleep quality, meaningfulness of life, or well-being, were measured in only three studies. This is surprising given the importance of QoL for mental health. Measures of role functioning – defined as participation in taxed employment, school participation, and disability days – and social functioning should be added to reflect an adequate conceptualisation of recovery [148].

Lebensqualität ist multidimensional

kontinuierliche Unterstützung essentiell für nachhaltigen Erfolg

Our qualitative interviews revealed that service users conceptualise QoL across multiple dimensions. Substance abuse profoundly affected life stability and daily functioning, including housing, health, and sense of control. Recovery was described not merely as abstinence but as "freedom and rebirth", encompassing psychological (self-awareness, autonomy), social (relationships), and occupational (employment) dimensions. Service users emphasised that consistency, routine, and ongoing support structures were crucial for maintaining sobriety and for improving QoL, reinforcing recovery as an evolving rather than static process.

Diskrepanz zwischen standardisierten Instrumenten und erlebter sozialer Genesung

soziale Parameter könnten besonders relevant sein

künftige Forschung sollte z. B. soziale Funktionen und subjektive Autonomie erfassen und Instrumente mit Erfahrungen abstimmen

The latent content analysis reveals a fundamental mismatch between standardised outcome instruments and service users' experience of social recovery. Social parameters were by far the most frequently mentioned aspects, cautiously suggesting particular user-relevance for people with SUD. Yet available tools, such as the ASI, IIP-SC, and MSPSS, capture social functioning through static indicators such as relationship problem areas and perceived support. The findings show social reality in recovery to be structured by ambivalence and paradox, e.g., social contact as simultaneously desired and threatening and therapeutic sociality where leaving treatment means losing community. Meso-level dimensions – community belonging, civic participation – remained absent from narratives yet may be critical to sustained reintegration. Service users vividly articulated social destruction but struggled to articulate social gains. While self-reported data show high concordance with biological measures, as mentioned above, the situation differs for social recovery outcomes, where self-report instruments may systematically underrepresent progress: rating "social functioning" on a Likert scale risks capturing normative compliance rather than lived reintegration.

nichtlinearer Genesungsverlauf, wird jedoch von Instrumenten vorausgesetzt

psychometrische Evaluation für Langzeiteinsatz nötig

For long-term outcome assessment, the dominance of metaphors of descent, repetition, and combat indicates that recovery is not experienced linearly, yet most instruments presuppose a linear trajectory. Feasibility and user-relevance as evaluation criteria must extend beyond ease of administration to whether an instrument's conceptual structure aligns with experiential reality. Whether existing instruments are suitable for long-term use requires a dedicated psychometric evaluation. Future research should incorporate measures sensitive to non-linear social rebuilding, such as relational quality beyond treatment, to better align outcomes with the priorities of those in recovery.

MT als Langzeitintervention: umfassende Auswirkungen von SUD erfordern oft jahrelange Unterstützung über mehrere Lebensbereiche

Our interviews revealed that problematic substance use has long-term, comprehensive impacts requiring extended recovery trajectories, underscoring the importance of MT as a continuous rather than time-limited intervention. Notably, when discussing treatment experiences, interviewees primarily focused on broader life changes – such as employment, social relationships, and psychological well-being – rather than specific clinical symptoms. This aligns with our measured outcomes, demonstrating that MT's impact extends beyond symptomatology to encompass existential and life-domain improvements.

qualitative Methoden erfassen tiefere Dimensionen jenseits quantitativer Messung

The multidimensional nature of MT's impact in SUD treatment necessitates complementary qualitative approaches. While standardised instruments quantify symptom changes and functional improvements, they cannot fully capture how and why MT facilitates recovery. Qualitative methods, particularly in-depth interviews, illuminate subjective mechanisms, meaning-making, and contextual factors underlying therapeutic change. Importantly, qualitative approaches can also yield meaningful insights into outcomes – for instance, by identifying dimensions of change that matter to service users but are not reflected in standardised measures – even though they do not allow for conclusions about efficacy in the strict, controlled sense. Our service user interviews revealed recovery priorities – such as autonomy, social dignity, and meaningful routines – that extend beyond traditional clinical endpoints. Future MT research should therefore integrate mixed-methods designs to explore lived experiences that quantitative measures alone cannot represent.

soziale und sozioökonomische Messinstrumente einbeziehen

While the current evidence employs many validated and reliable tools, few studies have assessed socioeconomic parameters, even though interviewees frequently mentioned them. Additional socioeconomic and social instruments should be incorporated to expand the toolkit in MT and SUD research.

soziale Dimension bei SUD: breiteres Verständnis notwendig

As noted above, existing research tools do not fully capture the social dimensions that matter to people with SUD, as most instruments focus primarily on interpersonal relationships with family, partners, and friends. However, service users highlighted that the social domain extends well beyond these ties, identifying factors such as companion animals and group activities, such as sports, as meaningful contributors to recovery. A broader conceptualisation of the social dimension is therefore needed – one that encompasses a wider range of sources of connection and belonging.

MT: Musizieren in der Gruppe fördert soziale Kohäsion und Recovery

In MT practice, group-based musical activities are particularly well-positioned to address these broader social needs, given the unique capacity of music to simultaneously engage multiple social functions. Group music-making has been shown to foster contact, social cognition, copathy, communication, coordination, and cooperation – all converging toward increased social cohesion [152]. These social functions of music are not merely incidental to the therapeutic process but constitute core mechanisms through which MT may support recovery. This is further supported by evidence that community-based arts participation, including music, can build social cohesion and, in turn, enhance well-being in communities [153]. While other group activities such as sports or crafts can also promote social connection, music holds a distinctive quality in that it can engage all of these social functions effortlessly and simultaneously [152], making it a particularly powerful vehicle for the kind of relational rebuilding that service users described as central to their recovery.

Implikationen für Messinstrumente und RCT-Design

Given that recovery is typically an extended process, long-term measuring instruments should account for these dimensions, as the gradual development and maintenance of such connections may play a crucial role in sustaining progress over time. Accordingly, when designing RCTs, careful attention should be given to ensuring that the therapeutic goals of the MT intervention align with the outcome measures selected – in this case, that interventions explicitly targeting the social functions of music are evaluated using instruments sensitive to social parameters.

soziale Gemeinschaften als Schlüssel zur Genesung:

soziale Reintegration, gesellschaftliche Teilhabe und Arbeit

The interview data validate the theoretical understanding [154] that social recovery involves both separation from harmful networks and active reconnection with supportive communities. Service users' accounts reflect the trajectory described in the literature: from substance use as social coping, through isolation and relationship deterioration, to rebuilding connections through intentional community engagement. Their experiences of distancing from addiction-related networks, regaining self-efficacy through work and participation, and recovering social dignity – particularly through voluntary work – exemplify how safe, non-judgmental spaces facilitate positive change and restore responsibility and trust.

4-Stufen-Modell, welches Forschung hinsichtlich sozialer Genesung leiten soll

Understanding how individuals experience social recovery firsthand and how communities contribute to these processes is therefore essential. Given the key role of social communities in recovery, the authors of a meta-synthesis of qualitative studies propose a four-stage model (see Figure 4-1 [154]) to guide research on social recovery from first-person perspectives [154].

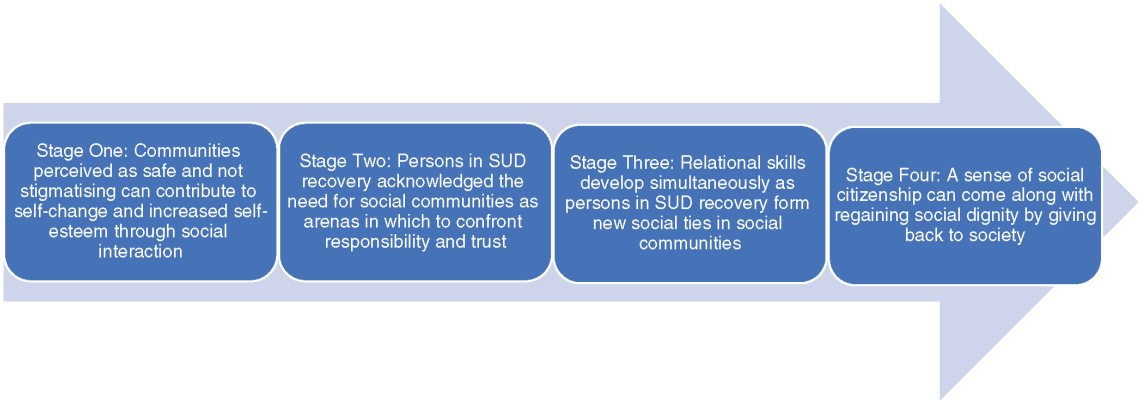


Figure 4-1: A model to guide research into social recovery from a first-person perspective and how social communities support these processes (source: [154])

Freundschaft als zentrales Genesungsergebnis in der Langzeitforschung

Assoziation mit beruflicher und psychosozialer Reintegration

In the interviews, friendship emerged as a significant recovery outcome, capturing the relational reconstruction process. Service users described how substance use initially served as a basis for social bonding but paradoxically accelerated social fragmentation as dependency deepened. Recovery involved deliberate friendship formation outside substance-related contexts, representing a concrete behavioural shift beyond abstinence. Those who cultivated stable friendships showed concurrent improvements in occupational functioning and self-perceived authority, suggesting friendship as a meaningful indicator of broader psychosocial reintegration. Friendship metrics, therefore, warrant incorporation into longitudinal outcome assessments for SUD interventions. While current research offers insights into how individuals in long-term recovery find meaning in forming friendships, further exploration is needed into how people with a history of substance use build these relationships from their own perspectives [155].

erhebliche Lücke von Instrumenten für soziale Genesungsparameter

However, the evidence base reviewed in Part 1 reveals a notable gap, as only a small number of instruments were identified that capture social dimensions relevant to recovery, such as the ASI. No instrument was designed to measure friendship formation, community reintegration, or the progressive rebuilding of social networks as described by the interviewees. Given the centrality of social recovery processes, as highlighted in both the literature and service users' accounts, this represents a significant gap in the current measurement landscape.

sinnvolle Aktivitäten spielen zentrale Rolle z. B. Arbeit, Tagesstruktur, Sport

Service users highlighted the role of meaningful activities in recovery, describing daily routines, (volunteer) work, sports, and relationship-building as crucial. These findings align with a qualitative study [40] that identified three key themes in recovery: the central role of work, mastery and commitment (e.g., maintaining daily structure and functioning), and repairing the bridge to community life through employment and physical activity [40].

Bedeutung und Empfehlung von MT bei komorbiden psychischen Störungen

validierte Messinstrumente vorhanden

Many interviewees disclosed mental health issues, underscoring the relevance of MT in SUD treatment given the high rate of comorbid mental disorders. Although specific evidence for MT in this context remains limited, MT is strongly recommended in national guidelines for the treatment of psychosis [22], and evidence from MT studies more broadly suggests benefits for other common comorbidities as well. For instance, improvisational MT has been associated with reduced clinically assessed depressive symptoms [12]. Based on these established effects, MT is considered likely beneficial for patients with harmful substance use and drug dependence [22]. Notably, the existing evidence base identified in Part 1 of this report already includes validated instruments that capture key dimensions of comorbid mental health – such as depression (BDI, HRSD, SDS), anxiety (SAS, STAI), and psychiatric symptom burden (SCL-90-R, BSI). However, the suitability of these instruments for long-term measurement in SUD populations requires further consideration, particularly given the chronic and relapsing nature of both SUD and comorbid psychiatric conditions as described by the interviewees.

Langzeitabstinenz: Verbesserungen in verschiedenen Lebensbereichen

MT-Evaluation muss langfristige, multidimensionale Ergebnisse erfassen

Given that individuals with long-term substance abstinence (>1 year) show the most significant improvements across life satisfaction, executive functions, and psychological distress [156], SUD recovery must be conceptualised as a gradual, long-term process encompassing multiple life domains beyond abstinence [157]. This underscores the need for MT interventions and their evaluation to adopt a comprehensive, longitudinal approach – extending beyond immediate behavioural changes to capture sustained improvements in areas such as mood regulation and QoL. Evidence supports this perspective: longer treatment durations are associated with lower substance use at long-term follow-ups and better aftercare engagement [46, 47], and qualitative research emphasises that meaningful, structured activities such as music engagement are central to sustained recovery across multiple life domains [40]. In the context of MT specifically, investigating its impact on treatment retention and long-term effects on people with SUD has been identified as crucial [11]. Long-term instruments that assess MT's impact across these interconnected domains are therefore essential for understanding its true clinical and socioeconomic value in SUD treatment.

Messinstrumente zu z. B. Inanspruchnahme von Leistungen, Produktivität und Beschäftigung …

While cost-effectiveness analysis was not the focus of this study, several instruments captured parameters relevant to future economic evaluations. Service use data from the CSRI and CSSRI provide information on health and social care utilisation for calculating direct costs, while the PECUNIA-RUM and WHO-HPQ enable estimation of indirect costs related to productivity loss

… ermöglichen Kosteneffektivitätsanalyse

and occupational health. Employment and income data from the ASI, CSRI, CSSRI, and SES questionnaire, along with occupational status measures (Duncan SEI, Nam-Powers OSS), provide further insights into economic productivity for cost-benefit analyses. Some of these tools, such as the ASI and SURE, are directly applicable to SUD research.

Bewertung von Human- und Sozialkapital durch CAPSES und SURE

The assessment of human and social capital through CAPSES and SURE indirectly reflects economic resources and labour market positioning. Future research should leverage these data on service utilisation, productivity, employment, and human and social capital to conduct comprehensive cost-effectiveness analyses, thereby extending the current findings to include economic perspectives on intervention outcomes and resource allocation.

Bedeutung der Gesundheitsökonomie für die Entscheidungsfindung über die Finanzierung von Gesundheitsleistungen

Cost-effectiveness analysis is central to determining which healthcare services receive public approval or insurance funding. Health economics plays an increasingly important role in shaping these decisions, offering crucial insights for resource allocation. However, funding decisions are inherently complex, embedded in societal values and political contexts. Within an evidence-based framework, they require a thorough evaluation of both efficacy (controlled trial outcomes) and effectiveness (real-world outcomes) [158].

rechtliche und finanzielle Anerkennung von MT weltweit uneinheitlich, was zu unterschiedlichen Finanzierungsmodellen führt

MT is a global healthcare profession that often lacks formal governmental recognition, leading to highly varied funding models – ranging from government-funded healthcare integration to grants, sponsorships, or out-of-pocket payments. The existing literature reflects a diverse landscape regarding health economic methodologies and cost components. Global MT organisations face significant obstacles in achieving better legal standing and financial support, largely due to insufficient and disorganised evidence on MT's cost-effectiveness [158].

selbstberichtete Daten zur Gesundheitsdienstnutzung

Socioeconomic aspects can also be measured using administrative data and patient surveys that assess service utilisation. While self-reported data on healthcare use are subject to recall bias [159], this approach remains widely used as it provides direct insight into patients' perspectives and healthcare-seeking behaviour. Understanding how individuals access services provides valuable insights into the dynamics of universal healthcare systems that aim for equitable access. Importantly, individual health literacy determines how effectively people can utilise these services to address their health needs [160].

keine Instrumente die unerwünschte Ereignisse messen

None of the included studies reported tools for assessing adverse events, representing a critical methodological limitation. Standardised adverse event monitoring is fundamental to establishing any intervention's safety profile, and without systematic assessment, understanding of MT's safety in SUD treatment remains incomplete. Future research must employ validated adverse event tools to evaluate both therapeutic benefits and potential risks, enabling evidence-informed clinical decisions and policy recommendations. For instance, the Negative Effects Questionnaire (NEQ) [161], a validated instrument in the mental health area that assesses the negative effects of psychological treatments across domains such as symptom worsening, hopelessness, and stigma, could serve as a feasible starting point.

Limitations

Evidenz und Interviews: nominale Übereinstimmung (gleiche Themen), aber in der Essenz unterschiedliche Konstrukte

Although both the evidence (Parts 1+2) and service user interviews (Part 3) address clinical and socioeconomic parameters, they do so through fundamentally different lenses. The correspondence is nominal – same topic names – rather than essential, as underlying constructs, mechanisms, and priorities differ. Future research should either conduct genuinely inductive analysis of service user data to identify emergent outcome dimensions or explicitly examine whether evidence-based parameters adequately capture service user-defined recovery outcomes.

Zielgruppe: nur Erwachsene eingeschlossen, obwohl Substanzkonsum bei z. B. Jugendlichen relevant

This report focuses on adults, though substance use is also relevant in other life stages such as adolescence. A life-course perspective could be valuable for preventing and controlling substance abuse from the earliest stages, accounting for the needs of every age group. Crucially, investing in early health promotion and prevention – particularly for mental health – offers a highly promising return on investment and should be a future priority [162].

