

Psychotherapy Works: Interpreting Response, Recovery, and Lasting Benefit

Christian Jonathan Haverkamp

Narrative Clinical Review

Abstract—Psychotherapy works across a range of mental disorders, yet the phrase is less simple than it sounds. A positive average effect does not say how many patients respond, remit, regain functioning, deteriorate, or remain well after treatment. This narrative clinical review examines evidence available through 30 November 2025 and asks what is justified when clinicians tell patients that psychotherapy is effective. Current umbrella, cross-disorder, and treatment-family syntheses support meaningful benefit while showing that estimates depend on diagnosis, comparator, outcome definition, follow-up, and study quality. Absolute response is often modest even when relative benefit is clear. Waitlist comparisons tend to produce larger contrasts than care-as-usual or active controls, and symptom scales do not exhaust the changes clients value. Evidence for sustained benefit is encouraging in several areas, including relapse prevention in depression, but long-term data remain thinner than acute data. Nonresponse, premature ending, and adverse events belong inside any balanced account of effectiveness. The practical conclusion is neither promotional nor skeptical: psychotherapy is effective enough to recommend and imperfect enough to monitor. A defensible claim specifies the patient group, treatment, comparator, desired outcome, time horizon, absolute probability, and plan if expected change does not occur.

Index Terms—psychotherapy effectiveness, treatment outcome, response and remission, recovery, long-term benefit, adverse events

I. INTRODUCTION

Psychotherapy works. That sentence is supported by a large body of research, but it becomes clinically useful only after several missing clauses are supplied. Works for which problem, delivered in what form, compared with what alternative, judged by which outcome, and assessed when? The patient who asks whether therapy will help is rarely

asking for a standardized mean difference. The underlying questions are closer to ordinary life: Will I sleep? Will I stop avoiding the train? Can I return to work, stay in a relationship without fear, or recover enough to make decisions again?

The evidence answers some of these questions with confidence and others only indirectly. Across common adult mental disorders, meta-analytic reviews generally find that bona fide psychotherapies improve symptoms more than control conditions. A field-wide umbrella review and later cross-disorder analyses also show why this conclusion cannot be converted into a guarantee. Effects vary by disorder, treatment family, comparator, and study quality; many treated patients do not meet conventional response thresholds; and the literature is much denser for acute symptoms than for participation, quality of life, relapse, or years of follow-up (Leichsenring et al., 2022; Cuijpers, Miguel, Ciharova, et al., 2024; Cuijpers et al., 2025).

A mature account can hold two facts without weakening either. Psychotherapy has real therapeutic effects, and its present effects leave substantial room for improvement. Overstatement is not needed to defend treatment, while relentless qualification can obscure a result that matters: on average, people offered an evidence-supported psychotherapy do better than comparable people in many control conditions. The task of this review is to make that result more exact and more usable.

The question is sometimes framed as whether improvement would have happened anyway. Symptoms can fluctuate, crises can resolve, relationships and circumstances can change, and patients often seek help near a period of unusual distress. Randomized controls estimate these influences imperfectly but far better than clinical memory. Therapists mainly remember the people they treated, not the comparable people who recovered, persisted, or worsened without that encounter. Effectiveness research supplies the missing comparison and should temper both therapeutic pride and therapeutic nihilism.

Christian Jonathan Haverkamp, M.D., LL.M., is a psychotherapist working in private practice in Dublin, Ireland. Correspondence: jonathanhaverkamp@gmail.com. Website: www.jonathanhaverkamp.com.
© 2025 Christian Jonathan Haverkamp.

II. SCOPE AND REVIEW METHOD

This narrative clinical review used a literature cutoff of 30 November 2025, exactly one calendar month before the nominal issue date. Searches emphasized umbrella reviews, systematic reviews, meta-analyses, participant-level analyses, randomized and pragmatic trials, routine-care evidence, and official implementation guidance. Complete online-first publications available by the cutoff were eligible. Newer work was given weight when it changed interpretation rather than merely repeating an established result.

The review concentrates on adult psychotherapy for depressive, anxiety-related, trauma-related, obsessive-compulsive, eating, psychotic, bipolar, somatic symptom, and personality presentations where cross-cutting evidence was available. It does not attempt to rank every named therapy or cover child, couple, family, group, or rehabilitation literatures comprehensively. Most numerical examples come from disorders with large trial databases. They should not be silently generalized to populations that were excluded, underrepresented, or treated for different outcomes.

No new meta-analysis was performed. Trial overlap is therefore important: an umbrella review, a network meta-analysis, and a later unified series may summarize some of the same studies. They are used to answer different questions, not counted as independent replications. Interpretation focused on population, treatment, comparator, outcome, follow-up, missing data, bias, and the difference between relative and absolute change. The aim is a disciplined reading of effectiveness evidence rather than a vote for or against psychotherapy.

Publication date and search date were kept separate. A review published in 2025 may contain trials only through 2022, while an older review may still provide the most direct analysis of a particular outcome. The cutoff identifies what evidence was publicly available; it does not make every eligible source equally current. Where a recent synthesis depended on an older evidence base, that lag is treated as part of the uncertainty rather than hidden by the year in the citation.

