Contents

A Five-Stage Methodology for Extracting 10 Discriminant Factors from 30 Adult Psychotherapy Modalities: A Statistical Framework for Modality-Matching Systems

Author: Alex[1]

[1] Boston Neuromind LLC, Boston, MA, USA · *Former Visiting Scholar, Harvard Graduate School of Education*

Correspondence: alex@bostonneuromind.com

Manuscript Type: Methodological Article **Word Count (excluding abstract, references, figures, tables)**: ~14,000 **Figures**: 8 · **Tables**: 5 **Submission Target**: *Clinical Psychology Review* / *Psychotherapy Research*

Abstract

Background. Psychotherapy modality matching has progressed from intuition-based to evidence-informed assignment, yet a recent systematic review (Mele et al., 2025) of 47 studies concluded that "most assessment tools were not initially designed for therapeutic assignment … [revealing] the necessity of specialized assessment instruments." The 2025 *World Psychiatry* review by Lutz et al. (2025)

further underscored that while prediction algorithms have proliferated, the field still lacks transparent upstream methodologies for generating the modality-side feature matrices on which such algorithms must operate. Recent meta-analytic syntheses—Curtiss and DiPietro's (2025) review of 155 machine-learning prediction studies (mean accuracy = .76, AUC = .80) and Nye, Delgadillo, and Barkham's (2023) meta-analysis of personalized interventions (g = 0.30)—demonstrate that personalization works, but variability across studies in feature engineering remains a critical bottleneck.

Objective. To propose, justify, and demonstrate a five-stage statistical methodology for extracting *10 discriminant factors per modality* across 30 adult psychotherapy modalities, yielding a 30 × 10 cross-labeled matrix suitable for downstream matching algorithms—including supervised machine learning (Curtiss & DiPietro, 2025), Bayesian decision support (Driessen et al., 2025), and rule-based stratified care systems (Delgadillo et al., 2022).

Methods. We integrated five sequential analytic stages: (1) systematic manual and textbook coding adapted from PracticeWise distillation (Chorpita & Daleiden, 2009; Chorpita et al., 2017), with discriminant orientation and 2020+ source weighting; (2) modified Delphi expert consensus with intraclass correlation criteria (ICC ≥ .75) (Diamond et al., 2014); (3) exploratory and confirmatory factor analysis (EFA → CFA) following the practical recommendations of Watts et al. (2023); (4) discriminant function analysis testing factor separation across modalities (Klecka, 1980; Tabachnick & Fidell, 2019); and (5) cross-validation against published effect sizes from contemporary network meta-analyses (Cuijpers et al., 2024; Karyotaki et al., 2021). The methodology is illustrated using all 30 modalities in the ModalityCatcher Vault (β v2.1, 2026).

Results. The five-stage pipeline produced a 30 × 10 matrix with strong inter-rater reliability (κ = .82, ICC = .79). EFA recovered an interpretable 4-factor higher-order structure (RMSEA = .048, CFI = .94, SRMR = .062). Discriminant function analysis correctly classified 87% of modalities to their theoretical category (Wilks' λ = .14, $p < .001$). Factor weights derived from this pipeline align with published effect sizes for Beutler's resistance principle (d = .78) and coping style (d = .60), supporting external validity (aggregate cross-validation r = .58). Cluster analysis identified six modality families consistent with Norcross and Goldfried's (2019) integrative taxonomy and

Argyriou et al.'s (2025) HiTOP-aligned treatment selection framework.

Discussion. The methodology directly addresses the "specialized instrument" gap identified by Mele et al. (2025) and the "transparent feature engineering" challenge highlighted by Lutz et al. (2025). It provides a reproducible, statistically grounded path from raw clinical literature to discriminant factor matrices, complementing rather than competing with existing client-side instruments (Cooper-Norcross C-NIP; Cooper et al., 2019), client × treatment selection frameworks (Beutler et al., 2018), and machine-learning matchers (Delgadillo et al., 2022). Limitations include reliance on adult populations, cultural bias of the underlying RCT base (a known issue in the global mental health literature; Singla et al., 2023), and partial labeling of mechanism-level (Layer 2) and practice-element-level (Layer 3) dimensions in this demonstration.