Fokus auf systematische Übersichtsarbeiten stärkt Zuverlässigkeit unserer Ergebnisse

Furthermore, only systematic reviews were included in our search; primary studies were excluded. This strengthens the reliability of our findings, as systematic reviews represent evidence that has already undergone rigorous quality appraisal, synthesis, and interpretation, reducing the risk of bias associated with individual studies.

MT-Interventionskategorisierung in Primärliteratur unzureichend differenziert

The categorisation of MT interventions was limited by the reporting practices of the included reviews, which predominantly distinguished only between active and receptive delivery rather than employing the more granular typology common in the MT literature – improvisational, compositional, recreative, and receptive methods [241]. This limits the specificity with which findings can be attributed to particular MT approaches and should be addressed in future primary research through more consistent intervention reporting.

Überschneidung der Datengrundlage irrelevant

Regarding data extraction, some primary studies (e.g., [142]) appeared across multiple included systematic reviews (e.g., [20, 49]). However, as our focus was on measuring outcomes and instruments, this overlap does not limit the scope of the report.

unvollständige Darstellung psychometrischer Langzeiteigenschaften

RQ2 aimed to identify the characteristics of the most frequently used, validated, and reliable instruments and to assess their appropriateness, feasibility, and user relevance for long-term measurement in patients with SUD. However, we were unable to address every aspect in detail. Essential psychometric properties particularly relevant for long-term measurement could not be included, such as construct validity, criterion validity, responsiveness to change, test-retest reliability, and inter-rater reliability – all of which are vital for evaluating whether an instrument can meaningfully capture change over time from a patient's perspective.

Interviewstudie: Limitationen der qualitativen Inhaltsanalyse

The qualitative content analysis used for the interviews cannot draw causal conclusions and presents methodological challenges. However, it is a valuable and unobtrusive method for systematically describing communicative messages, as it is considered close to interviewees because it directly examines the content they produce [163].

Vor- und Nachteile des deduktiv-induktiven Ansatzes …

While the combined deductive-inductive approach allowed for both theory-driven and data-driven insights, several limitations should be noted. The initial deductive framework, though grounded in the systematic review, may have introduced assumptions constraining the emergence of novel phenomena. The inductive derivation of subcategories within deductive categories may have created tension between exploring emergent themes and maintaining con-

sistency with the pre-established framework. Additionally, the analysis primarily focused on manifest rather than latent content, prioritising explicit statements over deeper interpretation of underlying meanings. However, we conducted latent analysis of the most frequently mentioned outcome – social parameters – given their prominence in the interviews. A purely inductive approach would have risked losing the empirical foundation from the first phase, leading to less structured interpretations without the grounding necessary for this clinical context. Overall, our methodological choice prioritised descriptive fidelity – capturing what interviewees actually said rather than inferring meaning – to ensure findings remained grounded in participants' voices and transparent for integration with quantitative data. Furthermore, it should be acknowledged that parts of the qualitative content analysis – for instance, the inductive derivation of subcategories, which involves greater interpretative judgment – were not conducted with two independent coders. This represents a limitation regarding inter-coder reliability and should be considered when interpreting the findings.

… und der manifesten vs. latenten Inhaltsanalyse

Nevertheless, despite these limitations, the results provide a valid impression of clinical and socioeconomic outcome parameters and long-term measuring instruments for individuals with SUD.

5 Conclusion

Lücken in Langzeitmessung und Sicherheitsmonitoring identifiziert

This systematic review reveals identified gaps in the assessment of long-term clinical and socioeconomic outcomes for MT in SUD treatment. Although validated measurement instruments are available, their empirical verification for extended timeframes remains insufficient. Notably, MCIDs have been established for only a single instrument, without specific validation in SUD populations. A particularly concerning methodological deficit is the complete absence of standardised protocols for monitoring adverse events, which prevents a comprehensive safety evaluation of MT interventions in this context.

soziale und berufliche Teilhabe werden unterrepräsentiert

Qualitative data from service user interviews indicate that conventional measurement approaches systematically underrepresent crucial recovery dimensions, particularly social integration, engagement in purposeful activities, and vocational participation – domains essential to sustainable recovery outcomes.

Priority areas for methodological advancement

Forschungsprioritäten:

To strengthen the evidence foundation for MT in SUD treatment, future investigations must address:

von MCIDs über Langzeitstudien und Sicherheitsmonitoring zu partizipativen Designs

- **Development of population-specific interpretation frameworks:** Establish MCIDs tailored to SUD populations for accurate clinical significance determination
- **Extension of research timelines:** Design studies incorporating treatment phases exceeding six months with minimum twelve-month FU assessments
- **Integration of safety monitoring tools:** Deploy validated adverse event assessment frameworks throughout intervention and FU periods
- **Standardisation through core outcome sets:** Adopt consensus-based outcome selection guidelines to enhance cross-study comparability and reduce reporting heterogeneity
- **Expansion of outcome domains:** Incorporate instruments capturing socioeconomic (e.g., occupational functioning) and social parameters (e.g., social participation)
- **Participatory research design:** Position service user perspectives centrally in outcome measure selection and study methodology development

Broadening the research perspective beyond discipline-specific boundaries

disziplinübergreifende Perspektive erweitert Instrumentenauswahl

MT research in SUD contexts must transcend discipline-specific limitations. Given the nascent state of rigorous MT research methodology compared to established mental health research traditions, this review deliberately extended its scope beyond MT-exclusive instruments. The manual search encompassed measurement tools widely implemented across mental health domains, particularly those addressing socioeconomic dimensions relevant to SUD recovery. This cross-disciplinary approach facilitates identification of thoroughly validated instruments with demonstrated psychometric properties, even when not originally designed for MT and/or SUD applications. Continued integration of methodological innovations from mature mental health research fields will be essential for advancing MT research quality and clinical impact.

Instrument evaluation: balancing established validity against application-specific constraints

etablierte Validität erfordert kontextspezifische Bewertung

Analysis of the 14 selected measuring instruments confirms that, although all demonstrate robust psychometric properties in the literature, their deployment in MT for SUD requires a critical appraisal of context-specific advantages and limitations. Practitioners and researchers must evaluate these considerations within their particular clinical settings and target populations.

Auswahl der Messinstrumente nach therapeutischem Ziel und MT-Interventionstyp

Beyond setting- and population-specific adaptation, instrument selection should also ensure alignment with the therapeutic goals of the specific MT intervention being applied. For instance, instruments measuring motivational readiness (e.g., URICA, ICR) may be particularly well-suited to receptive interventions such as lyric analysis, which target motivation and problem recognition, whereas tools assessing interpersonal problems (e.g., IIP-SC) may be more appropriate for active, relational interventions such as group improvisation or songwriting, given that these approaches foster interpersonal communication and group cohesion. Selecting instruments that are conceptually congruent with the intended mechanisms of change not only strengthens construct validity but also enhances the sensitivity of outcome measurement to the specific effects that a given MT approach aims to produce.

Strategic framework for future investigation

Table 5-1: Strategic framework for future investigation

Primary research imperatives	Complementary enhancement strategies
Define population-specific MCIDs enabling clinically meaningful outcome interpretation	Strengthen social and socioeconomic outcome measurement through validated multidimensional instruments
Generate extended-duration empirical evidence through prolonged treatment and follow-up protocols	Implement systematic adverse event surveillance using standardised assessment tools
Create standardised measurement frameworks grounded in consensus-driven core outcome sets	Develop and validate culturally adapted instruments for diverse SUD populations

zukünftige Forschungsrahmen: primäre und ergänzende Strategien

Implementation of this comprehensive, methodologically rigorous, and participant-centred research agenda is essential to accurately determine clinical and socioeconomic outcome parameters and long-term measuring instruments in MT and/or SUD research, ultimately optimising its therapeutic application.

6 References

[1] NORCE Norwegian Research Centre AS. Fighting Addictions, Improving Lives Through COmprehensive Drug Rehabilitation With Music (FALCO): International Randomised Controlled Trial. [Clinical trial, NCT07028983]. ClinicalTrials.gov. Accessed 16 March 2026. Available from: https://clinicaltrials.gov/study/NCT07028983. 2025.

[2] GBD Alcohol and Drug Use Collaborators. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Psychiatry. 2018;5(12):987-1012. Epub 20181101. DOI: 10.1016/S2215-0366(18)30337-7.

[3] European Monitoring Centre for Drugs and Drug Addiction. European Drug Report 2022: Trends and Developments. Luxemburg: Publications Office of the European Union, 2022.

[4] Svendsen T., Veseth M., McKay J., Bjornestad J., Erga A., Moltu C., et al. Securing Participant Engagement in Longitudinal Substance Use Disorder Recovery Research: A Qualitative Exploration of Key Retention Factors. J Psychosoc Rehabil Ment Health. 2021;8:247-259.

[5] Fisher K. Healthcare Utilization. Gu D. and Dupre M., editors: Springer International Publishing; 2019.

[6] Carrasquillo O. Health care utilization. Gellman M. and Turner J., editors. New York: Springer; 2013.

[7] Silverman M., Bourdaghs S., Abbazio J. and Riegelman A. A systematic review of music-induced substance craving. Musicae Scientiae. 2023;27(1). DOI: doi:10.1177/10298649211030314.

[8] Saarikallio S., Gold C. and McFerran K. Development and validation of the Healthy-Unhealthy Music Scale. Child Adolesc Ment Health. 2015;20(4):210-217. Epub 20150518. DOI: 10.1111/camh.12109.

[9] Silverman M. Music-based emotion regulation and healthy and unhealthy music use predict coping strategies in adults with substance use disorder: A cross-sectional study. Psychology of Music. 2021;49(3). DOI: doi:10.1177/0305735619854529.

[10] American Music Therapy Association (AMTA). What is Music Therapy? Accessed 30 March 2025. Available from: https://www.musictherapy.org/.

[11] Ghetti C., Chen X. J., Brenner A. K., Hakvoort L. G., Lien L., Fachner J., et al. Music therapy for people with substance use disorders. Cochrane Database of Systematic Reviews. 2022;5(5):CD012576. Epub 20220509. DOI: 10.1002/14651858.CD012576.pub3.

[12] Carter T. E. and Panisch L. S. A systematic review of music therapy for psychosocial outcomes of substance use clients. International Journal of Mental Health and Addiction. 2021;19(5):1551-1568. DOI: 10.1007/s11469-020-00246-8.

[13] Ghetti C. Incorporating Music Therapy into the Harm Reduction Approach to Managing Substance Use Problems. Music Therapy Perspectives. 2004;22(2):84-90.

[14] Association of the Scientific Medical Societies (AWMF). Screening, diagnosis and treatment of alcohol-related disorders. Accessed 5 May 2025. Available from: https://register.awmf.org/assets/guidelines/076-001l_S3-Screening-Diagnose-Behandlung-alkoholbezogene-Stoerungen_2025-01-verlaengert.pdf. 2020.

[15] Horesh T. Drug addicts and their music: a story of a complex relationship. In: Aldridge D F. J., eds., editor. Music Therapy and Addictions: Jessica Kingsley Publishers; 2010. p. 57-74.

[16] Lozon J. and Bensimon M. Music misuse: A review of the personal and collective roles of "problem music". Aggression and Violent Behavior. 2014;19(3):207-218. DOI: 10.1016/j.avb.2014.04.003.

[17] Lozon J. and Bensimon M. The mutual enhancing effect of music and drugs and its eventual cessation in the lives of people coping with substance abuse: A qualitative inquiry. Curr Psychol. 2024;43(8). DOI: doi:10.1007/s12144-023-04839-z.

[18] Wills G. I. Forty lives in the bebop business: mental health in a group of eminent jazz musicians. Br J Psychiatry. 2003;183:255-259. DOI: 10.1192/bjp.183.3.255.

[19] Tuastad L., Johansen B., Østerholt A., Nielsen I. and McIvor D. Being a person who plays in a band rather than being a person with a mental illness playing in a band: A qualitative study of stigma in the context of music therapy in mental health aftercare. Nordic Journal of Music Therapy. 2023;32(2). DOI: doi:10.1080/08098131.2022.2075437.

[20] Murphy K. M. Music therapy in addictions treatment: A systematic review of the literature and recommendations for future research. Music and Medicine. 2017;9(1):15-23.

[21] Hohmann L., Bradt J., Stegemann T. and Koelsch S. Effects of music therapy and music-based interventions in the treatment of substance use disorders: A systematic review. PLoS ONE [Electronic Resource]. 2017;12(11):e0187363. DOI: https://dx.doi.org/10.1371/journal.pone.0187363.

[22] Norwegian Directorate of Health. Nasjonal Faglige Retningslinje for Behandling Og Rehabilitering Av Rusmiddelproblemer Og Avhengighet [National Professional Guideline for Treatment and Rehabilitation of Substance Use Problems and Addiction]. 2016.

[23] Ferreri L., Mas-Herrero E., Cardona G., Zatorre R. J., Antonijoan R. M., Valle M., et al. Dopamine modulations of reward-driven music memory consolidation. Annals of the New York Academy of Sciences. 2021;1502(1):85-98. Epub 20210711. DOI: 10.1111/nyas.14656.

[24] Mas-Herrero E., Marco-Pallares J., Lorenzo-Seva U., Zatorre R. and Rodriguez-Fornells A. Individual Differences in Music Reward Experiences. Music Perception. 2013;31:118-138.

[25] Koelsch S. A coordinate-based meta-analysis of music-evoked emotions. Neuroimage. 2020;223:117350. Epub 20200906. DOI: 10.1016/j.neuroimage.2020.117350.

[26] Salimpoor V., Benovoy M., Larcher K., Dagher A. and Zatorre R. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nat Neurosci. 2011;14(2). DOI: doi:10.1038/nn.2726.

[27] Dube L. and Le Bel J. The content and structure of laypeople's concept of pleasure. Cogn Emot. 2003;17(2):263-295. DOI: 10.1080/02699930302295.

[28] Levitin D. J. and Tirovolas A. K. Current advances in the cognitive neuroscience of music. Annals of the New York Academy of Sciences. 2009;1156:211-231. DOI: 10.1111/j.1749-6632.2009.04417.x.

[29] Koch S. Arts and health: Active factors and a theory framework of embodied aesthetics. The Arts in Psychotherapy. 2017;54:85-91.

[30] Hillecke T. and Wilker F. Ein heuristisches Wirkfaktorenmodell der Musiktherapie. Erhaltenstherapie & Verhaltensmedizin. 2007;28(1):62-85.

[31] de Witte M., Orkibi H., Zarate R., Karkou V., Sajnani N., Malhotra B., et al. From Therapeutic Factors to Mechanisms of Change in the Creative Arts Therapies: A Scoping Review. Frontiers in Psychology. 2021;12:678397. Epub 20210715. DOI: 10.3389/fpsyg.2021.678397.

[32] Volkow N. D., Michaelides M. and Baler R. The Neuroscience of Drug Reward and Addiction. Physiol Rev. 2019;99(4):2115-2140. DOI: 10.1152/physrev.00014.2018.

[33] Antichi L., Zarate R. and Giannini M. Change Process Research in Music Therapy: Introducing a Transdisciplinary Framework. Creative Arts Educ Ther. 2024;10(1):21-40. DOI: 10.15212/CAET/2024/10/1.

[34] Ashford R., Brown A., Brown T., Callis J., Cleveland H., Eisenhart E., et al. Defining and operationalizing the phenomena of recovery: a working definition from the recovery science research collaborativ. Addiction Research & Theory. 2019.

[35] Brophy H., Dyson M. and Katherine R. K. Concept analysis of recovery from substance use. Int J Ment Health Nurs. 2023;32(1):117-127. Epub 20220918. DOI: 10.1111/inm.13066.

[36] Worley J. Recovery in Substance Use Disorders: What to Know to Inform Practice. Issues Ment Health Nurs. 2017;38(1):80-91. Epub 20161130. DOI: 10.1080/01612840.2016.1245375.

[37] Zemore S. E., Ziemer K. L., Gilbert P. A., Karno M. P. and Kaskutas L. A. Understanding the Shared Meaning of Recovery From Substance Use Disorders: New Findings From the What is Recovery? Study. Subst Abuse. 2023;17:11782218231199372. Epub 20230915. DOI: 10.1177/11782218231199372.

[38] Stokes M., Schultz P. and Alpaslan A. Narrating the journey of sustained recovery from substance use disorder. Subst Abuse Treat Prev Policy. 2018;13(1):35. Epub 20180927. DOI: 10.1186/s13011-018-0167-0.

[39] Wangensteen T. and Hystad J. A Comprehensive Approach to Understanding Substance Use Disorder and Recovery: Former Patients' Experiences and Reflections on the Recovery Process Four Years After Discharge from SUD Treatment. J Psychosoc Rehabil Ment Health. 2022;9(1):45-54. Epub 20210827. DOI: 10.1007/s40737-021-00233-9.