III. WHAT DOES IT MEAN TO SAY THAT PSYCHOTHERAPY WORKS?

The verb “works” can refer to at least four claims. A therapy may have efficacy under controlled trial conditions, effectiveness in ordinary care, efficiency relative to its cost and burden, or public-health impact after access and uptake are considered. These are related achievements, not synonyms. A treatment may be efficacious but reach very few people; it may be feasible in a national service while producing uneven results across teams; it may reduce symptoms without changing the outcome a particular patient most wants to change.

Effectiveness also depends on the unit being described. A group mean concerns a population. Response and remission concern proportions defined by thresholds. Reliable change asks whether movement exceeds likely measurement error. Deterioration concerns meaningful worsening. None of these alone is an individual prognosis. A patient can improve substantially without crossing a remission boundary, cross a scale threshold while remaining unable to work, or regain a valued role with a smaller symptom change than the trial designated as response.

It is therefore useful to treat “psychotherapy works” as the opening line of an evidence summary. The next line states the disorder or problem, the type and format of therapy, the comparison, the outcome, the interval, and the expected absolute probability. The final line addresses uncertainty and alternatives. This grammar makes consent more honest without turning a clinical conversation into a statistics seminar.

Absolute probability also depends on baseline risk. The same relative effect can yield a large absolute gain when spontaneous or usual-care response is common and a much smaller gain when it is rare. Number needed to treat can help communicate this difference, but it inherits every decision about outcome threshold, follow-up, and comparator. It is an estimate for a population, not the number of patients a particular clinician must treat before one named person benefits.

Completing the sentence “psychotherapy works”

- Name the population and the treatment actually studied, including format and intensity.
- State the comparator: waitlist, no treatment, care as usual, medication, or another credible therapy.
- Distinguish average symptom change from response, remission, reliable change, functioning, and deterioration.
- Give the time horizon and say whether treatment continued during follow-up.
- Use absolute probabilities where available; a relative advantage does not show how many people recovered.
- Include a plan for nonresponse, worsening, practical barriers, or a goal that the usual scale does not capture.

IV. WHAT RANDOMIZED COMPARISONS ESTABLISH

Random allocation is powerful because, with sufficient numbers and appropriate analysis, it reduces systematic differences between groups before treatment. If outcomes then diverge, the intervention is a more plausible cause than in an uncontrolled before-and-after series. Psychotherapy trials have used this design across diagnoses, modalities, delivery formats, and settings. Their accumulation is the main reason the statement that psychotherapy is effective does not rest on testimonials or spontaneous recovery.

Randomization does not make every trial equally informative. Blinding patients and therapists to psychotherapy is usually impossible. Expectations, therapist allegiance, treatment fidelity, selective publication, missing outcomes, small samples, and flexible analyses can alter estimates. Eligibility criteria may remove the instability, comorbidity, risk, or practical difficulty common in routine practice. A trial can be internally persuasive yet apply only narrowly, or broadly representative yet leave more uncertainty about what caused the observed change.

The question asked by the control group is especially consequential. A waitlist estimates the added benefit of receiving therapy now rather than later under the study arrangement. Care as usual estimates benefit beyond whatever the local system already provides. An active psychological control may isolate particular procedures from attention, structure, and expectation. A head-to-head trial compares packages, but a small or imprecise difference does not prove that they are universally equivalent. Each design supports a bounded conclusion.

Analysis after randomization also matters. Intention-to-treat methods preserve the original allocation and reduce the bias that would arise if only completers were compared. They cannot fully solve outcome data that are missing because a patient deteriorated, disengaged, disliked the treatment, or sought other care. Sensitivity analyses make assumptions explicit, but no statistical method can reconstruct an unobserved clinical course without uncertainty. Completion-only results should therefore be read as a selected view of treatment exposure.

Certainty also depends on whether a result survives less favorable assumptions. Effects often become smaller when analyses are restricted to larger or lower-risk-of-bias trials, when passive controls are replaced by active ones, or when publication bias is considered. A smaller credible effect can be more useful than a dramatic unstable one. It provides a sound basis for service planning and shared decisions without requiring every methodological correction to leave the original headline untouched.

Table 1. Different evidence designs answer different psychotherapy questions

Design	Primary question	Strength	Interpretive limit
Therapy versus waitlist	Does receiving therapy now outperform delayed access?	Clear allocation and simple contrast	May exaggerate the contrast with ordinary active care
Therapy versus care as usual	Does therapy add benefit in this service context?	Clinically relevant incremental effect	Usual care differs greatly across sites and periods
Therapy versus active control	Do specified procedures add to attention, structure, or expectancy?	Stronger test of specific benefit	Control credibility and therapeutic content are difficult to match

Design	Primary question	Strength	Interpretive limit
Head-to-head psychotherapy	Does one credible package outperform another?	Direct comparative information	Often underpowered for modest differences or subgroups
Pragmatic or routine-care study	Does benefit survive ordinary delivery?	Reach, retention, and service variation become visible	Causal inference may be weaker and missingness informative

V. THE COMPARATOR CHANGES THE ANSWER

The influence of the comparator is not a technical footnote. In depression trials, waiting-list groups often improve less than no-treatment, care-as-usual, or psychological-placebo groups. A network meta-analysis found that response differed among control conditions, and a later analysis of 333 trials reported a larger average psychotherapy contrast against waitlist than against care as usual (Furukawa et al., 2014; Cuijpers, Miguel, Harrer, et al., 2024). The treatment groups did not appear dramatically different; much of the contrast arose because improvement in the waitlist groups was smaller.