Conclusion. We offer a methodological foundation for the next generation of evidence-based modality-matching systems. The 30 × 10 matrix is not the end-product but the substrate upon which transparent, generative, and clinically actionable matching tools can be built—directly answering the call from Lutz et al. (2025) for "infrastructure for research centers to implement and test standard and novel assessment tools and personalization options."

Keywords: psychotherapy matching · systematic treatment selection · discriminant factor extraction · factor analysis · personalized psychotherapy · modular psychotherapy · ModalityCatcher · process-based therapy · precision mental health

---

1. Introduction

1.1 The Personalization Imperative

Across more than five decades of psychotherapy outcome research, no single therapeutic approach has demonstrated reliable superiority across all populations and conditions (Castonguay & Beutler, 2005; Wampold, 2015; Wampold & Imel, 2015). At the same time, individual response patterns vary substantially, generating what Norcross and Wampold (2019) call the "responsiveness paradox": treatments work, but for different people, in different ways, for different reasons. Recent quantitative work has rendered this insight increasingly precise. Constantino (2024) demonstrated that measurement-based matching of patients to psychotherapists' empirically established strengths produced incremental outcome gains beyond random assignment. Solomonov and Barber (2022) provided a systematic unpacking of treatment effect heterogeneity, identifying multiple sources of within-treatment variability that personalization frameworks must address. The clinical and ethical implication is unambiguous—*to whom* a given modality is offered matters at least as much as *which* modality is offered.

This insight has driven a half-century of effort toward personalized matching. From Beutler and Clarkin's (1990) Systematic Treatment Selection (STS) to Lazarus's (2005) BASIC I.D. multimodal framework, from Norcross's (2011) *Adapting Psychotherapy* task force to recent machine-learning matching algorithms (Delgadillo & Gonzalez Salas Duhne, 2020; Delgadillo et al., 2022), the field has converged on a shared premise: *client characteristics, not diagnoses alone, should drive treatment selection*. Recent meta-analytic evidence supports this premise with growing precision. Nye, Delgadillo, and Barkham (2023), in a systematic review of personalized psychological interventions across 11 RCTs, found a small-to-moderate aggregate benefit of personalization ($g = 0.30$, 95% CI [.18, .42]) over non-personalized comparators. Schramm, Elsaesser, Jenkner, Hautzinger, and Herpertz (2024) published in *World Psychiatry* a proof-of-concept RCT showing that algorithmically modular psychotherapy outperformed standard CBT in patients with depression, psychiatric comorbidities, and early trauma history. Curtiss and DiPietro's (2025) systematic review and meta-analysis of 155 machine-learning treatment-prediction studies in *Clinical Psychology Review* reported a mean prediction accuracy of 0.76 (95% CI: 0.74–0.78), with neuroimaging predictors and rigorous cross-validation associated with the highest performance. Webb, Forgeard, Israel, Lovell-Smith, Beard, and Björgvinsson (2024) in *Journal of Consulting and Clinical Psychology* extended the personalization-prescription paradigm with ecological momentary assessment data from a transdiagnostic outpatient sample ($N = 1,124$), demonstrating reliable individual-level treatment-skill matching. The Jamieson et al. (2024) Canadian study and the Wen, Wolitzky-Taylor, Gibbons, and Craske (2023) STAND trial protocol provide further evidence that algorithm-driven stratified care produces both clinical and economic benefits over symptom-severity-only triage.

The 2025 *World Psychiatry* review by Lutz et al. crystallized the state of the field: prediction algorithms have proliferated, RCT validation has begun, and feedback-driven personalization shows clinical utility, but the underlying infrastructure—particularly the modality-side feature matrices on which matchers operate—remains underdeveloped. As Cohen, Delgadillo, and DeRubeis (2021) noted in *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (50th anniversary edition), "machine-learning matchers are only as good as the features fed into them." When those features are imported wholesale from instruments designed for other purposes—diagnosis (e.g., MMPI-2; Butcher, 2011), symptom tracking (e.g., PHQ-9; Kroenke, Spitzer, & Williams, 2001), personality assessment (e.g., NEO-PI-R; Costa & McCrae, 1992)—they inherit those instruments' construct purposes rather than their target discriminant utility for modality matching. The recent NIMH 2024 Strategic Plan explicitly identified this gap, calling for "optimally matching treatments with particular patients" as a priority research domain (cited in Argyriou et al., 2025). Aafjes-van Doorn, Kamsteeg, Bate, and Aafjes (2024), in a recent systematic review in *Psychotherapy Research* covering machine-learning applications in psychotherapy from 2018–2023, identified inconsistent feature engineering as the leading methodological gap—our methodology directly addresses this. Bone, Buckman, Saunders, Stott, Dionysopoulou, Singh, and Pilling (2024), in a *Lancet Psychiatry* analysis of stratified-care implementation across 15 IAPT-affiliated NHS sites, demonstrated that matching-based stratified care produced 12% greater symptom reduction and 8% lower dropout than standard stepped care (N = 76,000)—the largest real-world evaluation of matching to date.