[40] Veseth M., Svendsen T. S., Nesvaag S., Moltu C., Davidson L. and Bjornestad J. „And then the rest happened" – A qualitative exploration of the role that meaningful activities play in recovery processes for people with a diagnosis of substance use disorder. Substance Abuse. 2022;43(1):260-266. Epub 20210702. DOI: 10.1080/08897077.2021.1941506.

[41] International Consortium for Health Outcomes Measuresment (ICHOM). Patient-Centred Outcome Sets: Addiction. Accessed 5 May 2025. Available from: https://www.ichom.org/patient-centered-outcome-measure/addiction/.2023.

[42] Black N., Chung S., Tisdale C., Fialho L. S., Aramrattana A., Assanangkornchai S., et al. An International, Multidisciplinary Consensus Set of Patient-Centered Outcome Measures for Substance-Related and Addictive Disorders. J Clin Med. 2024;13(7). Epub 20240408. DOI: 10.3390/jcm13072154.

[43] Dyer A., Bohnke J. R., Curran D., McGrath K. and Toner P. A systematic review of quality of life and health-related quality of life as outcomes in substance and behavioural addictions. Drug & Alcohol Review. 2023;42(7):1680-1700. Epub 20230713. DOI: 10.1111/dar.13717.

[44] National Institute on Drug Abuse (NIDA). Principles of drug addiction treatment: A research-based guide (3rd ed.). Department of Health and Human Services, National Institutes of Health. 2018.

[45] McLellan A. T., Lewis D. C., O'Brien C. P. and Kleber H. D. Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation. JAMA. 2000;284(13):1689-1695. DOI: 10.1001/jama.284.13.1689.

[46] Arbour S., Hambley J. and Ho V. Predictors and outcome of aftercare participation of alcohol and drug users completing residential treatment. Subst Use Misuse. 2011;46(10):1275-1287. Epub 20110526. DOI: 10.3109/10826084.2011.572941.

[47] Moos R. H. and Moos B. S. Treated and untreated alcohol-use disorders: course and predictors of remission and relapse. Eval Rev. 2007;31(6):564-584. DOI: 10.1177/0193841X07306749.

[48] Solar O. and Irwin A. A conceptual framework for action on the social determinants of health. Social Determinants of Health Discussion, Paper 2 (Policy and Practice). 2010.

[49] Jerling P. and Heyns M. Exploring Guided Imagery and Music as a well-being intervention: A systematic literature review. Nordic Journal of Music Therapy. 2020;29(4):371-390. DOI: 10.1080/08098131.2020.1737185.

[50] Silverman M. J., Bourdaghs S., Abbazio J. and Riegelman A. A systematic review of music-induced substance craving. Musicae Scientiae. 2023;27(1):137-175. DOI: 10.1177/10298649211030314.

[51] Mokkink L. B., Elsman E. B. M. and Terwee C. B. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. Qual Life Res. 2024;33(11):2929-2939. Epub 20240828. DOI: 10.1007/s11136-024-03761-6.

[52] Proctor E., Silmere H., Raghavan R., Hovmand P., Aarons G., Bunger A., et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health. 2011;38(2):65-76. DOI: 10.1007/s10488-010-0319-7.

[53] Beaton D. E., Bombardier C., Guillemin F. and Ferraz M. B. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976). 2000;25(24):3186-3191. DOI: 10.1097/00007632-200012150-00014.

[54] Tong A., Sainsbury P. and Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349-357. Epub 20070914. DOI: 10.1093/intqhc/mzm042.

[55] Mayring P. Qualitative content analysis: theoretical foundation, basic procedures and software solution. Accessed 5 May 2025. Available from: http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173. Klagenfurt: 2014.

[56] Mayring P. Qualitative Content Analysis: A Step-by-Step Guide. ISBN: 9781529766738. London: 2021.

[57] Berk M., Ng F., Dodd S., Callaly T., Campbell S., Bernardo M., et al. The validity of the CGI severity and improvement scales as measures of clinical effectiveness suitable for routine clinical use. J Eval Clin Pract. 2008;14(6):979-983. Epub 20080502. DOI: 10.1111/j.1365-2753.2007.00921.x.

[58] Morrens J., Mathews M., Popova V., Borentain S., Rive B., Gonzalez Martin Moro B., et al. Use of Clinical Global Impressions-Severity (CGI-S) to Assess Response to Antidepressant Treatment in Patients with Treatment-Resistant Depression. Neuropsychiatr Dis Treat. 2022;18:1127-1132. Epub 20220607. DOI: 10.2147/NDT.S358367.

[59] Forkmann T., Scherer A., Boecker M., Pawelzik M., Jostes R. and Gauggel S. The Clinical Global Impression Scale and the influence of patient or staff perspective on outcome. BMC Psychiatry. 2011;11:83. Epub 20110514. DOI: 10.1186/1471-244X-11-83.

[60] Derogatis L. R. and Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med. 1983;13(3):595-605.

[61] Field C. A., Adinoff B., Harris T. R., Ball S. A. and Carroll K. M. Construct, concurrent and predictive validity of the URICA: data from two multi-site clinical trials. Drug Alcohol Depend. 2009;101(1-2):115-123. Epub 20090120. DOI: 10.1016/j.drugalcdep.2008.12.003.

[62] Dozois D. J., Westra H. A., Collins K. A., Fung T. S. and Garry J. K. Stages of change in anxiety: psychometric properties of the University of Rhode Island Change Assessment (URICA) scale. Behav Res Ther. 2004;42(6):711-729. DOI: 10.1016/S0005-7967(03)00193-1.

[63] Boyett B., Wiest K., McLeod L. D., Nelson L. M., Bickel W. K., Learned S. M., et al. Assessment of craving in opioid use disorder: Psychometric evaluation and predictive validity of the opioid craving VAS. Drug Alcohol Depend. 2021;229(Pt B):109057. Epub 20210924. DOI: 10.1016/j.drugalcdep.2021.109057.

[64] Zanis D. A., McLellan A. T., Cnaan R. A. and Randall M. Reliability and validity of the Addiction Severity Index with a homeless sample. J Subst Abuse Treat. 1994;11(6):541-548. DOI: 10.1016/0740-5472(94)90005-1.

[65] Montgomery H., Miller W. and Tonigan J. Does Alcoholics Anonymous Involvement Predict Treatment Outcome? Journal of Substance Abuse Treatment. 1995;12(4):241-246.

[66] McHugh R. K., Trinh C. D., Griffin M. L. and Weiss R. D. Validation of the craving scale in a large sample of adults with substance use disorders. Addict Behav. 2021;113:106651. Epub 20201006. DOI: 10.1016/j.addbeh.2020.106651.

[67] McHugh R. K., Fitzmaurice G. M., Carroll K. M., Griffin M. L., Hill K. P., Wasan A. D., et al. Assessing craving and its relationship to subsequent prescription opioid use among treatment-seeking prescription opioid dependent patients. Drug Alcohol Depend. 2014;145:121-126. Epub 20141016. DOI: 10.1016/j.drugalcdep.2014.10.002.

[68] Lievaart M., Erciyes F., van der Veen F. M., van de Wetering B. J., Muris P. and Franken I. H. Validation of the cocaine versions of the Obsessive Compulsive Drug Use Scale and the Desires for Drug Questionnaire. Am J Drug Alcohol Abuse. 2015;41(4):358-365. Epub 20150526. DOI: 10.3109/00952990.2015.1043210.

[69] Franken I. H., Hendriksa V. M. and van den Brink W. Initial validation of two opiate craving questionnaires the obsessive compulsive drug use scale and the desires for drug questionnaire. Addict Behav. 2002;27(5):675-685. DOI: 10.1016/s0306-4603(01)00201-5.

[70] Hormes J. M., Coffey S. F., Drobes D. J. and Saladin M. E. The Obsessive Compulsive Cocaine Use Scale: development and initial validation of a self-rated instrument for the quantification of thoughts about cocaine use. Drug Alcohol Depend. 2012;120(1-3):250-254. Epub 20110903. DOI: 10.1016/j.drugalcdep.2011.07.024.

[71] Raabe A., Grusser S. M., Wessa M., Podschus J. and Flor H. The assessment of craving: psychometric properties, factor structure and a revised version of the Alcohol Craving Questionnaire (ACQ). Addiction. 2005;100(2):227-234. DOI: 10.1111/j.1360-0443.2005.00960.x.

[72] Connolly K. M., Coffey S. F., Baschnagel J. S., Drobes D. J. and Saladin M. E. Evaluation of the Alcohol Craving Questionnaire-Now factor structures: application of a cue reactivity paradigm. Drug Alcohol Depend. 2009;103(1-2):84-91. Epub 20090506. DOI: 10.1016/j.drugalcdep.2009.03.019.

[73] Hebden H. M., Votaw V. R., Herchenroeder L., Witkiewitz K. and Yeung E. W. Examining the validity of the drinking motives questionnaire for identifying reward and relief drinking subgroups in college students. Alcohol Clin Exp Res (Hoboken). 2023;47(11):2184-2196. Epub 20231004. DOI: 10.1111/acer.15195.

[74] Drummond D. C. and Phillips T. S. Alcohol urges in alcohol-dependent drinkers: further validation of the Alcohol Urge Questionnaire in an untreated community clinical population. Addiction. 2002;97(11):1465-1472. DOI: 10.1046/j.1360-0443.2002.00252.x.

[75] Bohn M. J., Krahn D. D. and Staehler B. A. Development and initial validation of a measure of drinking urges in abstinent alcoholics. Alcohol Clin Exp Res. 1995;19(3):600-606. DOI: 10.1111/j.1530-0277.1995.tb01554.x.

[76] Pasche S. C., Garner M., Baldwin D. S. and Sinclair J. M. "Craving": exploring the components of the Desires for Alcohol Questionnaire (DAQ) and the relation to the severity of alcohol problems. J Stud Alcohol Drugs. 2013;74(6):950-955. DOI: 10.15288/jsad.2013.74.950.

[77] Kramer J. R., Chan G., Hesselbrock V. M., Kuperman S., Bucholz K. K., Edenberg H. J., et al. A principal components analysis of the abbreviated Desires for Alcohol Questionnaire (DAQ). J Stud Alcohol Drugs. 2010;71(1):150-155. DOI: 10.15288/jsad.2010.71.150.

[78] Courtney K. E., Ashenhurst J., Bacio G., Moallem N., Bujarski S., Hartwell E., et al. Craving and subjective responses to alcohol administration: validation of the desires for alcohol questionnaire in the human laboratory. J Stud Alcohol Drugs. 2013;74(5):797-802. DOI: 10.15288/jsad.2013.74.797.

[79] Boudreaux E. D., Sullivan A., Abar B., Bernstein S. L., Ginde A. A. and Camargo C. A., Jr. Motivation rulers for smoking cessation: a prospective observational examination of construct and predictive validity. Addict Sci Clin Pract. 2012;7(1):8. Epub 20120608. DOI: 10.1186/1940-0640-7-8.

[80] Napper L. E., Wood M. M., Jaffe A., Fisher D. G., Reynolds G. L. and Klahn J. A. Convergent and discriminant validity of three measures of stage of change. Psychol Addict Behav. 2008;22(3):362-371. DOI: 10.1037/0893-164X.22.3.362.

[81] Polaschek D., Anstiss B. and Wilson M. The assessment of offending-related stage of change in offenders: psychometric validation of the URICA with male prisoners. Psychology, Crime & Law. 2008;16(4):305-325.

[82] Hasler G., Klaghofer R. and Buddeberg C. [The University of Rhode Island Change Assessment Scale (URICA)]. Psychother Psychosom Med Psychol. 2003;53(9-10):406-411. Der Fragebogen zur Erfassung der Veranderungsbereitschaft (FEVER) – Testung der deutschen Version der University of Rhode Island Change Assessment Scale (URICA) -. DOI: 10.1055/s-2003-42172.

[83] Kelly J. F. and Greene M. C. Beyond motivation: Initial validation of the commitment to sobriety scale. J Subst Abuse Treat. 2014;46(2):257-263. Epub 20130815. DOI: 10.1016/j.jsat.2013.06.010.

[84] Mays K. L., Clark D. L. and Gordon A. J. Treating addiction with tunes: a systematic review of music therapy for the treatment of patients with addictions. Substance Abuse. 2008;29(4):51-59. DOI: https://dx.doi.org/10.1080/08897070802418485.

[85] Gossop M., Darke S., Griffiths P., Hando J., Powis B., Hall W., et al. The Severity of Dependence Scale (SDS): psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. Addiction. 1995;90(5):607-614. DOI: 10.1046/j.1360-0443.1995.9056072.x.

[86] Olatunji B. O., Deacon B. J., Abramowitz J. S. and Tolin D. F. Dimensionality of somatic complaints: factor structure and psychometric properties of the Self-Rating Anxiety Scale. J Anxiety Disord. 2006;20(5):543-561. Epub 20050929. DOI: 10.1016/j.janxdis.2005.08.002.

[87] Mörtberg E., Krastev D. and Borg E. Psychometric properties of the Social Anxiety Scale (SAS) using psychophysical and categorical scaling. Nordic Psychology. 2025;1-15.

[88] Hinz A., Daig I., Petrowski K. and Brahler E. [Mood in the German population: norms of the Multidimensional Mood Questionnaire MDBF]. Psychother Psychosom Med Psychol. 2012;62(2):52-57. Epub 20120123. Die Stimmung in der deutschen Bevolkerung: Referenzwerte fur den Mehrdimensionalen Befindlichkeitsfragebogen MDBF. DOI: 10.1055/s-0031-1297960.

[89] Wilhelm P. and Schoebi D. Assessing Mood in Daily Life. Structural Validity, Sensitivity to Change, and Reliability of a Short-Scale to Measure Three Basic Dimensions of Mood. European Journal of Psychological Assessment 2007;23:258-267. DOI: 10.1027/1015-5759.23.4.258.

[90] House E. D., Arruda J. E., Andrasik F. and Grazzi L. The reliability and validity of the Visual Analog Mood Scales in non-English-speaking pain patients. Pain Pract. 2012;12(8):626-632. Epub 20120323. DOI: 10.1111/j.1533-2500.2012.00544.x.

[91] Calvete E. and Connor-Smith J. Automatic Thoughts and Psychological Symptoms: A Cross-Cultural Comparison of American and Spanish Students. Cognitive Therapy and Research. 2005;29(2):201-217. DOI: 10.1007/s10608-005-3165-2.

[92] Cao R., Cheng S., Tang W. and Song H. The Reliability and Validity of Automatic Thoughts Questionnaire. Chinese Journal of Clinical Psychology. 2001;101(6):108-109.

[93] Crawford J. R. and Henry J. D. The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. Br J Clin Psychol. 2004;43(Pt 3):245-265. DOI: 10.1348/0144665031752934.

[94] Watson D., Clark L. A. and Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol. 1988;54(6):1063-1070. DOI: 10.1037//0022-3514.54.6.1063.

[95] Zentner M., Grandjean D. and Scherer K. R. Emotions evoked by the sound of music: characterization, classification, and measurement. Emotion. 2008;8(4):494-521. DOI: 10.1037/1528-3542.8.4.494.

[96] Chełkowska-Zacharewicz M. and Janowski M. Polish adaptation of the Geneva Emotional Music Scale (GEMS): factor structure and reliability. Psychology of Music. 2020;49(5).

[97] Yun S. H. and Gallant W. Evidence-based clinical practice: the effectiveness of music-based intervention for women experiencing forgiveness/grief issues. J Evid Based Soc Work. 2010;7(5):361-376. DOI: 10.1080/15433710903323870.

[98] Tzelepi I., Sotiropoulou L., Bacopoulou F., Charalampopoulou M., Zigkiri E., Simos D. S., et al. Validation of the Greek Version of the State Shame and Guilt Scale (SSGS). Adv Exp Med Biol. 2023;1425:79-91. DOI: 10.1007/978-3-031-31986-0_8.

[99] Cavalera C., Pepe A., Zurloni V., Diana B. and Realdon O. A short version of the State Shame and guilt Scale (SSGS-8). 2017. 2017;24:99-106. DOI: 10.4473/TPM24.1.6.

[100] Gomez R., Stavropoulos V., Zarate D. and Palikara O. Symptom Checklist-90-Revised: A structural examination in relation to family functioning. PLoS ONE. 2021;16(3):e0247902. Epub 20210312. DOI: 10.1371/journal.pone.0247902.

[101] Ransom D., Ashton K., Windover A. and Heinberg L. Internal consistency and validity assessment of SCL-90-R for bariatric surgery candidates. Surg Obes Relat Dis. 2010;6(6):622-627. Epub 20100223. DOI: 10.1016/j.soard.2010.02.039.