Several processes may contribute. Being told to wait can affect expectation, help-seeking, or use of other care. Care as usual can include medication, supportive contact, monitoring, or spontaneous changes that a simple label conceals. Active controls can themselves be therapeutic. It follows that an effect size is not a fixed property stored inside a treatment. It describes a difference produced in a particular comparison.

This does not make waitlist trials useless. They show that timely access to the tested package is better than the delayed-access arrangement used in the trial. The error begins when this contrast is presented as the effect beyond competent active care or as the difference between psychotherapy and doing nothing in ordinary life. Reviews should keep control conditions separate where possible, and clinicians should be wary of a single pooled figure assembled from unlike comparisons.

Care as usual is not a stable treatment. In one service it may be regular primary-care review and medication; in another it may amount to a name on a waiting list. Its intensity can change during a long recruitment period or differ by country. A modest incremental effect over well-organized usual care may represent a demanding test, whereas a large effect over sparse usual care may say more about service absence. Reports that leave usual care undescribed make clinical translation needlessly uncertain.

VI. EVIDENCE ACROSS DISORDERS AND THERAPY FAMILIES

Cross-disorder reviews support psychotherapy across a broad clinical range, although the evidence is not evenly distributed. An overview of meta-analyses found benefits of

psychological and pharmacological treatments across adult psychiatric disorders, with sizeable variation between diagnoses and literatures (Huhn et al., 2014). One later umbrella synthesis drew on 102 meta-analyses encompassing 3,782 randomized trials and over 650,000 participants. Across heterogeneous disorders, estimated benefits over placebo or treatment as usual were small on average, and risk of bias remained consequential (Leichsenring et al., 2022).

The 2024 cross-disorder analysis translated benefit into response rates for eight disorders. Psychotherapy response estimates ranged from about one quarter to just over two fifths, depending on diagnosis, while relative response was significantly better than control for most conditions examined. The absence of a significant relative signal for borderline personality disorder in that particular analysis should not be read as a verdict on specialist treatments, because treatment goals, duration, control conditions, and definitions did not align neatly with the other disorders. The broader lesson is that a field average does not replace diagnosis-specific evidence (Cuijpers, Miguel, Ciharova, et al., 2024).

Cognitive behavioral therapy has the largest cross-disorder trial base. In a unified 2025 series, evidence was judged probable for major depression and several anxiety-related conditions, including post-traumatic stress and obsessive-compulsive disorders, as well as for eating disorders; findings for psychotic and bipolar disorders were classified as possible. The authors also documented common limitations: risk of bias, heterogeneity, possible publication bias, and dropout. The label CBT covered related but not identical interventions, so the result supports a family of treatments rather than a generic script (Cuijpers et al., 2025).

Evidence is not confined to CBT. Reviews of adult depression found several therapies better than controls and relatively few robust differences among credible treatments after study characteristics were considered (Barth et al., 2013; Cuijpers et al., 2020). An updated umbrella review judged psychodynamic psychotherapy empirically supported for several common disorders, while retaining the same methodological cautions that apply elsewhere (Leichsenring et al., 2023). Therapeutic pluralism is defensible when it is tied to evidence and formulation; it is not a claim that every therapy is suitable for every target.

Heterogeneity is not only statistical noise. It can reflect real differences in severity, comorbidity, recruitment, therapist training, culture, delivery, and the outcome selected. A pooled estimate remains useful when those studies address a coherent question, but it should prompt examination of the distribution rather than a search for one universal value. Subgroup findings can generate hypotheses, yet small and repeated moderator analyses are especially vulnerable to chance. Variation should invite better prediction, not retrospective certainty.

VII. FROM AVERAGE EFFECTS TO RESPONSE AND RECOVERY

Standardized mean differences are useful when trials use different scales, but they do not tell a patient the probability of a recognizable change. A modest mean difference can arise because many people improve a little, a smaller group improves greatly, or outcomes vary widely. Response and remission make interpretation more concrete, yet their cutoffs are conventional. A 50% symptom reduction can leave severe residual symptoms when baseline severity was high; a remission score can be crossed by one point without restoring social or occupational functioning.

For adult depression, a 2021 meta-analysis estimated short-term response after psychotherapy at about 41%, compared with roughly 16–17% in waitlist or care-as-usual groups, and remission in about one third of treated participants. Reliable improvement and deterioration supplied additional information that mean scores would have hidden (Cuijpers et al., 2021b). The 2024 analysis extended the absolute-response perspective across eight disorders and again found a consistent combination: meaningful relative advantage alongside modest absolute response (Cuijpers, Miguel, Ciharova, et al., 2024).

This is not a pessimistic reinterpretation. If treatment approximately doubles a patient's chance of response in a relevant comparison, the benefit is important even though many people still need further care. The figures also protect against moral injury. A patient who does not respond is not an exception that invalidates the evidence, nor evidence that they failed to try. Nonresponse is present inside the trial result and should be anticipated in the treatment pathway.

Recovery is broader still. It may include sustained remission, restored roles, less avoidance, safer behavior, more choice, renewed trust, or the ability to manage recurrence. Some of these outcomes are difficult to compare across diagnoses and are measured inconsistently. They belong in evaluation nonetheless, because symptom reduction is a means to living and sometimes only one part of the change a patient seeks. For example, an earlier JPPC review treated credible relational repair as recognition, responsibility, responsiveness, and observable amendment—changes that a symptom total alone cannot capture (Haverkamp, 2020).