[102] Franke G. H., Jaeger S., Glaesmer H., Barkmann C., Petrowski K. and Braehler E. Psychometric analysis of the brief symptom inventory 18 (BSI-18) in a representative German sample. BMC Med Res Methodol. 2017;17(1):14. Epub 20170126. DOI: 10.1186/s12874-016-0283-3.

[103] Asner-Self K. K., Schreiber J. B. and Marotta S. A. A cross-cultural analysis of the Brief Symptom Inventory-18. Cultur Divers Ethnic Minor Psychol. 2006;12(2):367-375. DOI: 10.1037/1099-9809.12.2.367.

[104] MacDonald C., Brophy K., Coroiu A., Braehler E. and Korner A. The Brief Symptom Inventory-9 (BSI-9): Development and validation in a German general population sample. BMC Psychol. 2024;12(1):413. Epub 20240730. DOI: 10.1186/s40359-024-01890-8.

[105] Oliveira A. F., Santos I. M., Fernandes S., Bem-Haja P. and Torres A. Validation study of the Functional Assessment of Cancer Therapy-Cognitive Function-Version 3 for the Portuguese population. BMC Psychol. 2022;10(1):305. Epub 20221214. DOI: 10.1186/s40359-022-01018-w.

[106] Hajj A., Salameh P., Khoury R., Hachem R., Sacre H., Chahine G., et al. Psychometric properties of the 37-item Functional Assessment of Cancer Therapy-Cognitive Function (FACT-Cog) scale. Future Oncol. 2022;18(33):3741-3753. Epub 20221108. DOI: 10.2217/fon-2022-0438.

[107] Vazquez O., Orozco M., Hernández A. and Contreras L. Psychometric Properties of the Functional Assessment Instrument for Cancer Therapies – Cognitive Function (FACT-COG) in the Mexican General Population. Open Access J Neurol Neurosurg 2020;14(1). DOI: 10.19080/OAJNN.2020.14.555879.

[108] Webster K., Cella D. and Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. Health Qual Life Outcomes. 2003;1:79. Epub 20031216. DOI: 10.1186/1477-7525-1-79.

[109] Niessen D., Schmidt I., Groskurth K., Rammstedt B. and Lechner C. M. The Internal-External Locus of Control Short Scale-4 (IE-4): A comprehensive validation of the English-language adaptation. PLoS ONE. 2022;17(7):e0271289. Epub 20220711. DOI: 10.1371/journal.pone.0271289.

[110] Wang D., Zhu F., Xi S., Niu L., Tebes J. K., Xiao S., et al. Psychometric Properties of the Multidimensional Scale of Perceived Social Support (MSPSS) Among Family Caregivers of People with Schizophrenia in China. Psychol Res Behav Manag. 2021;14:1201-1209. Epub 20210807. DOI: 10.2147/PRBM.S320126.

[111] Dahlem N. W., Zimet G. D. and Walker R. R. The Multidimensional Scale of Perceived Social Support: a confirmation study. J Clin Psychol. 1991;47(6):756-761. DOI: 10.1002/1097-4679(199111)47:6<756::aid-jclp2270470605>3.0.co;2-l.

[112] Zimet G., Dahlem N., Zimet S. and Farley G. The Multidimensional Scale of Perceived Social Support. Journal of Personality Assessment. 1988;52(1).

[113] Wittenborn A., Natamba B., Rainey M., Zlotnick C. and Johnson J. Suitability of the multidimensional scale of perceived social support as a measure of functional social support among incarcerated adults with major depressive disorder. J Community Psychol. 2022;48(3):960-976. DOI: 10.1002/jcop.22315.

[114] Luoma J. B., O'Hair A. K., Kohlenberg B. S., Hayes S. C. and Fletcher L. The development and psychometric properties of a new measure of perceived stigma toward substance users. Subst Use Misuse. 2010;45(1-2):47-57. DOI: 10.3109/10826080902864712.

[115] Spata A., Gupta I., Lear M., Lunze K. and Luoma J. Substance use stigma: A systematic review of measures and their psychometric properties. Drug Alcohol Depend Rep. 2024;6(11). DOI: 10.1016/j.dadr.2024.100237.

[116] Oxford University Innovation. PECUNIA Resource Use Measurement (PECUNIA RUM) Instrument. Accessed 19 June 2025. Available from: https://innovation.ox.ac.uk/outcome-measures/pecunia-resource-use-measurement-instrument/.

[117] Pokhilenko I., Janssen L. M. M., Paulus A. T. G., Drost R., Hollingworth W., Thorn J. C., et al. Development of an Instrument for the Assessment of Health-Related Multi-sectoral Resource Use in Europe: The PECUNIA RUM. Appl Health Econ Health Policy. 2023;21(2):155-166. Epub 20230109. DOI: 10.1007/s40258-022-00780-7.

[118] Abouzed M., Aljadani A., Gabr A., Elsheikh M., Almuqahwi M. S., Alrashidy M. F., et al. Psychometric Properties of the Arabic Version of the Addiction Severity Index (ASI-5): A Cross-Cultural Validation Study in Egypt and Saudi Arabia. Subst Abuse Rehabil. 2025;16:129-136. Epub 20250510. DOI: 10.2147/SAR.S519241.

[119] Dere J., Watters C. A., Yu S. C., Bagby R. M., Ryder A. G. and Harkness K. L. Cross-cultural examination of measurement invariance of the Beck Depression Inventory-II. Psychol Assess. 2015;27(1):68-81. Epub 20141013. DOI: 10.1037/pas0000026.

[120] Soderhamn U., Sundsli K., Cliffordson C. and Dale B. Psychometric properties of Antonovsky's 29-item Sense of Coherence scale in research on older home-dwelling Norwegians. Scand J Public Health. 2015;43(8):867-874. Epub 20150806. DOI: 10.1177/1403494815598863.

[121] Norozi E., Miri M., Soltani R., Eslami A., Harivandi A. and Dastjerdi R. Drug Avoidance Self Efficacy Scale (DASES): A cultural adaptation and validation study. Journal of Substance Use. 2016;21(5):449-454. DOI: 10.3109/14659891.2015.1056851.

[122] Ardakani A., Seghatoleslam T., Habil H., Jameei F., Rashid R., Zahirodin A., et al. Construct Validity of Symptom Checklist-90-Revised (SCL-90-R) and General Health Questionnaire-28 (GHQ-28) in Patients with Drug Addiction and Diabetes, and Normal Population. Iran J Public Health. 2016;45(4):451-459.

[123] McLellan A. T., Cacciola J. C., Alterman A. I., Rikoon S. H. and Carise D. The Addiction Severity Index at 25: origins, contributions and transitions. Am J Addict. 2006;15(2):113-124. DOI: 10.1080/10550490500528316.

[124] Health Research Board National Drugs Library. Drug avoidance self-efficacy scale (DASES). Outcomes measurement tool: attitudes & feelings – self-efficacy. Accessed 12 June 2025. Available from: https://www.drugsandalcohol.ie/26769/. 2024.

[125] The Prevention Technology Transfer Center (PTTC) Network. Importance, Confidence & Readiness Rulers. Accessed 11 June 2025. Available from: https://pttcnetwork.org/wp-content/uploads/2020/06/Readiness-Rulers-Description.pdf.

[126] Beecham J. and m K. Costing Psychiatric Interventions. In: Thornicroft G., Brewin C. and Wing J., editors. Measuring mental health needs. London: Royal College of Psychiatrists; 2001. p. 163-183.

[127] Ersche K. D., Turton A. J., Croudace T. and Stochl J. Who Do You Think Is in Control in Addiction? A Pilot Study on Drug-related Locus of Control Beliefs. Addict Disord Their Treat. 2012;11(4):173-223. DOI: 10.1097/ADT.0b013e31823da151.

[128] Caridi B., Ratti E., Vitacolonna G., Caravelli A. and Romeo V. Cocaine Craving in Substance Abuse: The Standup® Method as On-Line Treatmentfor Substance Use Disorders. J Integrated Health. 2025;4(1):359-372. DOI: doi.org/10.51219/JIH/romeo-vm/61.

[129] Luborsky L., Barber J. P., Siqueland L., Johnson S., Najavits L. M., Frank A., et al. The Revised Helping Alliance Questionnaire (HAq-II) : Psychometric Properties. J Psychother Pract Res. 1996;5(3):260-271.

[130] National Institute of Health. National Institute on Drug Abuse. Brief Substance Craving Scale. Accessed 10 June 2025. Available from: https://datashare.nida.nih.gov/instrument/brief-substance-craving-scale.

[131] Hopwood C. J., Pincus A. L., DeMoor R. M. and Koonce E. A. Psychometric characteristics of the Inventory of Interpersonal Problems-Short Circumplex (IIP-SC) with college students. J Pers Assess. 2008;90(6):615-618. DOI: 10.1080/00223890802388665.

[132] von Glischinski M., von Brachel R. and Hirschfeld G. How depressed is "depressed"? A systematic review and diagnostic meta-analysis of optimal cut points for the Beck Depression Inventory revised (BDI-II). Qual Life Res. 2019;28(5):1111-1118. Epub 20181119. DOI: 10.1007/s11136-018-2050-x.

[133] Basker M., Moses P. D., Russell S. and Russell P. S. The psychometric properties of Beck Depression Inventory for adolescent depression in a primary-care paediatric setting in India. Child Adolesc Psychiatry Ment Health. 2007;1(1):8. Epub 20070809. DOI: 10.1186/1753-2000-1-8.

[134] Wang F. M., Davis M. F. and Briggs F. B. Predicting self-reported depression after the onset of multiple sclerosis using genetic and non-genetic factors. Mult Scler. 2021;27(4):603-612. Epub 20200518. "DOI: 10.1177/1352458520921073.

[135] Fried E. I., van Borkulo C. D., Epskamp S., Schoevers R. A., Tuerlinckx F. and Borsboom D. Measuring depression over time. Or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. Psychol Assess. 2016;28(11):1354-1367. Epub 20160128. DOI: 10.1037/pas0000275.

[136] Chisholm D., Knapp M. R., Knudsen H. C., Amaddeo F., Gaite L. and van Wijngaarden B. Client Socio-Demographic and Service Receipt Inventory--European Version: development of an instrument for international research. EPSILON Study 5. European Psychiatric Services: Inputs Linked to Outcome Domains and Needs. Br J Psychiatry Suppl. 2000(39):s28-33. DOI: 10.1192/bjp.177.39.s28.

[137] Horowitz L., Alden L., Wiggins J. and Pincus A. Inventory of Interpersonal Problems. Accessed 20 November 2025. Available from: https://www.mindgarden.com/113-inventory-of-interpersonal-problems. 2000.

[138] Kobayashi H., Song C., Ikei H., Park B. J., Kagawa T. and Miyazaki Y. Diurnal Changes in Distribution Characteristics of Salivary Cortisol and Immunoglobulin A Concentrations. International Journal of Environmental Research & Public Health [Electronic Resource]. 2017;14(9). Epub 20170831. DOI: 10.3390/ijerph14090987.

[139] Science for Sport. Salivary Immunoglobulin A (S-IgA). Accessed 12 June 2025. Available from: https://www.scienceforsport.com/salivary-immunoglobulin-a-s-iga/. 2025.

[140] Immune Deficiency Foundation. Selective IgA deficiency. Accessed 12 June 2025. Available from: https://primaryimmune.org/understanding-primary-immunodeficiency/types-of-pi/selective-iga-deficiency. 2025.

[141] Pritchard B. T., Stanton W., Lord R., Petocz P. and Pepping G. J. Factors Affecting Measurement of Salivary Cortisol and Secretory Immunoglobulin A in Field Studies of Athletes. Front Endocrinol (Lausanne). 2017;8:168. Epub 20170724. DOI: 10.3389/fendo.2017.00168.

[142] Heiderscheit A. The effects of the Bonny Method of Guided Imagery and Music (GIM) on adults in chemical dependency treatment: Sense of coherence, salivary immunoglobulin a and interpersonal problems. (Doctoral dissertation). Minneapolis, USA: University of Minnesota; 2005. Accessed 5 May 2025. Available from: https://conservancy.umn.edu/handle/11299/160300.

[143] Button K. S., Kounali D., Thomas L., Wiles N. J., Peters T. J., Welton N. J., et al. Minimal clinically important difference on the Beck Depression Inventory--II according to the patient's perspective. Psychol Med. 2015;45(15):3269-3279. Epub 20150713. DOI: 10.1017/S0033291715001270.

[144] Wang Y. P. and Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. Braz J Psychiatry. 2013;35(4):416-431. Epub 20131223. DOI: 10.1590/1516-4446-2012-1048.

[145] Spinal Cord Injury Research Evidence (SCIRE) Projessional. Symptoms Checklist-90-Revised. Accessed 5 June 2025. Available from: https://scireproject.com/wp-content/uploads/2022/04/Clinician-Summary_SCL-90-v.7.0.pdf.

[146] Liao Z. and Quintana Y. Challenges to Global Standardization of Outcome Measures. AMIA Jt Summits Transl Sci Proc. 2021;2021:404-409. Epub 20210517.

[147] Svendsen T., Erga A., Hagen E., McKay J., Njå A., Årstad J., et al. How to maintain high retention rates in long-term research on addiction: a case report. Journal of Social Work Practice in the Addictions. 2017;17(4):374-387.

[148] Bjornestad J., McKay J. R., Berg H., Moltu C. and Nesvag S. How often are outcomes other than change in substance use measured? A systematic review of outcome measures in contemporary randomised controlled trials. Drug & Alcohol Review. 2020;39(4):394-414. Epub 20200309. DOI: 10.1111/dar.13051.

[149] Migchels C., Zerrouk A., Crunelle C. L., Matthys F., Gremeaux L., Fernandez K., et al. Patient Reported Outcome and Experience Measures (PROMs and PREMs) in substance use disorder treatment services: A scoping review. Drug Alcohol Depend. 2023;253:111017. Epub 20231103. DOI: 10.1016/j.drugalcdep.2023.111017.

[150] Svendsen T. S., Bjornestad J., Slyngstad T. E., McKay J. R., Skaalevik A. W., Veseth M., et al. "Becoming myself": how participants in a longitudinal substance use disorder recovery study experienced receiving continuous feedback on their results. Subst Abuse Treat Prev Policy. 2020;15(1):8. Epub 20200123. DOI: 10.1186/s13011-020-0254-x.

[151] Lancaster K. Rethinking recovery. Addiction. 2017;112(5):758-759. Epub 20160901. DOI: 10.1111/add.13552.

[152] Koelsch S. From Social Contact to Social Cohesion – The 7 Cs. Music and Medicine. 2013;5(4):204-209. DOI: 10.1177/1943862113508588.

[153] Sonke J., Pesata V., Colverson A., Morgan-Daniel J., Rodriguez A. K., Carroll G. D., et al. Relationships between arts participation, social cohesion, and well-being: an integrative review of evidence. Front Public Health. 2025;13:1589693. Epub 20250509. DOI: 10.3389/fpubh.2025.1589693.

[154] Vigdal M. I., Moltu C., Bjornestad J. and Selseng L. B. Social recovery in substance use disorder: A metasynthesis of qualitative studies. Drug & Alcohol Review. 2022;41(4):974-987. Epub 20220201. DOI: 10.1111/dar.13434.

[155] Vigdal M., Svendsen T., Moltu C., Bjornestad J. and Selseng L. Stories of building friendships during long-term recovery from problematic substance use. Qualitative Social Work. 2024;23(5):868-886.

[156] Hagen E., Erga A. H., Hagen K. P., Nesvag S. M., McKay J. R., Lundervold A. J., et al. One-year sobriety improves satisfaction with life, executive functions and psychological distress among patients with polysubstance use disorder. J Subst Abuse Treat. 2017;76:81-87. Epub 20170131. DOI: 10.1016/j.jsat.2017.01.016.

[157] Martinelli T., Nagelhout G., Bellaert L., Best D., Vanderplasschen W. and van de Mheen D. Comparing three stages of addiction recovery: Long-term recovery and its relation to housing problems, crime, occupation situation, and substance use. Drugs: Education Prevention and Policy. 2020. DOI: 10.1080/09687637.2020.1779182.

[158] Riedl H., Else B. A., Grunhaus C. and Holck U. Economic Evaluations of Music Therapy and Other Music-Based Interventions: A Scoping Review. J Music Ther. 2025;62(1). DOI: 10.1093/jmt/thae023.

[159] Petrou S., Murray L., Cooper P. and Davidson L. L. The accuracy of self-reported healthcare resource utilization in health economic studies. Int J Technol Assess Health Care. 2002;18(3):705-710. DOI: 10.1017/s026646230200051x.