The denominator should remain visible throughout. Response among people who supplied a post-treatment measure is not the same as response among everyone randomized, referred, or offered care. Services can improve apparent results by excluding nonstarters, transfers, or missing outcomes. Trials can also differ in whether missing participants are assumed not to have responded. Whenever possible, the reader should ask who entered the denominator, who disappeared from it, and whether disappearance plausibly relates to outcome.

Functioning can lag behind symptoms. Avoidance may have narrowed work, friendships, physical activity, or confidence over years; symptom relief does not instantly rebuild those structures. The reverse can also occur: a person may resume important activity while distress remains elevated. Measuring both trajectories helps distinguish an intervention that reduces suffering from one that also restores participation, and it identifies when social, occupational, physical, or rehabilitative support is needed alongside psychotherapy.

Table 2. Outcome terms that should not be treated as interchangeable

Outcome	What it describes	Clinical value	Common mistake
Mean symptom change	Average movement on a continuous scale	Sensitive to change and useful across groups	Converting the average directly into an individual prognosis
Response	A prespecified proportional or categorical improvement	Approximates noticeable benefit	Assuming response means wellness or restored function
Remission	A score below a symptom threshold	Signals low symptom burden at that point	Treating a threshold crossing as durable recovery
Reliable change	Change unlikely to reflect measurement error	Separates likely signal from scale noise	Ignoring whether the final state remains clinically severe
Deterioration	Meaningful worsening over the interval	Makes harm and failed care visible	Recording only adverse events judged attributable to therapy
Functioning or participation	Change in roles, relationships, and daily activity	Connects treatment to ordinary life	Assuming symptom change automatically produces it

VIII. OUTCOMES PATIENTS RECOGNIZE

Most trials begin with researcher-selected primary outcomes, usually symptoms. Patients often describe a wider field of change. A qualitative meta-analysis of 177 studies across 24 countries organized accounts from 2,908 psychotherapy clients into domains including emotional and social functioning, self-understanding, agency, behavior, self-attitude, engagement with life, wellbeing, and symptom change (Ladmanová et al., 2025). These reports do not estimate causal effect, but they expose outcomes that standardized trials may neglect.

Patient-valued outcomes should not replace safety or diagnostic measurement. Someone can feel more self-aware while continuing a dangerous pattern, and a therapy can improve symptoms while creating dependence, conflict, or practical burden. A useful outcome set holds several perspectives together: the patient's priorities, symptom and functioning measures, risk, adverse experiences, and

observations from relevant others when consent and context make those observations appropriate.

Goals also change during therapy. A person may begin by wanting anxiety removed and later choose to tolerate some anxiety in order to travel, speak, grieve, or set a boundary. That shift is not automatically avoidance of the original outcome; it may reflect more accurate learning about what recovery entails. It should be discussed rather than inferred after the fact. Measurement is most informative when numbers remain connected to the life they were meant to describe.

Satisfaction is relevant but not a substitute for benefit. A patient may value being understood during a period when symptoms remain severe, or dislike an effective intervention that asks for difficult exposure and practice. Conversely, rapid symptom relief can coexist with feeling unheard or coerced. Outcome, experience, and process overlap without collapsing into one measure. Collecting them separately makes it possible to improve care without forcing every judgment into a single favorable score.

IX. DOES BENEFIT LAST?

Durability is one of psychotherapy's most appealing promises and one of the easiest to state too broadly. Follow-up studies are less numerous than acute trials, attrition accumulates, treatments may continue, and control participants may later receive psychotherapy. A difference observed a year after randomization can reflect skills learned during acute treatment, additional care, differential relapse, or all three. The time origin and subsequent treatment exposure matter.

A network meta-analysis of major psychotherapies for adult depression found short-term efficacy and acceptability across several treatment types, while long-term comparisons were fewer and less certain (Cuijpers et al., 2021a). This pattern is common: acute evidence is stronger than evidence distinguishing treatments years later. Lack of a precise long-term comparison does not show that benefit vanished, but neither does an acute effect establish enduring protection.

Relapse-prevention evidence is more direct in selected populations. A 2025 individual participant data meta-analysis of people in partial remission from depression found that psychological interventions reduced relapse risk over twelve months and reduced residual symptoms, though quality-of-life effects were uncertain and no convincing moderators of relapse prevention emerged (Gülpen et al., 2025). Another participant-level synthesis found psychological interventions helpful for subthreshold depression through twelve months, with effects less certain by two years (Harrer et al., 2025). These are encouraging, bounded findings rather than a general claim that psychotherapy immunizes against recurrence.

Long-term benefit may consist of more than maintaining a low score. A patient may notice recurrence earlier, use a previously learned response, seek help sooner, or prevent one episode from becoming a prolonged crisis. Trials do not always capture such capability. Clinically, durability is best treated as something to support: identify warning signs, rehearse future use, plan booster or renewed care, and make re-entry straightforward rather than presenting discharge as proof that treatment will never again be needed.

Follow-up timing deserves precision. “Long term” has referred to six months in some literatures and several years in others. An assessment six months after randomization may occur only weeks after a lengthy treatment ended, while twelve months after treatment completion is a different interval. Relapse, recurrence, and sustained response also require definitions. A summary that combines them can sound firmer than the underlying studies allow, particularly when few participants remain at the latest time point.

What happened during follow-up is equally important. Participants may continue the original therapy, receive booster sessions, begin medication, enter another psychotherapy, or experience major changes in health and circumstances. Preventing these events would make follow-up artificial, but ignoring them can attribute all later difference to the acute treatment. Trials and reviews should report subsequent care, while clinicians should understand durable outcome as the course following an initial strategy rather than the isolated afterlife of one package.

X. NONRESPONSE, PREMATURE ENDING, AND HARM

Because average benefit is real, clinicians can be tempted to interpret stalled treatment as an unusual patient problem. The absolute-response literature shows otherwise. Many participants in successful trials do not reach response or remission. The reasons are heterogeneous: the formulation may be wrong, the method poorly matched or delivered, the dose inadequate, attendance interrupted, the environment unsafe, another condition untreated, or the chosen scale insensitive to the relevant change. Sometimes a well-delivered treatment simply does not help enough.

Premature ending is equally difficult to interpret. A meta-analysis found substantial dropout across adult psychotherapy, but definitions varied and often conflated disengagement, practical loss, transfer, dissatisfaction, and mutually agreed early completion (Swift & Greenberg, 2012). Completion rates should not be pursued as a quality indicator without context. Retaining someone in an ineffective or unwanted therapy is not success, while losing contact during escalating risk is not an ordinary preference outcome.

Psychotherapy can be associated with distress, symptom worsening, dependency, relationship strain, recovered or

newly salient memories, shame, or other adverse events. A systematic review of 85 randomized trials found explicit harm reporting in only 60%, with heterogeneous definitions and uncertain attribution; adverse events were not rare, although occurrence during a trial did not establish that therapy caused each event (Klatte et al., 2025). Deterioration rates can be reassuring in comparative analyses and still leave a duty to monitor individual harm.

A balanced service records symptom deterioration, serious events, unwanted effects, dropout circumstances, and barriers to continued care. It also creates a response: reassess diagnosis and risk, ask directly about the therapeutic relationship, change pace or method, coordinate medical or social intervention, seek supervision, transfer when appropriate, and document why. The ethical answer to imperfect effectiveness is not defensiveness but a route to revision.

Early change can inform that route without becoming a rigid stopping rule. Some problems respond quickly, others require preparatory work, and apparent early gains may not last. Still, an expected trajectory belongs in the formulation. If the proposed maintaining process has not shifted after it has been targeted competently, the absence of movement is evidence. The clinician should explain why continuing unchanged is reasonable or alter the plan; hope alone is not a treatment rationale.

Table 3. Signals that the statement “therapy is working” needs review

Signal	Possible explanations	Immediate question	Reasonable response
No meaningful early change	Wrong target, low intensity, poor fit, interrupted use	What should have changed by now if the formulation were right?	Clarify target, assess implementation, and set a time-limited test
Worsening symptoms or risk	Expected temporary distress, adverse effect, illness progression, external crisis	What changed, when, and how is safety affected?	Reassess risk and causation; change or stop treatment when indicated
Repeated missed sessions	Practical barrier, avoidance, rupture, low credibility, competing need	What makes returning difficult or not worthwhile?	Address access and relationship; avoid moralizing attendance
Improved scale, unchanged life	Measure mismatch, environmental constraint, residual avoidance	Which valued activity remains impossible?	Add a functional goal and test change outside sessions
Supportive relationship, static treatment	Comfort without target movement, dependency, unclear method	What work is the relationship making possible?	Renegotiate aims and tasks; consult or transfer if needed

XI. FROM TRIALS TO ORDINARY SERVICES

Efficacy can be lost before the first session. Referral thresholds, waiting times, fees, transport, childcare, work, language, disability access, stigma, and mistrust shape who begins. A service that reports outcomes only among people

who completed treatment excludes several clinically important groups from its denominator. Public benefit depends on need, referral, uptake, attendance, treatment quality, and outcome together.

Large-scale services show that psychotherapy can be organized beyond specialist research clinics. The English Improving Access to Psychological Therapies programme combined evidence-based protocols, workforce training, routine outcome measurement, and public reporting at national scale (Clark, 2018). Its achievement does not erase variation in case mix, missing data, local capacity, therapist outcomes, or recovery definitions. Routine measurement makes these questions visible; it does not answer them automatically.

Digital and lower-intensity formats can extend reach. Routine-service evaluations of internet-delivered CBT for depression and anxiety generally favored treatment, but guidance, adherence, and setting varied across studies (Etzelmueller et al., 2020). A randomized trial in Indian primary care showed that a brief behavioral treatment delivered by trained lay counsellors benefited adults with severe depression when embedded in supervision and referral arrangements (Patel et al., 2017). Format and workforce are therefore modifiable parts of care, not proxies for lesser seriousness.

Implementation requires more than distributing a manual. The World Health Organization's implementation guidance asks services to prepare the workforce, adapt to context, identify and support intended beneficiaries, and monitor delivery (World Health Organization, 2024). These conditions belong to effectiveness because they determine whether the intervention encountered in practice resembles the one supported by evidence. A technically excellent treatment that is inaccessible, culturally unintelligible, or unsupported after risk emerges has not achieved its public purpose.

Equity cannot be inferred from a favorable average. A service may produce good outcomes for those who enter while language, location, digital access, disability, immigration status, or previous discrimination keeps others outside. Reporting uptake and outcome by relevant groups can reveal where benefit is being lost, but small numbers and crude categories require care. Adaptation should preserve the treatment's plausible active tasks while changing how those tasks are explained, delivered, and supported in context.