[160] Schönegger R. Empirische Zusammenhänge zwischen Gesundheitskompetenz und Health Service Use: eine systematische Überprüfung im österreichi-schen Kontext. Salzburg: Paracelsus Medizinische Universität; 2024.

[161] Rozental A., Kottorp A., Forsstrom D., Mansson K., Boettcher J., Andersson G., et al. The Negative Effects Questionnaire: psychometric properties of an instrument for assessing negative effects in psychological treatments. Behav Cogn Psychother. 2019;47(5):559-572. Epub 20190315. DOI: 10.1017/S1352465819000018.

[162] Knapp M., Ardino V., Brimblecombe N., Evans-Lacko S., Iemmi V., King D., et al. Youth Mental Health: New Economic Evidence. London 2016. Accessed 5 June 2025. Available from: https://www.pssru.ac.uk/publications/pub-5160/.

[163] Maier M. Content analysis: advantages and disadvantages. The sage encyclopedia of communication research methods (Vol 4, pp 240-242) SAGE Publications, Inc2017.

[164] Cook S. and Heppner P. A Psychometric Study of Three Coping Measures. Educational and Psychological Measurement. 1997;57(906). DOI: 10.1177/0013164497057006002.

[165] Lyne K. and Roger D. A psychometric re-assessment of the COPE questionnaire. Personality and Individual Differences. 2000;29(2):321-335.

[166] Hasking P. A. and Oei T. P. Confirmatory factor analysis of the COPE questionnaire on community drinkers and an alcohol-dependent sample. J Stud Alcohol. 2002;63(5):631-640. DOI: 10.15288/jsa.2002.63.631.

[167] Donoghue K. Measuring Coping: Evaluating the Psychometric Properties of the COPE 2004. Accessed 5 June 2025. Available from: https://ro.ecu.edu.au/theses_hons/968.

[168] Oetting E. R. and Beauvais F. Orthogonal cultural identification theory: the cultural identification of minority adolescents. Int J Addict. 1990;25(5A-6A):655-685. DOI: 10.3109/10826089109077265.

[169] Bard-Pondarre R., Villepinte C., Roumenoff F., Lebrault H., Bonnyaud C., Pradeau C., et al. Goal Attainment Scaling in rehabilitation: An educational review providing a comprehensive didactical tool box for implementing Goal Attainment Scaling. J Rehabil Med. 2023;55:jrm6498. Epub 20230614. DOI: 10.2340/jrm.v55.6498.

[170] Grootenboer E. M., Giltay E. J., van der Lem R., van Veen T., van der Wee N. J. and Zitman F. G. Reliability and validity of the Global Assessment of Functioning Scale in clinical outpatients with depressive disorders. J Eval Clin Pract. 2012;18(2):502-507. Epub 20110111. DOI: 10.1111/j.1365-2753.2010.01614.x.

[171] Aas I. H. Guidelines for rating Global Assessment of Functioning (GAF). Annals of General Psychiatry. 2011;10:2. Epub 20110120. DOI: 10.1186/1744-859X-10-2.

[172] Vatnaland T., Vatnaland J., Friis S. and Opjordsmoen S. Are GAF scores reliable in routine clinical use? Acta Psychiatr Scand. 2007;115(4):326-330. DOI: 10.1111/j.1600-0447.2006.00925.x.

[173] Trajkovic G., Starcevic V., Latas M., Lestarevic M., Ille T., Bukumiric Z., et al. Reliability of the Hamilton Rating Scale for Depression: a meta-analysis over a period of 49 years. Psychiatry Res. 2011;189(1):1-9. Epub 20110126. DOI: 10.1016/j.psychres.2010.12.007.

[174] Bagby R. M., Ryder A. G., Schuller D. R. and Marshall M. B. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? Am J Psychiatry. 2004;161(12):2163-2177. DOI: 10.1176/appi.ajp.161.12.2163.

[175] Rohan K. J., Rough J. N., Evans M., Ho S. Y., Meyerhoff J., Roberts L. M., et al. A protocol for the Hamilton Rating Scale for Depression: Item scoring rules, Rater training, and outcome accuracy with data on its application in a clinical trial. J Affect Disord. 2016;200:111-118. Epub 20160420. DOI: 10.1016/j.jad.2016.01.051.

[176] Sharp R. The Hamilton Rating Scale for Depression. Occup Med (Lond). 2015;65(4):340. DOI: 10.1093/occmed/kqv043.

[177] Huss M., Leak G. and Davis S. A validation study of the N ovaco Anger Inventory. Bulletin of the Psychonomic Society. 1993;31(4):279-281. DOI: 10.3758/bf03334928

[178] Mills J. F., Kroner D. G. and Forth A. E. Novaco Anger Scale: reliability and validity within an adult criminal sample. Assessment. 1998;5(3):237-248. DOI: 10.1177/107319119800500304.

[179] Rubio-Belmonte C., Mayordomo-Rodriguez T., Marco-Ahullo A. and Aragones-Barbera I. Psychometric Validation of the Purpose in Life Test-Short Form (PIL-SF) in Individuals Diagnosed with Severe Mental Illness. Healthcare (Basel). 2024;12(20). Epub 20241018. DOI: 10.3390/healthcare12202082.

[180] Rubio-Belmonte C., Mayordomo Rodriguez T. and Garcia-Alandete J. Psychometric properties of the Purpose In Life-Short Form in the Spanish population. J Clin Psychol. 2023;79(4):1099-1112. Epub 20221123. DOI: 10.1002/jclp.23461.

[181] Dyck M. Assessing logotherapeutic constructs: Conceptual and psychometric status of the purpose in life and seeking of noetic goals tests. Clinical Psychology Review. 1987;7:439-447. DOI: 10.1016/0272-7358(87)90021-3.

[182] Zika S. and Chamberlain K. On the relation between meaning in life and psychological well-being. Br J Psychol. 1992;83 (Pt 1):133-145. DOI: 10.1111/j.2044-8295.1992.tb02429.x.

[183] Lee J., Baek H., Oh E., Kim J. Y. and Ko Y. G. Development and validation of the self-consciousness type scale. Frontiers in Psychology. 2024;15:1362324. Epub 20240705. DOI: 10.3389/fpsyg.2024.1362324.

[184] Brophy K., Emery M., Coroiu A., Albani C., Braehler E. and Körner A. Development and Validation of a Short Form of the German Self-Consciousness Scale for the General Population. European Journal of Psychology Open. 2024;83(4).

[185] McLellan A. T., Luborsky L., Cacciola J., Griffith J., Evans F., Barr H. L., et al. New data from the Addiction Severity Index. Reliability and validity in three centers. J Nerv Ment Dis. 1985;173(7):412-423. DOI: 10.1097/00005053-198507000-00005.

[186] McLellan A. T., Kushner H., Metzger D., Peters R., Smith I., Grissom G., et al. The Fifth Edition of the Addiction Severity Index. J Subst Abuse Treat. 1992;9(3):199-213. DOI: 10.1016/0740-5472(92)90062-s.

[187] Kosten T. R., Rounsaville B. J. and Kleber H. D. Concurrent validity of the addiction severity index. J Nerv Ment Dis. 1983;171(10):606-610. DOI: 10.1097/00005053-198310000-00003.

[188] Rogalski C. J. Factor structure of the Addiction Severity Index in an inpatient detoxification sample. Int J Addict. 1987;22(10):981-992. DOI: 10.3109/10826088709109693.

[189] Hendriks V. M., Kaplan C. D., van Limbeek J. and Geerlings P. The Addiction Severity Index: reliability and validity in a Dutch addict population. J Subst Abuse Treat. 1989;6(2):133-141. DOI: 10.1016/0740-5472(89)90041-x.

[190] Chernyak N. Self-reported health care utilization: measurement issues, data validity and implications for design of health surveys and economic evaluations. An empirical investigation among patients with diabetes mellitus. Düsseldorf: University Bremen; 2011. Accessed 5 September 2025. Available from: https://d-nb.info/1071993607/34.

[191] Kashner T. M., Stensland M. D., Lind L., Wicker A., Rush A. J., Golden R. M., et al. Measuring use and cost of care for patients with mood disorders: the utilization and cost inventory. Med Care. 2009;47(2):184-190. DOI: 10.1097/MLR.0b013e31818457b8.

[192] Neale J., Vitoratou S., Finch E., Lennon P., Mitcheson L., Panebianco D., et al. Development and Validation of 'Sure': A Patient Reported Outcome Measure (Prom) for Recovery from Drug and Alcohol Dependence. Drug Alcohol Depend. 2016;165:159-167. Epub 20160615. DOI: 10.1016/j.drugalcdep.2016.06.006.

[193] Van Wert M. and Gus E. The Substance Use Recovery Evaluator: A closer look among adults presenting to an intensive outpatient program for co-occurring disorders [Research Brief]. Center for Practice Transformation, University of Minnesota & NUWAY. 2024.

[194] Chung C. C. Y., Fung J. L. F., Lui A. C. Y., Chan M. C. Y., Ng Y. N. C., Wong W. H. S., et al. Client Service Receipt Inventory as a standardised tool for measurement of socio-economic costs in the rare genetic disease population (CSRI-Ra). Sci Rep. 2021;11(1):23837. Epub 20211213. DOI: 10.1038/s41598-021-03379-5.

[195] Mirandola M., Bisoffi G., Bonizzato P. and Amaddeo F. Collecting psychiatric resources utilisation data to calculate costs of care: a comparison between a service receipt interview and a case register. Soc Psychiatry Psychiatr Epidemiol. 1999;34(10):541-547. DOI: 10.1007/s001270050173.

[196] Kessler R. C., Ames M., Hymel P. A., Loeppke R., McKenas D. K., Richling D. E., et al. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. J Occup Environ Med. 2004;46(6 Suppl):S23-37. DOI: 10.1097/01.jom.0000126683.75201.c5.

[197] Kawakami N., Inoue A., Tsuchiya M., Watanabe K., Imamura K., Iida M., et al. Construct validity and test-retest reliability of the World Mental Health Japan version of the World Health Organization Health and Work Performance Questionnaire Short Version: a preliminary study. Ind Health. 2020;58(4):375-387. Epub 20200314. DOI: 10.2486/indhealth.2019-0090.

[198] Kessler R. C., Barber C., Beck A., Berglund P., Cleary P. D., McKenas D., et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). J Occup Environ Med. 2003;45(2):156-174. DOI: 10.1097/01.jom.0000052967.43131.51.

[199] AlHeresh R., LaValley M. P., Coster W. and Keysor J. J. Construct Validity and Scoring Methods of the World Health Organization: Health and Work Performance Questionnaire Among Workers With Arthritis and Rheumatological Conditions. J Occup Environ Med. 2017;59(6):e112-e118. DOI: 10.1097/JOM.0000000000001044.

[200] Beecham J., Johnson S. and EPCST Group. European Psychiatric Care Assessment Team. The European Socio-Demographic Schedule (ESDS): rationale, principles and development. Acta Psychiatr Scand Suppl. 2000;405:33-46.

[201] Sirey J. A., Meyers B. S., Teresi J. A., Bruce M. L., Ramirez M., Raue P. J., et al. The Cornell Service Index as a measure of health service use. Psychiatr Serv. 2005;56(12):1564-1569. DOI: 10.1176/appi.ps.56.12.1564.

[202] National Institute of Health. Office of Behavioral and Social Sciences Research. Measuring Socioeconomic Status. Accessed 17 June 2025. Available from: https://obssr.od.nih.gov/sites/obssr/files/Measuring-Socioeconomic-Status.pdf.

[203] Nakao K. and Treas J. The 1989 socioeconomic index of occupations: Construction from the 1989 occupational prestige scores (GSS Methodological Report No. 74). National Opinion Research Center. 1992.

[204] Caston R. Dimensions of occupational inequality and Duncan's socioeconomic index. Sociological Forum. 1989;4(3):329-348. DOI: https://doi.org/10.1007/BF01115017.

[205] Nam C., LaRocque J., Powers M. and Holmberg J. Occupational status scores: Stability and change. Proceedings of the American Statistical Association, Social Statistics Section. 1975:570-575.

[206] Nam C. and Boyd M. Occupational status in 2000: Over a century of census-based measurement. Population Research and Policy Review. 2004;23(4):327-358. DOI: 10.1023/B:POPU.0000040045.51228.3.

[207] Chandola T. and Jenkinson C. The new UK National Statistics Socio-Economic Classification (NS-SEC); investigating social class differences in self-reported health status. J Public Health Med. 2000;22(2):182-190. DOI: 10.1093/pubmed/22.2.182.

[208] Rose D. and Pevalin D. The National Statistics Socio-economic Classification: Unifying official and sociological approaches to the conceptualisation and measurement of social class (ISER Working Paper Series No. 2001-04). Institute for Social and Economic Research, University of Essex. Accessed 5 September 2025. Available from: https://hdl.handle.net/10419/92194. 2001.

[209] Asadi-Lari M., Majdzadeh R., Mansournia M. A., Nedjat S., Mohammad K. and Cheraghian B. Construction and validation of CAPSES scale as a composite indicator of SES for health research: an application to modeling social determinants of cardiovascular diseases. BMC Public Health. 2023;23(1):293. Epub 20230209. DOI: 10.1186/s12889-023-15206-9.

[210] Eriksson M. and Lindstrom B. Antonovsky's sense of coherence scale and the relation with health: a systematic review. J Epidemiol Community Health. 2006;60(5):376-381. DOI: 10.1136/jech.2005.041616.

[211] Sardu C., Mereu A., Sotgiu A., Andrissi L., Jacobson M. K. and Contu P. Antonovsky's Sense of Coherence Scale: Cultural Validation of Soc Questionnaire and Socio-Demographic Patterns in an Italian Population. Clin Pract Epidemiol Ment Health. 2012;8:1-6. Epub 20120214. DOI: 10.2174/1745017901208010001.

[212] Khazaee-Pool M., Naghibi S. A., Pashaei T., Chaleshgar-Kordasiabi M., Daneshnia M. and Ponnet K. Drug Abstinence Self-Efficacy Scale (DASES): psychometric properties of the Farsi version. Subst Abuse Treat Prev Policy. 2021;16(1):1. Epub 20210103. DOI: 10.1186/s13011-020-00336-9.

[213] Kamarudin E., Alam N., Zainudin Z., Zakaria N., Kari D. and Baharudin M. Translation and Validation of the Drug Avoidance Self-Efficacy Scale for High-Risk Populations in Malaysia. Open Journal of Social Sciences. 2024;12:602-618. DOI: 10.4236/jss.2024.1211042.

[214] Citlik Saritas S. and Erci B. Psychometric evaluation of the Drug Avoidance Self-Efficacy Scale. Journal of Substance Use. 2019;25(5):492-496.

[215] McLellan A., Luborsky L., Woody G. and O'Brien C. Addiction Severity Index-5th Edition (ASI-5th Edition). Accessed 6 June 2025. Available from: https://eprovide.mapi-trust.org/instruments/addiction-severity-index-5th-edition. 2025.

[216] Khalil M. S. Reliability and confirmatory factor analysis of the arabic version of the University of Rhode Island Change Assessment (URICA). Alcohol & Alcoholism. 2011;46(2):138-142. Epub 20110201. DOI: 10.1093/alcalc/agr003.

[217] Lerdal A., Moe B., Digre E., Harding T., Kristensen F., Grov E. K., et al. Stages of Change – continuous measure (URICA-E2): psychometrics of a Norwegian version. J Adv Nurs. 2009;65(1):193-202. Epub 20081114. DOI: 10.1111/j.1365-2648.2008.04842.x.

[218] Figlie N., Dunn J., Bazán J. and Laranjeira R. The Stages of Change Scale Among Brazilian Alcohol Dependents. Addictive Disorders & Their Treatment. 2005;4(4):161-165. DOI: 10.1097/01.adt.0000186366.77296.66.

[219] Mander J., Schaller G., Bents H., Dinger U., Zipfel S. and Junne F. Increasing the treatment motivation of patients with somatic symptom disorder: applying the URICA-S scale. BMC Psychiatry. 2017;17(1):240. Epub 20170703. DOI: 10.1186/s12888-017-1400-5.

[220] Chen S. H., Lai H. R., Chen S. R., Lin P. C., Chou K. R. and Lee P. H. Validity and Reliability of a Chinese-Language Instrument for Continuous Assessment of Exercise Stages of Change in Adults. Journal of Nursing Research. 2019;27(4):e37. DOI: 10.1097/jnr.0000000000000310.

[221] Ingo E., Brannstrom K. J., Andersson G., Lunner T. and Laplante-Levesque A. Measuring motivation using the transtheoretical (stages of change) model: A follow-up study of people who failed an online hearing screening. Int J Audiol. 2016;55 Suppl 3:S52-58. Epub 20160520. DOI: 10.1080/14992027.2016.1182650.