XII. MEDICATION, COMBINED CARE, AND PREFERENCE

Psychotherapy and medication are often presented as competing philosophies. Direct comparisons in depression and anxiety have generally found no large overall efficacy difference, with variation by diagnosis, study quality, and treatment (Cuijpers et al., 2013). Broader overviews likewise support both modalities across disorders but do not justify a

universal ranking because their trials, outcomes, adverse effects, and time horizons differ (Huhn et al., 2014). The clinically relevant question is which option, sequence, or combination fits the present situation.

Combined treatment can improve outcomes in some depressive and anxiety presentations. Meta-analysis has found benefit when psychotherapy is added to antidepressant medication and, in examined comparisons, when combined care is compared with psychotherapy alone (Cuijpers et al., 2014). Additional treatment also adds appointments, cost, coordination, and potential conflict between explanatory models. Combination should have an explicit rationale rather than functioning as accumulation: what is each component expected to change, and how will benefit or burden be recognized?

Preference has modest but meaningful relevance. An updated meta-analysis found small favorable effects of receiving a preferred treatment or active choice on satisfaction, completion, and clinical outcome (Bennett et al., 2025). Preference cannot decide every safety or evidence question, and it may change after a credible explanation or previous nonresponse. It is still data. Shared choice can improve fit and makes it more likely that the patient understands what the treatment asks of them.

Burden belongs beside efficacy in that choice. Weekly travel, exposure-related distress, practice between sessions, medication adverse effects, privacy, and time away from work are not extraneous inconveniences. They alter whether a treatment can be used. A theoretically optimal option that the patient cannot sustain may have less real benefit than a somewhat less intensive option that can be completed, monitored, and stepped up if needed. This is a pragmatic judgment, not permission to offer inadequate care.

Sequence can preserve choice when the first decision is uncertain. A patient may begin with psychotherapy because adverse effects or pregnancy make medication less attractive, add medication when severity prevents useful engagement, or begin combined care when rapid stabilization and longer-term psychological work are both priorities. The evidence sets reasonable starting points; response, burden, and preference then supply new information. Treatment choice is better understood as a revisable pathway than a once-only declaration of allegiance.

XIII. MEASUREMENT, FEEDBACK, AND TIMELY REVISION

Routine outcome measurement is sometimes criticized as reductive and sometimes promoted as though the act of measurement were treatment. Both positions miss the practical issue. A scale can reveal unexpected deterioration, discrepant views, or lack of change that an unstructured conversation overlooks. It can also narrow attention, be completed

mechanically, or misrepresent a culturally unfamiliar experience. Its value depends on what follows.

A multilevel meta-analysis found that progress feedback produced small average symptom benefits and small favorable effects on dropout, with particular relevance when patients were not on track (de Jong et al., 2021). A small average effect is appropriate to the intervention: feedback does not supply a new treatment by itself. It creates an opportunity to ask why progress differs from expectation and whether the clinician and patient will act on the answer.

Therapist variation is another reason to monitor. Routine-practice analyses have found meaningful differences in outcomes between therapists, including interaction with patient severity and risk (Saxon & Barkham, 2012). Such findings should not be used for crude league tables without case-mix adjustment and adequate numbers. They do support supervision that examines outcomes, treatment process, and difficult cases rather than assuming that a therapy's evidence base transfers automatically through every clinician.

A review point should end with a decision. Continue because the current trajectory is credible; clarify the goal or measure; address access; repair a misunderstanding; increase intensity; change method; add another treatment; seek consultation; or reconsider the formulation. Measurement without a decision becomes documentation. The purpose is not to make care algorithmic but to make inertia harder.

XIV. BOUNDARIES OF THE EVIDENCE

The strongest psychotherapy literatures concern selected diagnoses, outcomes, and settings. Evidence is thinner for complex multimorbidity, unstable housing, coercive environments, significant cognitive impairment, acute medical illness, or people excluded because of risk. A treatment supported for a disorder may remain reasonable in a more complex case, but the certainty is lower and the surrounding care becomes more important. Guideline status should not conceal the distance between the study population and the person being treated.

Average comparisons also tell us little about reliable individual selection. Prognostic factors identify who tends to do better; moderators identify whether a characteristic changes the relative benefit of two treatments. Many proposed personalized-treatment rules have not replicated strongly enough for routine use. A sensible formulation can guide a choice without pretending to be a validated prediction model. Where uncertainty is material, transparent preference and early monitoring may be more honest than false precision.

The evidence base is also shaped by what is funded, manualized, measured, and published. Approaches with fewer trials are not thereby proven ineffective, but absence of evidence cannot be converted into evidence of benefit.

Conversely, a large literature can repeat the same biases. The right response is proportional confidence: make the strongest claim the design supports, identify what remains unknown, and avoid filling gaps with allegiance.

Communication should be proportional too. Saying that a therapy “usually works” may imply a probability the evidence does not support; reciting every limitation may leave a patient unable to hear that benefit is more likely with treatment than without it. A useful explanation gives a best estimate or range, names the outcome and interval, acknowledges that individual prediction is limited, and describes how progress will be reviewed. Confidence becomes practical when it includes a route forward if the estimate does not fit this person.