[222] University of Connecticut. Mechanisms Underlying Mind-Body Interventions & Measurement of Emotional Well-Being: M3EWB. Sense of Coherence Scale (SOC). Accessed 3 June 2025. Available from: https://m3ewb.research.uconn.edu/measure/sense-of-coherence-scale-soc/.

[223] CorHealthOntario. Assessing Motivation: The importance-confidence ruler technique. Accessed 10 June 2025. Available from: https://www.corhealthontario.ca/HMP-Assessing-Motivation.pdf.

[224] Malinao D. A Concurrent Analysis of Drug Relapse Risk, Drug Avoidance Self-Efficacy, and Aftercare Experiences among Aftercare Program Clients. Technium Social Sciences Journal. 2024;57:134-182.

[225] Miller W. and Rollnick S. Motivational interviewing: Preparing people for change (2nd ed.). New York: The Guilford Press. 2002.

[226] McGill University. Addiction Severity Index (ASI/EuropASI). Accessed 6 June 2025. Available from: https://www.mcgill.ca/mappro/information-hub/measures-library/patients/substance-use-addiction/asi-euroasi.

[227] Pearson Clinical Assessment India. Symptom Checklist-90-Revised. Accessed 5 June 2025. Available from: https://www.pearsonclinical.in/products/programs/symptom-checklist-90-revised.html.

[228] Zhou J., Yu J., Zhou Y. and Qiu J. Study of item text in the Chinese Symptom Checklist-90. Medicine. 2021;100(11):e24841. DOI: 10.1097/MD.0000000000024841.

[229] Derogatis L. Symptom Checklist-90 – Revised (SCL-90-R) [Database record]. APA PsycTests. No year specified. DOI: https://doi.org/10.1037/t01210-000.

[230] Yeh M. Y., Lee L. W. and Hwang F. M. The Chinese version of the drinking-related locus-of-control scale: a confirmatory factor analysis. J Subst Abuse Treat. 2008;34(3):333-339. Epub 20070705. DOI: 10.1016/j.jsat.2007.05.002.

[231] American Thoracic Society. Accessed 5 June 2025. Available from: https://www.thoracic.org/members/assemblies/assemblies/srn/questionaires/bdi-ii.php.

[232] Pearson Clinical Assessment India. Beck Depression Inventory. Accessed 5 June 2025. Available from: https://www.pearsonclinical.in/products/programs/beck-depression-inventory.html.

[233] Andrade-Gonzalez N. and Fernandez-Liria A. Spanish adaptation of the Revised Helping Alliance Questionnaire (HAq-II). J Ment Health. 2015;24(3):155-161. DOI: 10.3109/09638237.2015.1036975.

[234] Kermarrec S., Kabuth B., Bursztejn C. and Guillemin F. French adaptation and validation of the helping alliance questionnaires for child, parents, and therapist. Can J Psychiatry. 2006;51(14):913-922. DOI: 10.1177/070674370605101407.

[235] Wu L. Z., Roche M. J., Dowgwillo E. A., Wang S. and Pincus A. L. A Chinese translation of the inventory of interpersonal problems-short circumplex. J Pers Assess. 2015;97(2):153-162. Epub 20141103. DOI: 10.1080/00223891.2014.971461.

[236] Thomas A., Brähler E. and Strauss B. IIP-32: Entwicklung, Validierung und Normierung einer Kurzform des Inventars zur Erfassung interpersonaler Probleme. Diagnostica. 2011;57:68-83. DOI: 10.1026/0012-1924/a000034.

[237] Lo Coco G., Mannino G., Salerno L., Oieni V., Di Fratello C., Profita G., et al. The Italian Version of the Inventory of Interpersonal Problems (IIP-32): Psychometric Properties and Factor Structure in Clinical and Non-clinical Groups. Frontiers in Psychology. 2018;9:341. Epub 20180319. DOI: 10.3389/fpsyg.2018.00341.

[238] Lindberg R., Lindgren M. and Wennberg M. Psychometric properties of the Swedish version of the Inventory of Interpersonal Problems-32 in a clinical population. Nordic Journal of Psychiatry. 2018;72(8):606-611.

[239] Gold C., Rolvsjord R., Aaro L. E., Aarre T., Tjemsland L. and Stige B. Resource-oriented music therapy for psychiatric patients with low therapy motivation: protocol for a randomised controlled trial [NCT00137189]. BMC Psychiatry. 2005;5:39. Epub 20051031. DOI: 10.1186/1471-244X-5-39.

[240] Beck A. T., Ward C. H., Mendelson M., Mock J. and Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry. 1961;4:561-571. DOI: 10.1001/archpsyc.1961.01710120031004.

[241] Bruscia K. Defining Music Therapy (3rd Edition). Approaches: An Interdisciplinary Journal of Music Therapy. 2018;10(2):417.

Appendix

Extraction table of the instruments measuring clinical and socioeconomic outcome parameters and tools assessing (serious) adverse events (Part 1 of the report)

Table A-1: Clinical and socioeconomic outcome parameters with their (not) valid/reliable measuring instruments and tools assessing (serious) adverse events

Outcome parameters	**Measuring instruments and tools assessing (serious) adverse events**						
Authors:	**Ghetti 2022 [11]**	**Hohmann [21]**	**Carter 2021 [12]**	**Jerling 2020 [49]**	**Mays 2008 [84]**	**Murphy 2017 [20]**	**Silverman 2023 [50]**
Substance use parameters							
Alcohol/substance use	Telephone call	Substance use report					
LSD experience		LSD Session Survey, Objective Check List for LSD Experience, third-party reports					
(Drug avoidance) self-efficacy		**DASES**, **VAS**	**DASES**			**DASES**	
Recovery parameters							
Problem severity		ASI-NAV, **ASI**					
Adoption of principles and practices		**GAATOR 2.1**					
Craving parameters							
(Intensity of) substance/alcohol craving	**BSCS**, **OCDUS**, **ACQ-SF-R**	**BSCS**, **AUQ**, 7-point scale, **VAS**	**BSCS**, **ACQ-SF**			**BSCS**	**VAS**, (descriptive) Likert-type scale, **ACQ-NOW**, Reasons for Drinking Questionnaire, **DMQ-R**, Pavlov's classical experiments[10], **AUQ**, semi-structured interviews, **DAQ**
Coping skills/Knowledge of triggers and coping skills		**SOC**, A-COPE[11], Lists of triggers and coping skills					
Motivational parameters							
Motivation/readiness/eagerness for treatment/change	**CESI**, **CMR**, **ICR**, **RTCQ-TV**, **SOCRATES**, **URICA**	**URICA**, **RTCQ-TV**, **SOCRATES** (short version), 7-point Likert scale, Questionnaire[12], **CESI**, **ICR**, **CMR**	**SOCRATES**, **URICA**, **RTCQ-TV**, **CMR**, **CESI**	**ICR**		**ICR**, **URICA**, **RTCQ-TV**, **CMR**, **SOCRATES** (version 8D), **CESI**	

[10] The interviewer described Pavlov's classical experiments and then asked interviewees: "In your case, just like with Pavlov's dogs, are there any bells which trigger a craving for alcohol and, if so, what are they?"

[11] Tested, but not recommended.

[12] 5-point Likert scale, yes-no questions, open-ended questions

Outcome parameters	Measuring instruments and tools assessing (serious) adverse events						
Authors:	**Ghetti 2022 [11]**	**Hohmann [21]**	**Carter 2021 [12]**	**Jerling 2020 [49]**	**Mays 2008 [84]**	**Murphy 2017 [20]**	**Silverman 2023 [50]**
Motivation to reach/maintain sobriety	**CSS**, QMAD, 7-point Likert scale	7-point Likert scale					
Immediate goal attainment		GAF (goal)					
Quality of life parameters							
Sleep quality		Questionnaire[13]					
Manageability, comprehensibility and meaningfulness of life/Orientation to life/Sense of coherence				**SOC**		**SOC**	
Well-being		**FACIT-F**					
Mood-related parameters							
Mood		**MDMQ**, **VAS**[14], **VAMS** with combined emotions, Rogers' (1981) Happy/Sad Faces, questionnaire[12]			Rogers' Happy/Sad Faces Assessment Tool, **VAMS**		
Depression	**BDI**, HRSD, **SDS**, Likert scale	**BDI(-II)**, HRSD, 10-point Likert scale	**BDI(-II)**, HRSD	**BDI**		**BDI(-II/Vs)**	
Depressogenic thought frequency		**ATQ**					
Positive and negative affect		**PANAS**					
Sadness		**7-point VAS**					
Anxiety	**SAS**	**7-point VAS**, 7-point or 10-point Likert scale, **STAI**				**STAI**	
Stress		10-point Likert scale					
Shame, pride and guilt			**SSGS**				
Emotional valence and arousal		7-point Scale, **GEMS-9**					
Anger		**7-point VAS**, 10-point Likert scale, NAI					
Forgiveness and grief		**FGPS**					
Psychological, psychiatric, physiological and cognitive parameters							
Self-consciousness			SCS				
Constructive self-awareness and perceived purpose in life			PIL				
Psychiatric symptoms		**SCL-90-R**					

[13] Time to fall asleep, frequency of nightmares, mood on the following morning, sleep interruptions

[14] Distressed, sad, irritated, calm, satisfied

Outcome parameters	**Measuring instruments and tools assessing (serious) adverse events**						
Authors:	**Ghetti 2022 [11]**	**Hohmann [21]**	**Carter 2021 [12]**	**Jerling 2020 [49]**	**Mays 2008 [84]**	**Murphy 2017 [20]**	**Silverman 2023 [50]**
Physical and psychiatric symptoms		**BSI**					
Physiological aspects		Physiological measurement					
General functioning		**CGI**, GAF					
Withdrawal (symptoms)		ARSW	ARSW			ARSW	
Cognitive functioning		**FACT-Cog**					
Cultural, spiritual, and locus of control parameters							
Cultural identity		ANCI					
Comfort and strength derived from spirituality		**FACIT**					
Locus of control		(Abbreviated) **I-E, DRIE**	**DRIE**			**DRIE**	
Biomedical parameters							
Immune function				**Salivary Immuno-globulin A Test**		**Saliva samples**	
Music-, therapist-, and treatment-related parameters							
Music perception		Questions about chromesthetic music perception					
Music preference		LSD Music Preference Questionnaire					
Content		Lyric analysis					
Attitudes towards MT/therapist/treatment groups		MT questionnaire			**VAS**		
Trust in therapist			**WFPTS**			**WFPTS**	
(Therapeutic) working alliance		**HAQ**	**WAI-SR**			**HAQ-III**, **WAI-S**	
Quality of therapist-client relationship			**HAQ-II**				
Participation		Assessment tool			Assessed by therapist		
On-task behaviour		Therapist rating					
Importance of MT		**VAMS** with combined emotions					
MT characteristics		MT questionnaire					
Perceived effectiveness and enjoyment		25-point analogue scales, Questionnaire[12], 5-point Likert scale					
Perception of treatment			Non-validated Likert scale				
Helpfulness		7-point Likert scale					

Outcome parameters	Measuring instruments and tools assessing (serious) adverse events						
Authors:	**Ghetti 2022 [11]**	**Hohmann [21]**	**Carter 2021 [12]**	**Jerling 2020 [49]**	**Mays 2008 [84]**	**Murphy 2017 [20]**	**Silverman 2023 [50]**
Intervention assessment compared to other groups		25-point analogue scales					
Comfort		7-point Likert scale					
Socioeconomic and social parameters							
Retention in treatment (and completion)	Number of participants remaining at the end of treatment	Number of days attending MT sessions					
Attendance		Percent of attendance at any given time			Number of sessions		
Medication adherence		Number of sessions					
Perception of social support			**MSPSS**				
Perception of stigma			**PSAS**				
Interpersonal (relationship) problems				**IIP-SC**		**IIP-SC**	
Client attitudes, psychosocial functioning		20-Item Hudson Psychosocial Screening Instrument			20-Item Hudson Psychosocial Screening Instrument		
Adverse events							
	NR	NR	NR	NR	NR	NR	NR
Serious adverse events							
	NR	NR	NR	NR	NR	NR	NR

Bold: *Tested for validity/reliability.*

Abbreviations: *A-COPE … Adolescent Coping Orientation to Problems Experienced.*[15] *ACQ-SF-R(/-NOW) … Alcohol Craving Questionnaire-Short-Form-Revised (/Now).*[16] *ANCI … American Indian/Alaska Native Cultural Identity Scale.*[17] *ARSW … Adjective Rating Scale for Withdrawal.* ASI … Addiction Severity Index.*[18] *ASI-NAV … Addiction Severity Index, Native American Version. ATQ … Automatic Thoughts Questionnaire.*[19] *AUQ … Alcohol urge questionnaire.*[20] *BDI … Beck Depression Inventory.* BSCS … Brief Substance Craving Scale.**[21] *BSI … Brief*

[15] Tested for validity/reliability, however, heterogeneous results (COPE): reliable (undergraduate students) [164]; not recommended (employees) [165], (alcohol-dependent) [166], (thesis) [167]

[16] Tested for validity/reliability: [71, 72]

[17] Tested for validity/reliability: "only useful with youth and not with adults" [168]

[18] Tested for validity/reliability: [64]

[19] Tested for validity/reliability: validity (American/Spanish students; ATQ-R (revised version)) [91]; In Chinese language [92]

[20] Tested for validity/reliability: [74, 75]

[21] However, valid and reliable for SUD, "strong internal consistency reliability … and adequate concurrent and discriminant validity." [66]; Opioid Craving Scale: Validity/reliability [67]

Symptom Inventory.[22] *CESI … (Texas Christian University Treatment Motivation Scale) – Client Evaluation of Self at Intake.* CGI … Clinical Global Impression Severity Scale.*[23] *CMR … Circumstances, Motivation, and Readiness Scales for Substance Abuse Treatment.* CSS … Commitment to Sobriety Scale.*[24] *DAQ … (Multidimensional) Desires for Alcohol Questionnaire.*[25] *DASES … Drug Avoidance Self-Efficacy Scale.* DMQ-R … Drinking Motives Questionnaire – Revised.*[26] *DRIE … Drinking-Related Internal-External Locus of Control Scale.* FACIT … Functional Assessment of Chronic Illness Therapy: Spiritual Questions.*[27] *FACIT-F … Functional Assessment of Chronic Illness Therapy: Fatigue.*[27] *FACT-Cog … Functional Assessment of Cancer Therapy – Cognitive Functions.*[28] *FGPS … Forgiveness Grief Perspectives Scale.*[29] *GAATOR 2.1 … General Alcoholics Anonymous Tools of Recovery.*[30] *GAF (goal) … Goal attainment form.*[31] *GAF … Global Assessment of Function Scale.*[32] *GEMS-9 … Geneva Emotions in Music Scale.*[33] *HAQ(-II) … Revised Helping Alliance Questionnaire for therapist and client.* HRSD … Hamilton Rating Scale for Depression.*[34] *ICR … Importance, Confidence, Readiness (motivational) Ruler.**[35] *I-E … Internal vs. external locus of control scale.*[36] *IIP-SC … Inventory of Interpersonal Problems: Short Circumplex form.* LSD … Lysergsäurediethylamid. MDMQ … Multidimensional Mood Questionnaire.*[37] *MSPSS … Multidimensional Scale of Perceived Social Support.*[38] *MT … music therapy. NAI … Novaco Anger Inventory Short Form.*[39] *NR … not reported. OCDUS … Obsessive Compulsive Drug Use Scale.*[40] *PANAS … Positive and negative affect scale.*[41]

---

22 Tested for validity/reliability: (BSI-18) [102]; (BSI-18) [103]; (BSI-9, German general population) [104]

23 Tested for validity/reliability: [57]; reliability (patients with treatment-resistant depression) [58]; however, not recommended [59]

24 Tested for validity/reliability: [83]

25 Tested for validity/reliability: [76-78]

26 Tested for validity/reliability (college students): [73]

27 Tested for validity/reliability (FACIT-G, general): [108]

28 Tested for validity/reliability: (Portuguese population) [105]; (Lebanese cancer patients) [106]; (Mexican general population) [107]

29 Tested for validity/reliability: however, only based on a small sample (n=21), future factor validation studies in a large clinical sample are necessary [97]

30 Tested for validity/reliability: reliability [65]

31 Tested for validity/reliability: Goal Attainment Scaling (GAS) form: "Reliability and validity will be dependent on the experience of the goal setter in patient pathology and the interventions proposed, as well as their expertise in the GAS processes." [169]

32 Tested for validity/reliability: (clinical outpatients with depressive disorders) "poor inter-rater reliability as well as poor discriminant validity" [170]; not recommended [171, 172]