XV. CONCLUSION

Psychotherapy works in the scientifically ordinary and clinically important sense that evidence-supported psychological treatments improve outcomes for many people across a range of mental disorders. That conclusion survives reasonable scrutiny. It does not depend on every comparison being large, every therapy being equal, or every patient recovering. It rests on replicated randomized evidence, converging syntheses, and successful delivery in settings beyond the original trials.

The conclusion becomes more credible when its limits are included. Absolute response is modest, control conditions change effect estimates, follow-up is thinner than acute evidence, and outcomes valued by patients extend beyond symptoms. Nonresponse, dropout, deterioration, and adverse events are not embarrassments outside the evidence; they are part of it. A professional account of benefit should make room for them before treatment begins.

The practical message is straightforward. Offer psychotherapy when the evidence, formulation, and patient's circumstances support it. Explain the likely benefit in terms the patient can use. Measure symptoms, functioning, risk, and the goal that brought the person to care. Review progress early enough to change course. Psychotherapy is effective enough to warrant confidence and incomplete enough to require attention.

XVI. CONFLICTS OF INTEREST AND DISCLOSURES

Conflicts of interest—The author developed and writes about Communication-Focused Therapy (CFT). Where CFT is discussed, it is presented as a formulation perspective and not as a substitute for established assessment, evidence-based treatment, or independent replication. No other conflict of interest was reported.

Funding—No specific funding was reported for this article.

Author contributions—Christian Jonathan Haverkamp is the sole named author and is responsible for the conception,

literature review, manuscript preparation, and approval of the final draft.

Patient material—No identifiable patient material is presented. Any brief examples are generic composites created to illustrate clinical reasoning and are not case reports.

XVII. REFERENCES

- Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., Jüni, P., & Cuijpers, P. (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: A network meta-analysis. *PLOS Medicine*, 10(5), e1001454. <https://doi.org/10.1371/journal.pmed.1001454>
- Bennett, C. B., Toevs, E. K., Angus, A., Krawczak, R., Trentacosta, C. J., Jagadeesh, D., Tomlinson, C. S., & Lindhiem, O. (2025). Meta-analysis on the effects of client preferences on treatment satisfaction, completion, and clinical outcome: A 10-year update. *Clinical Psychology Review*, 121, 102644. <https://doi.org/10.1016/j.cpr.2025.102644>
- Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: The IAPT program. *Annual Review of Clinical Psychology*, 14, 159–183. <https://doi.org/10.1146/annurev-clinpsy-050817-084833>
- Cuijpers, P., Harrer, M., Miguel, C., Ciharova, M., Papola, D., Basic, D., Botella, C., Cristea, I., de Ponti, N., Donker, T., Driessen, E., Franco, P., Gómez-Gómez, I., Hamblen, J., Jiménez-Orenga, N., Karyotaki, E., Keshen, A., Linardon, J., Motrico, E., Matbouriahi, M., Panagiotopoulou, O. M., Pfund, R. A., Plessen, C. Y., Riper, H., Schnurr, P. P., Sijbrandij, M., Toffolo, M. B. J., Tong, L., van Ballegooijen, W., van der Ven, E., van Straten, A., Wang, Y., & Furukawa, T. A. (2025). Cognitive behavior therapy for mental disorders in adults: A unified series of meta-analyses. *JAMA Psychiatry*, 82(6), 563–571. <https://doi.org/10.1001/jamapsychiatry.2025.0482>
- Cuijpers, P., Karyotaki, E., Ciharova, M., Miguel, C., Noma, H., & Furukawa, T. A. (2021b). The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: A meta-analysis. *Acta Psychiatrica Scandinavica*, 144(3), 288–299. <https://doi.org/10.1111/acps.13335>
- Cuijpers, P., Karyotaki, E., de Wit, L., & Ebert, D. D. (2020). The effects of fifteen evidence-supported therapies for adult depression: A meta-analytic review. *Psychotherapy Research*, 30(3), 279–293. <https://doi.org/10.1080/10503307.2019.1649732>
- Cuijpers, P., Miguel, C., Ciharova, M., Harrer, M., Basic, D., Cristea, I. A., de Ponti, N., Driessen, E., Hamblen, J., Larsen, S. E., Matbouriahi, M., Papola, D., Pauley, D., Plessen, C. Y., Pfund, R. A., Setkowski, K., Schnurr, P. P., van Ballegooijen, W., Wang, Y., Riper, H., van Straten, A., Sijbrandij, M., Furukawa, T. A., & Karyotaki, E. (2024). Absolute and relative outcomes of psychotherapies for eight mental disorders: A systematic review and meta-analysis. *World Psychiatry*, 23(2), 267–275. <https://doi.org/10.1002/wps.21203>
- Cuijpers, P., Miguel, C., Harrer, M., Ciharova, M., & Karyotaki, E. (2024). The overestimation of the effect sizes of psychotherapies for depression in waitlist controlled trials: A meta-analytic comparison with usual care controlled trials. *Epidemiology and Psychiatric Sciences*, 33, e56. <https://doi.org/10.1017/S2045796024000611>
- Cuijpers, P., Quero, S., Noma, H., Ciharova, M., Miguel, C., Karyotaki, E., Cipriani, A., Cristea, I. A., & Furukawa, T. A. (2021a). Psychotherapies for depression: A network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry*, 20(2), 283–293. <https://doi.org/10.1002/wps.20860>
- Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: A meta-analysis of direct comparisons. *World Psychiatry*, 12(2), 137–148. <https://doi.org/10.1002/wps.20038>
- Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2014). Adding psychotherapy to antidepressant medication in depression and anxiety disorders: A meta-analysis. *World Psychiatry*, 13(1), 56–67. <https://doi.org/10.1002/wps.20089>
- de Jong, K., Conijn, J. M., Gallagher, R. A. V., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, 85, 102002. <https://doi.org/10.1016/j.cpr.2021.102002>
- Etzelmueller, A., Vis, C., Karyotaki, E., Baumeister, H., Titov, N., Berking, M., Cuijpers, P., Riper, H., & Ebert, D. D. (2020). Effects of internet-based cognitive behavioral therapy in routine care for adults in treatment for depression and anxiety: Systematic review and meta-analysis. *Journal of Medical Internet Research*, 22(8), e18100. <https://doi.org/10.2196/18100>
- Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., Chen, P., Hunot, V., & Churchill, R. (2014). Waiting list may be a nocebo condition in psychotherapy trials: A contribution from network meta-analysis. *Acta Psychiatrica Scandinavica*, 130(3), 181–192. <https://doi.org/10.1111/acps.12275>
- Gülpen, J., Breedvelt, J. J. F., van Dis, E. A. M., Geurtsen, G. J., Warren, F. C., van Heeringen, C., Hitchcock, C., Holländare, F., Huijbers, M. J., Jarrett, R. B., Jermann, F., de Jonge, M., Klein, D. N., Klein, N. S., Ma, S. H., Moore, M. T., Denys, D. A. J. P., Williams, J. M. G., Kuyken, W., & Bockting, C. L. (2025). Psychological interventions for preventing relapse in individuals with partial remission of depression: A systematic review and individual participant data meta-analysis. *Psychological Medicine*, 55, e50. <https://doi.org/10.1017/S0033291725000157>
- Harrer, M., Sprenger, A. A., Illing, S., Adriaanse, M. C., Albert, S. M., Allart, E., Almeida, O. P., Basanovic, J., van Bastelaar, K. M. P., Batterham, P. J., Baumeister, H., Berger, T., Blanco, V., Bø, R., Casten, R. J., Chan, D., Christensen, H., Ciharova, M., Cook, L., et al. (2025). Psychological intervention in individuals with subthreshold depression: Individual participant data meta-analysis of treatment effects and moderators. *The British Journal of Psychiatry*. Advance online publication. <https://doi.org/10.1192/bjp.2025.56>
- Haverkamp, C. J. (2020). Demand, withdrawal, and repair: Communication patterns in close relationships. *J Psychiatry Psychotherapy Communication*, 9(3), 81–89.
- Huhn, M., Tardy, M., Spineli, L. M., Kissling, W., Förstl, H., Pitschel-Walz, G., Leucht, C., Samara, M., Dold, M., Davis, J. M., & Leucht, S. (2014). Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: A systematic overview of meta-analyses. *JAMA Psychiatry*, 71(6), 706–715. <https://doi.org/10.1001/jamapsychiatry.2014.112>