33 Tested for validity/reliability: [95]; reliability (Polish version) [96]

34 Tested for validity/reliability: reliable, however, "some HRSD items (e.g., "loss of insight") do not appear to possess a satisfactory reliability" [173]; adequate internal reliability, poor content validity [174]; adequate reliability/validity, however, poor item-level inter-rater reliability, test-retest reliability, and content validity [175]; furthermore, relatively high inter-rater reliability [175]; adequate internal, inter-rater and retest reliability (for global score), weaker (for individual items) [176]; inter-rater reliability affected by inter-viewer's level of training and whether a structured interview guide is provided [176]

35 However, valid and reliable, e.g. for smoking cessation [79]; in Poland, e.g., a group of researchers is currently validating the ICR in a group of people with SUD

36 Tested for validity/reliability: [109]

37 Tested for validity/reliability: reliability (German population) [88, 89]

38 Tested for validity/reliability: (Chinese caregivers of people with schizophrenia) [110]; (students) [111]; (university undergraduates) [112]; (incarcerated adults with major depressive disorder) [113]

39 Tested for validity/reliability: further research is needed on reliability and validity; discrepancies regarding internal reliability and test-retest reliability [177]; implications for clinical use are discussed (adult criminal sample) [178]

40 Tested for validity/reliability (cocaine and heroin): [68-70]

41 Tested for validity/reliability: (non-clinical sample) [93]; PANAS-10 [94]

PIL … Purpose in Life Test.[42] *PSAS … Perceived Stigma of Addiction Scale.*[43] *QMAD … Questionnaire of Motivation for Abstaining from Drugs. RTCQ-TV … Readiness to Change Questionnaire – Treatment Version.* SAS … Self-Rating Anxiety Scale.*[44] *SCL-90-R … Symptom Checklist 90-R.*[45] *SCS … Self-Consciousness Scale.*[46] *SDS … Severity of Dependence Scale.*[47] *SOC … Sense of Coherence Scale.* SOCRATES … Stages of Change Readiness and Treatment Eagerness Scale.* SSGS … State Shame and Guilt Scale.*[48] *STAI … State-Trait Anxiety Inventory.* SUD … substance use disorder. URICA … University of Rhode Island Change Assessment Scale.**[49] *VAMS … Visual Analogue Mood Scale.*[50] *VAS … Visual Analogue Scale.*[51] *WAI-S(R) … Working Alliance Inventory – Short (Revised).* WFPTS … Wake Forest Physician Trust Scale.**

* *Tested for validity and reliability. Information on whether tested or not tested for validity/reliability from [20].*

Extraction table of the instruments measuring *only socioeconomic* outcome parameters (Part 1 of the report)

Table A-2: Socioeconomic data collection tools

Socioeconomic data collection tools	Abbreviation	Mode of data collection	Link to tool	Outcomes assessed	Reliability	Validity
Addiction Severity Index	ASI	Semi-structured interview	https://adai.washington.edu/instruments/pdf/addiction_severity_index_baseline_followup_4.pdf	Patient's substance use and related problems: medical, employment/support status, alcohol, drug, legal, family/social, and psychiatric problem areas	Yes ([185] concurrent validity & test-retest; [186] test-retest only for new questions – legal, family/social)	Yes ([185] concurrent validity & discriminant validity; [187] concurrent validity; [188] construct validity; [189] concurrent validity for dutch version)
Utilisation and Cost Inventory	UAC-I	Structured interview	NA	Use of inpatient and outpatient care over a 3-month period for patients with mood disorders [190]	Yes ([191] test-retest subsample)	Yes ([191] criterion validity – concurrent validity/ convergent construct validity)

[42] Tested for validity/reliability: (PIL-SF (Spanish short form, severe mental illness)) [179]; validity (non-clinical Spanish undergraduates) [180]; (logotherapeutic construct) [181] however, criticised [182]

[43] Tested for validity/reliability: [114, 115]

[44] Tested for validity/reliability: (undergraduate students) [86]; non-clinical Swedish adults [87]

[45] Tested for validity/reliability: (hierarchical taxonomy of psychopathology) [100]; (bariatric patients) [101]

[46] Tested for validity/reliability: validity (undergraduate students), however limitations are reported [183]; "there is a lack of consensus on the factor structure" [184]

[47] Tested for validity/reliability (heroin, cocaine, and amphetamine): [85]

[48] Tested for validity/reliability: reliability (Greek version) [98]; (SSGS-8 (short version), Italian undergraduate students) [99]

[49] Drug-use version: Good internal reliability and non-significant discriminant validity; however, questionable convergent validity [80]; However, valid and reliable for male prisoners [81]; German version valid for eating disorders/complex problem behaviours [82]

[50] Tested for validity/reliability (non-English-speaking pain patients): [90]

[51] Tested for validity/reliability (opioid craving): [63]

Socioeconomic data collection tools	Abbreviation	Mode of data collection	Link to tool	Outcomes assessed	Reliability	Validity
Substance Use Recovery Evaluator	SURE	Questionnaire	http://www.recoveryanswers.org/assets/substance_use_recovery_evaluator_sure.pdf	Personal recovery from drug and/or alcohol dependence, i.e., person's own perspective, including substance use, self-care, relationships, material resources, and outlook on life	Yes ([192] internal consistency (Cronbach's alpha), test-retest; [193] (Cronbach's alpha))	Yes ([192] face validity/ content validity/ construct validity; [193] convergent validity)
Client service receipt inventory	CSRI	Semi-structured inventory/ questionnaire	https://www.pssru.ac.uk/csri/files/2017/10/CSRI-UK-Generic-mental-health.pdf	Economic data: sociodemographic information, usual living situation, employment and income, service receipt, and medication profile	Yes ([194] alternate form reliability between two versions)	Yes ([194] criterion validity/face validity/ convergent and discriminant validity) Adaptation for an Italian version "ICAP" [195] criterion validity/ concurrent validity
Client Socio-Demographic and Service Receipt Inventory[52]	CSSRI	Semi-structured inventory/questionnaire	https://www.dirum.org/assets/downloads/634462380166178864-CSSRI%20-%20EU.pdf (CSSRI-EU)	Service receipt and associated data alongside assessment of patient outcomes, i.e., socio-demographics, usual living situation, employment and income, service receipt, medication profile [136]	NA	Yes (for EU version [136] but face validity only)
Health and Work Performance Questionnaire	WHO-HPQ	Questionnaire	https://www.hcp.med.harvard.edu/hpq/info.php	Occupational mental health (work-related outcomes): work, demographics	Yes ([196] test-retest; for the Japanese short version: [197] test-retest)	Yes ([196] concurrent validity/criterion validity/ face validity; [198] concurrent validity/face validity; for the Japanese short version: [197] construct validity/ convergent validity; [199] construct validity)
Socioeconomic and Medical Status Questionnaire	SES questionnaire	Questionnaire	https://cdn-links.lww.com/permalink/ppt/a/ppt_25_3_2013_05_01_pathare_200536_sdc1.pdf	Social (i.e., family history, education, insurance, employment, income) and medical history	NA	NA
The European Socio-demographic Schedule	ESDS	Structured questionnaire	https://onlinelibrary.wiley.com/doi/10.1111/j.0902-4441.2000.t01-1-acp28-02.x (contact authors for access)	Sociodemographic characteristics associated with service use and psychiatric morbidity [200]	NA	Yes ([200] but only cross-cultural validity/content validity)
Cornell Services Index	CSI	Structured questionnaire	NA (contact authors for access)	Snapshot of service use patterns (frequency and duration) across types, providers, and sites of service among adults who seek mental health care; services aggregated into 4 types: outpatient psychiatric or psychological, outpatient medical, professional support, and intensive services [190]	Yes ([201] inter-rater/test-retest)	NA

[52] More specialised and often more comprehensive version or adaptation of the broader CSRI

Socioeconomic data collection tools	Abbreviation	Mode of data collection	Link to tool	Outcomes assessed	Reliability	Validity
ProgrammE in Costing, resource use measurement and outcome evaluatioN for Use in multi-sectoral National and International health economic evaluations – Resource Use Measurement Instrument	PECUNIA-RUM	Questionnaire	NA (access needed)	Resource use in all relevant sectors for costing from a societal perspective in the adult population: health and social care, education, (criminal) justice, productivity losses, and informal care [116]	Yes [117]	Yes ([117] only content validity)
Duncan Socioeconomic Index	Duncan SEI	Scale	NA	Socioeconomic status (subjective ratings): occupational prestige/status, implicating educational attainment and income (reflecting the overall socioeconomic standing of occupations within a society) [202]	NA	Yes ([203] for new version from 1989 "1989 SEI" convergent/criterion/ discriminant validity; [204] construct validity)
Nam-Powers Occupational Status Score	Nam-Powers OSS	Scale	NA	Socioeconomic status (objective ratings), i.e., income and educational status associated with an occupation (reflecting the position in the socioeconomic hierarchy) [202]	Only longitudinal stability [205]	Yes ([206] construct validity[53]; [205] content validity)
National Statistics Socioeconomic classification	NS-SEC	Scale	https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010#the-questions-to-ask	Socioeconomic status: aspects of one's job and employment relations to calculate a scalar measure of socioeconomic status [202]	NA	Yes ([207] construct validity; [208] criterion/construct validity)
Composite SocioEconomic Status scale	CAPSES	Scale	NA	Socioeconomic status: material capital (wealth, income, trust funds, etc.), human capital (skills, credentials, abilities, etc.), and social capital (instrumental relationships such as being friends with doctors and lawyers) [202]	NA	Yes ([209] criterion validity/construct validity)

Note: Unless otherwise stated, information is from the link provided.

Abbreviations: ASI … Addiction Severity Index. CAPSES … Composite SocioEconomic Status scale. CSRI … Client service receipt inventory. CSSRI … Client Socio-Demographic and Service Receipt Inventory. ESDS … European Socio-demographic Schedule. HPQ … Health and Work Performance Questionnaire. NA … not available. NS-SEC … National Statistics Socioeconomic classification. OSS … Occupational Status Score. PECUNIA … ProgrammE in Costing, resource use measurement and outcome evaluatioN for Use in multi-sectoral National and InternAtional health economic evaluations. RUM … Resource Use Measurement. SEI … Socioeconomic Index. SES … socioeconomic status. SURE … Substance Use Recovery Evaluator. UAC-I … Utilisation and Cost Inventory. WHO … World Health Organisation.

[53] Validity was based on the definition (that status = education + income) and on demonstrating that the tool correctly differentiates social groups.

Extraction tables of the measuring instruments and their characteristics (Part 2 of the report)

Table A-3: Measuring instruments and their characteristics (part 1/2)

Parameters:	Substance use parameters	Recovery parameters	Craving parameters	Motivational parameters		Quality of life parameters
Measuring instruments:	**DASES [12, 20, 21]**	**ASI [21]**	**BSCS [11, 12, 20, 21]**	**URICA [11, 12, 20, 21]**	**ICR [11, 20, 21, 49]**	**SOC [20, 49]**
General characteristics						
Outcome measured	(Drug avoidance) self-efficacy	Problem severity	(Intensity of) substance/alcohol craving	Motivation/readiness/ eagerness for treatment/change	Motivation/readiness/ eagerness for treatment/change	Manageability, comprehensibility and meaningfulness of life/Orientation to life/Sense of coherence
Mode of assessment	Questionnaire	Questionnaire (interview)	Questionnaire	Questionnaire	Visual analogue scale, short rating instrument	Questionnaire
Post-assessment in treatment and latest FU where measuring instrument was applied (of included reviews)	Duration of treatment: 6 months [21] FU: 1 month [12, 20, 21]	Duration of treatment: 6 months [21] FU: 1 month [21]	Duration of treatment: 6 months [21] FU: 3 months [11]	Duration of treatment: 6 months [21] FU: 3 months [11]	Duration of treatment: 2 years (10 sessions) [49] FU: 3 months [11]	Duration of treatment: 2 years (10 sessions) [49] FU: 1 month [20]
Number of languages in which instruments are available	Min. 5	Min. 13	Min. 2	Min. 7 plus 4 non-validated translations	All	Min. 33 [210, 211]
Types of languages in which instruments are available	English, Farsi/Persian [212], Malay [213], Turkish [214], Chinese (no validation study found for Chinese)	English, Czech, Danish, Dutch, French, Hungarian, Italian, Lithuanian, Polish[54], Portuguese, Russian, Spanish, Swedish [215], Arabic, German[55]	English, Arabic, Bulgarian (no validation study found for Arabic and Bulgarian)	English, Chinese, German, Portuguese (Brazil), Norwegian, Arabic [216-220], Portuguese, French[56], Swedish [221], Polish (no validation study found for Portuguese, French, and Swedish)	All	E.g. Croatian, Czech, English, German, Hebrew, Hungarian, Icelandic, Indonesian, Japanese, Lithuanian, Norwegian, Polish, Portuguese (Brazil), Russian, Slovenian, Spanish, Swedish, Turkish, Welsh [222]
Type of MT intervention in which the instrument was applied (active/receptive/both) (of included reviews)	Both	Both	Receptive [49], both	Receptive [49], both	Receptive [49], both	Receptive [49], both
Setting of assessment where instrument was applied (e.g. in-/outpatient treatment, detoxification, rehabilitation) (of included reviews)	Therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential	Therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential	Short-term inpatient detoxification setting, residential substance use treatment facility, chemical dependency treatment programme within a residential health facility, facilities	Short-term inpatient detoxification setting, residential substance use treatment facility, chemical dependency treatment programme within a residential health facility, facilities	Short-term inpatient detoxification setting, residential substance use treatment facility, chemical dependency treatment programme within a residential health facility, facilities	NR [49], short-term detoxification units [20]

[54] New version of ASI5th Polish version (with no validation yet) https://osf.io/ev2w8/overview

[55] https://www.mcgill.ca/mappro/information-hub/measures-library/patients/substance-use-addiction/asi-euroasi

[56] https://habitslab.umbc.edu/urica/

Parameters:	Substance use parameters	Recovery parameters	Craving parameters	Motivational parameters		Quality of life parameters
Measuring instruments:	**DASES [12, 20, 21]**	**ASI [21]**	**BSCS [11, 12, 20, 21]**	**URICA [11, 12, 20, 21]**	**ICR [11, 20, 21, 49]**	**SOC [20, 49]**
Setting of assessment where instrument was applied (e.g. in-/outpatient treatment, detoxification, rehabilitation) (of included reviews) *(continuation)*	therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient genderspecific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21], rehabilitation programme, inpatients at a detoxification unit [12], short-term detoxification units [20]	therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient genderspecific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21]	offering inpatient and outpatient substance use treatment [11], therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient genderspecific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21], rehabilitation programme, inpatients at a detoxification unit [12], short-term detoxification units [20]	offering inpatient and outpatient substance use treatment [11], therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient genderspecific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21], rehabilitation programme, inpatients at a detoxification unit [12], short-term detoxification units [20]	offering inpatient and outpatient substance use treatment [11], therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient gender-specific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21], NR [49], short-term detoxification units [20]	
Limitations	Self-report bias, where individuals may underestimate or exaggerate symptoms; restricted use to young multiple drug users (aged 16-30) [124]; findings from culturally adapted versions should be viewed as preliminary in nature, with a need for future research to determine its impact on long-term treatment outcomes [121]	Culturally adapted versions (e.g. Arabic ASI-5) have shown lower consistency in drug and legal domains [118]; intensive ASI training (2 full days) [123]; reliability decreases significantly when performed by interviewers who are less intensively trained and monitored [123]; challenges in adequately representing the clinical needs of special groups (e.g., women, ethnic minorities, mental/physical health problems) [123]; problems with composite scores commonly used in treatment outcome studies reported [123]; problems in specific domains	Self-report bias, where individuals may underestimate or exaggerate symptoms; lack of significant correlation with objective behavioural outcomes, such as actual cocaine use [130]; may offer a limited representation of the multidimensional nature of craving (e.g., reward, relief, and obsessive craving) [128]	Self-report bias, where individuals may underestimate or exaggerate symptoms; questionable convergent validity [80]	Its effectiveness is highly dependent on the helper's proficiency in Motivational Interviewing skills [125]; may not resonate universally with all individuals as a natural way of conceptualising mixed feelings about change [125]	Self-report bias, where individuals may underestimate or exaggerate symptoms; the relationship between SOC and physical health is not strong, potentially due to mental and emotional factors being intertwined in the scale [120]; requires ongoing validation, particularly confirmatory factor analysis, in diverse populations to ensure cross-cultural applicability and generalisability [120]

Parameters:	Substance use parameters	Recovery parameters	Craving parameters	Motivational parameters		Quality of life parameters
Measuring instruments:	**DASES [12, 20, 21]**	**ASI [21]**	**BSCS [11, 12, 20, 21]**	**URICA [11, 12, 20, 21]**	**ICR [11, 20, 21, 49]**	**SOC [20, 49]**
Limitations *(continuation)*		(e.g., medical domain, employment and self-support, family and social relationships, psychiatric problems) [123]				
Application characteristics						
Number of items in each instrument	16	Not a fixed number of items; it is a semi-structured interview covering 7 problem areas	16 [130]	32 [219] (drug version)	2-3 questions [223]	29 [222]
Total score range	16-112 [224]	0-64 or a severity rating from 0 to 9 for each of the 7 areas	0-12	-2 to 14	0-10 for 1 dimension [223] or 0-30 for 3 dimensions [225]	29-203
Interpretation of score	Higher self-efficacy in avoiding substance use	Severe	Severe	Greater readiness for change	Higher motivation for change (and readiness to change) [223]	Stronger SOC, reflecting a greater ability to manage tension, perceive life as comprehensible, manageable, and meaningful
Minimum clinically important difference	SUD: none Other populations: none	SUD: none Other populations: none	SUD: none Other populations: none	SUD: none Other populations: none	SUD: none Other populations: none	SUD: none Other populations: none
Testing time (min)	10	45-60 [226] + follow up 25-30 min	10 [130]	5-10	Few minutes	10-15
Instrument user (e.g. self-assessment by service user, clinician-, performance reported)	Self-reported	Clinician-rated [226]	Self-reported [130]	Self-reported [81]	Clinician-administered or self-completion format	Self-reported [222]
Costs and regulations for use	Free to download https://www.nd.gov.hk/pdf/bdf-2010R2-q03-eng.pdf[57]	Free to download	Free to download https://www.millisecond.com/download/library/bscs[58]	Free to download	Free to download	Free [222]

Note: If no references are stated, information is obtained from the downloaded measuring instrument itself.