Haverkamp, C. J. (2025). Psychotherapy Works: Interpreting Response, Recovery, and Lasting Benefit. *J Psychiatry Psychotherapy Communication*, 2025 Dec 31;14(4):111-120

- Klatte, R., Strauss, B., Flückiger, C., & Rosendahl, J. (2025). Adverse events in psychotherapy randomized controlled trials: A systematic review. *Psychotherapy Research*, 35(1), 84–99. <https://doi.org/10.1080/10503307.2023.2286992>
- Ladmanová, M., Řiháček, T., Timulak, L., Jonášová, K., Kubantová, B., Mikoška, P., Polakovská, L., & Elliott, R. (2025). Client-identified outcomes of individual psychotherapy: A qualitative meta-analysis. *The Lancet Psychiatry*, 12(1), 18–31. [https://doi.org/10.1016/S2215-0366\(24\)00356-0](https://doi.org/10.1016/S2215-0366(24)00356-0)
- Leichsenring, F., Abbass, A., Heim, N., Keefe, J. R., Kisely, S., Luyten, P., Rabung, S., & Steinert, C. (2023). The status of psychodynamic psychotherapy as an empirically supported treatment for common mental disorders: An umbrella review based on updated criteria. *World Psychiatry*, 22(2), 286–304. <https://doi.org/10.1002/wps.21104>
- Leichsenring, F., Steinert, C., Rabung, S., & Ioannidis, J. P. A. (2022). The efficacy of psychotherapies and pharmacotherapies for mental disorders in adults: An umbrella review and meta-analytic evaluation of recent meta-analyses. *World Psychiatry*, 21(1), 133–145. <https://doi.org/10.1002/wps.20941>
- Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., Dimidjian, S., Araya, R., Hollon, S. D., King, M., Vijayakumar, L., Park, A.-L., McDaid, D., Wilson, T., Velleman, R., Kirkwood, B. R., & Fairburn, C. G. (2017). The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: A randomised controlled trial. *The Lancet*, 389(10065), 176–185. [https://doi.org/10.1016/S0140-6736\(16\)31589-6](https://doi.org/10.1016/S0140-6736(16)31589-6)
- Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology*, 80(4), 535–546. <https://doi.org/10.1037/a0028898>
- Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 80(4), 547–559. <https://doi.org/10.1037/a0028226>
- World Health Organization. (2024). Psychological interventions implementation manual: Integrating evidence-based psychological interventions into existing services. <https://www.who.int/publications/i/item/9789240087149>