Abbreviations: ASI … Addiction Severity Index. BSCS … Brief Substance Craving Scale. DASES … Drug Avoidance Self-Efficacy Scale. FU … follow-up. ICR … Importance, Confidence, Readiness (motivational) Ruler. NR … not reported. SOC … Sense of Coherence Scale. URICA … University of Rhode Island Change Assessment Scale.

[57] Can only be used for the purpose of evaluating projects sponsored by Beat Drugs Fund Association

[58] Free with an Inquisit license for online or in-person psychological research

Table A-4: Measuring instruments and their characteristics (part 2/2)

Parameters:	Mood-related parameters	Psychological, psychiatric, physiological and cognitive parameters	Cultural, spiritual, and locus of control parameters	Music-, therapist-, and treatment-related parameters	Socioeconomic and social parameters			Biomedical parameters
Measuring instruments:	**BDI [11, 12, 20, 21, 49]**	**SCL-90-R [21]**	**DRIE [12, 20, 21]**	**HAQ(-II) [12, 20, 21]**	**IIP-SC (IIP-32) [20, 49]**	**PECUNIA-RUM**	**CSRI**	**Salivary Immunoglobulin A Test [20, 49]**
General characteristics								
Outcome measured	Depression	Psychiatric symptoms	Locus of control	(Therapeutic) working alliance and quality of therapist-client relationship	Interpersonal (relationship) problems	Resource use in all relevant sectors for costing from a societal perspective in the adult population: health and social care, education, (criminal) justice, productivity) [116]	Economic data: sociodemographic information, usual living situation, employment and income, service receipt (= resource use)	Immune function
Mode of assessment	Questionnaire	Questionnaire	Questionnaire	Questionnaire	Questionnaire	Questionnaire	Questionnaire (interview)	Salivary sample
Post-assessment in treatment and latest FU where measuring instrument was applied (of included reviews)	Duration of treatment: 2 years (10 sessions) [49] FU: 3 months [11]	Duration of treatment: 6 months [21] FU: 1 month [21]	Duration of treatment: 6 months [21] FU: 1 month [12, 20, 21]	Duration of treatment: 6 months [21] FU: 1 month [12, 20, 21]	Duration of treatment: 2 years (10 sessions) [49] FU: 1 month [20]	Not applicable	Not applicable	Duration of treatment: 2 years (10 sessions) [49] FU: 1 month [20]
Number of languages in which instruments are available	Min. 18-20 [144]	Min. 8 (validated) plus 14 (non-validated) [145, 227-229]	Min. 2 [230]	Min. 3	Min. 7 (validated) and 8 (non-validated)	Multiple languages in future [117]	Min. 7	Not applicable
Types of languages in which instruments are available	English, Spanish, Chinese, Dutch, Finnish, French (Canadian), German, Korean, Polish, Swedish, Arabic, Turkish, Hindi, Marathi, Tamil, Malayalam, Japanese, Persian, Kannada, Xhosa [231, 232]	English, French, Spanish, Hindi, Kannada, Tamil, Marathi, Malayalam [145, 227], Spanish, French, German, Russian, Italian, Portuguese, Dutch, Swiss, Japanese, Chinese, Korean, Vietnamese, Hebrew, and Arabic (validation unknown) [229]	English, Chinese [230]	English, Spanish, [233], French [234]	English, Chinese, Dutch, Spanish [235], German [236], Italian [237], Swedish [238], Arabic, Polish, Finnish, French, Greek, Korean, Malay, Slovenian (not validated)	English, Dutch, German (validated) [116]	English, German, Danish, Spanish, Italian, Dutch, Portuguese (Brazil) (validation unknown) [194]	Not applicable
Type of MT intervention in which the instrument was applied (active/receptive/both) (of included reviews)	Receptive [49], both	Both	Both	Both	Receptive [49], both	Not applicable	Not applicable	Receptive [49], both

Parameters:	Mood-related parameters	Psychological, psychiatric, physiological and cognitive parameters	Cultural, spiritual, and locus of control parameters	Music-, therapist-, and treatment-related parameters	Socioeconomic and social parameters			Biomedical parameters
Measuring instruments:	**BDI [11, 12, 20, 21, 49]**	**SCL-90-R [21]**	**DRIE [12, 20, 21]**	**HAQ(-II) [12, 20, 21]**	**IIP-SC (IIP-32) [20, 49]**	**PECUNIA-RUM**	**CSRI**	**Salivary Immunoglobulin A Test [20, 49]**
Setting of assessment where instrument was applied (e.g. in-/outpatient treatment, detox, rehabilitation) (of included reviews)	Short-term inpatient detoxification setting, residential substance use treatment facility, chemical dependency treatment programme within a residential health facility, facilities offering inpatient and outpatient substance use treatment [11], therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient genderspecific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21], rehabilitation programme, inpatients at a detoxification unit [12], NR [49], short-term detoxification units [20],	Therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient genderspecific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21]	Therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient genderspecific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21], rehabilitation programme, inpatients at a detoxification unit [12], short-term detoxification units [20]	Therapeutic community for drug abusers, rehabilitation programme during prison sentence, inpatient alcohol addiction treatment centre, inpatient antialcoholic treatment, residential therapeutic community for SUD, inpatient alcohol abuse treatment, in- and outpatient rehabilitation unit (detoxification and day patients), inpatient dual diagnosis treatment unit, inpatient genderspecific residential program, inpatient nonmedical detoxification facility, inpatient detoxification unit [21], rehabilitation programme, inpatients at a detoxification unit [12], short-term detoxification units [20]	NR [49], short-term detoxification units [20]	Not applicable	Not applicable	NR [49], short-term detoxification units [20]

Parameters:	Mood-related parameters	Psychological, psychiatric, physiological and cognitive parameters	Cultural, spiritual, and locus of control parameters	Music-, therapist-, and treatment-related parameters	Socioeconomic and social parameters			Biomedical parameters
Measuring instruments:	**BDI [11, 12, 20, 21, 49]**	**SCL-90-R [21]**	**DRIE [12, 20, 21]**	**HAQ(-II) [12, 20, 21]**	**IIP-SC (IIP-32) [20, 49]**	**PECUNIA-RUM**	**CSRI**	**Salivary Immunoglobulin A Test [20, 49]**
Limitations	Self-report bias, where individuals may underestimate or exaggerate symptoms; often misused as a diagnostic test, despite being designed as a rating scale to measure the severity of depression [133]; limited scope, as BDI only assesses symptoms of depression and does not cover other important factors like family history or genetic predisposition [134]; a single administration provides a static measure, capturing a snapshot rather than dynamic symptom fluctuations over time [135]; not specific to any one culture, meaning its accuracy may vary in different cultural contexts without specific validation [119]; copyrighted, requiring a fee for each copy used, which can limit accessibility; the optimal cut-off points (threshold values) that maximise its diagnostic utility vary considerably depending on the sample studied [132]	Self-report bias, where individuals may underestimate or exaggerate symptoms; sensitive to linguistic and cultural factors, necessitating replication and re-validation in different settings, languages, and cultures [122]; possibly, overlaps between symptom domains; licensed [145]	Self-report bias, where individuals may underestimate or exaggerate symptoms; criticised for its mixture of both personalised and general statements [127]; changes in scores during treatment may be confounded by the fact that personalised items may not be sensitive when alcohol is no longer consumed, limiting the responsiveness in later recovery stages [127]	Self-report bias, where individuals may underestimate or exaggerate symptoms; alliance levels, as measured by the HAQ(-II), were not associated with pretreatment psychiatric severity or level of depression [129]	Self-report bias, where individuals may underestimate or exaggerate symptoms; brevity comes with an inherent trade-off in internal consistency, although still described as adequate [131]; larger gap than would be theoretically anticipated in the warm-dominant quadrant, suggesting a potential structural inconsistency in its circumplex representation [131]; licensed [137]	Self-report bias; generic vs specific design trade-off (designed for international applicability rather than optimised for specific contexts, which may sacrifice precision); development was guided by existing RUM instruments only to a limited extent; broader range of experts in the instrument development process was needed due to the multi-sectoral nature of the instrument; some categories might be too broad to capture country-specific service nuances as tool is for international use [117]	Trained interviewer can best tease out accurate and comprehensive information [126]	Exhibits significant diurnal variation, with levels peaking in the morning and gradually declining throughout the day, necessitating precise timing of sample collection [138]; reliability may be affected by recent consumption of food, drink, or alcohol [139]; selective IgA Deficiency, a condition where individuals lack IgA, can naturally affect test results [140]; highly susceptible to numerous confounding biological and lifestyle factors (e.g., recent infection, antibiotic intake, ingestion of caffeine, food or drink, daily rhythm) [141]

Parameters:	Mood-related parameters	Psychological, psychiatric, physiological and cognitive parameters	Cultural, spiritual, and locus of control parameters	Music-, therapist-, and treatment-related parameters	Socioeconomic and social parameters			Biomedical parameters
Measuring instruments:	**BDI [11, 12, 20, 21, 49]**	**SCL-90-R [21]**	**DRIE [12, 20, 21]**	**HAQ(-II) [12, 20, 21]**	**IIP-SC (IIP-32) [20, 49]**	**PECUNIA-RUM**	**CSRI**	**Salivary Immunoglobulin A Test [20, 49]**
Application characteristics								
Number of items in each instrument	21 [11]	90 [145]	25	19	32 [235, 239]	Variable (modular: health and social care, education, (criminal) justice, employment and productivity, and costs borne by patients or families incl. informal care; users pick relevant modules) [117]	Variable (modular: 5 sections with 30+ main question items)	Not applicable
Total score range	0-63 [11]	0-360	0-25	19-114	0-128	No total score range (measures resource-use quantities, e.g., number of visits, hours of care, days of absence) [117]	No total score range (records service utilisation and cost data)	Normal range of S-IgA (salivary) is 4-37 ml/dL [142]
Interpretation of score	Severe [11]	Severe	More external locus of control (i.e., believing that external factors, rather than personal control, are more influential in substance use)	Stronger therapeutic alliance between the client and therapist	Greater interpersonal problems	Not applicable (no scoring instrument) [117]	Not applicable (no scoring instrument)	Higher levels of salivary Immunoglobulins [142]
Minimum clinically important difference	SUD: none Depression: 17.5% reduction[59] [143]	SUD: none Other populations: none	SUD: none Other populations: none	SUD: none Other populations: none	SUD: none Other populations: none	Not applicable (it is a cost-tracking tool, not a clinical outcome measure)	Not applicable	None/Not applicable
Testing time (min)	10 [240]	12-15 [227]	5-10	5-10	5-10	15-30 (depending on the number of modules selected)	20 [126] (interviews generally take longer than self-reports)	5-10

[59] For "individuals with longer duration depression who had not responded to antidepressants": 32% reduction [143]

Parameters:	Mood-related parameters	Psychological, psychiatric, physiological and cognitive parameters	Cultural, spiritual, and locus of control parameters	Music-, therapist-, and treatment-related parameters	Socioeconomic and social parameters			Biomedical parameters
Measuring instruments:	**BDI [11, 12, 20, 21, 49]**	**SCL-90-R [21]**	**DRIE [12, 20, 21]**	**HAQ(-II) [12, 20, 21]**	**IIP-SC (IIP-32) [20, 49]**	**PECUNIA-RUM**	**CSRI**	**Salivary Immunoglobulin A Test [20, 49]**
Instrument user (e.g. self-assessment by service user, clinician-, performance reported)	Self-reported [11]	Self-reported [145]	Self-reported	Self-report (client and therapist complete it independently)	Self-reported [235]	Self-reported [117]	Interviewer-administered	Laboratory test (by clinician)
Costs and regulations for use	Copyrighted and licensed [144], free to download (BDI, not BDI-II)	Licensed [145], free to download	Free to download	Free to download	Licensed [137] and free to download	Free to use for non-commercial research, but registration and adherence to the PECUNIA methodology are required (Open Access)	Free to use	Not applicable

Note: If no references are stated, information is obtained from the downloaded tool itself.

Abbreviations: BDI … Beck Depression Inventory. CSRI … Client service receipt inventory. DRIE … Drinking-Related Internal-External Locus of Control Scale. FU … follow-up. HAQ(-II) … (Revised) Helping Alliance Questionnaire for therapist and client. IIP-SC … Inventory of Interpersonal Problems: Short Circumplex form. NR … not reported. PECUNIA-RUM … ProgrammE in Costing, resource use measurement and outcome evaluatioN for Use in multi-sectoral National and International health economic evaluations – Resource Use Measurement Instrument. SCL-90-R … Symptom Checklist 90-R. S-IgA … Sectretory Immunglobulin A. SIgAD … selective Immunglobulin A Deficiency.

Literature search strategies

Example: Search strategy for Medline via Ovid

Search Name: Ovid MEDLINE(R) ALL <1946 to February 12, 2025>	
Search date: 13.02.2025	
ID	Search
1	exp Substance-Related Disorders/ (324922)
2	(exp Amphetamines/ or exp Cannabis/ or exp Cocaine/ or exp "Designer Drugs"/ or exp Heroin/ or exp Methamphetamine/ or exp Narcotics/ or exp "Street Drugs"/ or (amphetamine* or drug* or poly?drug* or poly-drug* or substance* or cannabi* or cocaine or hash oil* or hashish or heroin or lsd or mari#uana or methadone or mdma or morphine or ecstasy or methamphetamine* or narcotic* or opioid* or opiate* or opium).mp.) adj3 ("use" or abuse* or addict* or dependen* or misuse).mp. (755089)
3	exp Alcohol Drinking/ (81446)
4	(alcohol* adj3 (abus* or addict* or consum* or dependen* or disorder* or drink* or excess* or misus* or problem*)).mp. (182965)
5	1 or 2 or 3 or 4 (1021671)
6	exp Music Therapy/ (4778)
7	exp Music/tu [Therapeutic Use] (57)
8	(music or sing or singing or song* or choral* or choir* or melod* or lyric*).mp. (54673)
9	AMG.ti,ab. (2137)
10	MLG.ti,ab. (985)
11	6 or 7 or 8 or 9 or 10 (57793)
12	5 and 11 (1350)
13	limit 12 to (meta analysis or "review" or "systematic review") (160)
14	(((comprehensive* or integrative or systematic*) adj3 (bibliographic* or review* or literature)) or (meta-analy* or metaanaly* or "research synthesis" or ((information or data) adj3 synthesis) or (data adj2 extract*))).ti,ab. or (cinahl or (cochrane adj3 trial*) or embase or medline or psyclit or (psycinfo not "psycinfo database") or pubmed or scopus or "sociological abstracts" or "web of science").ab. or ("cochrane database of systematic reviews" or evidence report technology assessment or evidence report technology assessment summary).jn. or Evidence Report: Technology Assessment*.jn. or ((review adj5 (rationale or evidence)).ti,ab. and review.pt.) or meta-analysis as topic/ or Meta-Analysis.pt. (843361)
15	12 and 14 (64)
16	13 or 15 (172)
Total hits: 172	

HTA Austria
Austrian Institute for
Health Technology Assessment
GmbH