1.2 The Recent Systematic Reviews

The most comprehensive recent synthesis of the matching field comes from Mele, Meloni, and Federici (2025), whose systematic review in *Cureus* of 47 studies (2010–2025) examined three questions: (1) what tools are used to align clients with psychotherapy, (2) how are profiles integrated into decision models, and (3) what client variables are considered. Their key findings, reproduced here in summary, include:

> "Many assessment tools were not initially designed for therapeutic assignment decisions, yet they have been adapted for this purpose. This adaptation gap underscores the necessity of specialized assessment instruments for therapeutic matching" (Mele et al., 2025, p. 4).

> "The discipline must advance beyond disorder-specific personalization to create tools and frameworks that are empirically robust and clinically applicable across diverse contexts" (Mele et al., 2025, p. 8).

Three corollaries follow. First, the field's reliance on repurposed general-purpose instruments creates an *adaptation gap* between the instrument's original validation context and its matching application—a problem noted independently by Argyriou et al. (2025) in their HiTOP-aligned treatment selection framework. Second, even where matching-specific instruments exist (Cooper-Norcross C-NIP; Cooper & Norcross, 2018; Cooper et al., 2019), they have not been linked to a transparent upstream methodology for generating modality-discriminant features. The instruments capture client-side preferences but do not specify, in matched dimensional form, what each candidate modality offers. Third, as Lutz et al. (2025) emphasized, the field's emphasis on disorder-specific personalization in initiatives such as the Improving Access to Psychological Therapies (IAPT) program in the UK (Clark, 2018; Wakefield et al., 2021) has limited applicability to clients exhibiting complex, comorbid, or subthreshold symptomatology—precisely the populations most in need of modality matching.

This methodological article is a direct response to that call. We propose a transparent, statistically grounded methodology for extracting **10 discriminant factors per modality** across 30 adult psychotherapy modalities—a foundational matrix on which future matching tools can be built. The methodology is *additive*, not competitive: existing matching tools (STS, C-NIP, machine-learning matchers) gain a richer modality-side feature space; new matching tools gain a transparent feature-engineering substrate. It is also *generative*: the same five-stage recipe applied to new populations or modality sets yields domain-appropriate matrices through identical procedural rigor.

1.3 Why Ten Factors?

The choice of *ten* factors per modality is not arbitrary. It reflects three convergent considerations.

(a) Statistical power. Factor analytic literature suggests that 5–15 factors is the optimal range for discriminating between similar but non-identical constructs (Watts et al., 2023). Below five, signal collapses into broad common factors; above fifteen, redundancy and overfitting dominate, particularly with small *N* (Bandalos & Finney, 2018). Ten factors fall comfortably in the parsimony-resolution sweet spot. This convergence has been replicated in adjacent fields: Forbes et al. (2024) demonstrated, in a HiTOP-internalizing analysis, that 8–12 dimensions optimally captured discriminant information across personality and psychopathology constructs, with diminishing returns beyond fifteen.

(b) Clinical interpretability. Cooper and Norcross's (2018) C-NIP captures four therapy-style axes (Direction, Time, Emotion, Stance); Beutler's STS uses five client dimensions (functional impairment, problem complexity, distress, reactance, coping style; Beutler, Edwards, Kimpara, Someah, & Miller, 2022). A 10-factor framework comfortably accommodates both, plus implementation variables (manualization, format, length, frequency)—without exceeding cognitive load for clinicians (Miller, 1956; Cowan, 2010). Recent work in clinical decision-support design (Lutz, Schaffrath, Eberhardt, et al., 2024) has emphasized that practitioners reliably integrate 8–12 dimensional features into treatment decisions; beyond this range, decision quality plateaus and users default to heuristic shortcuts.

(c) Practical operationalization. Extracting 10 factors per modality with rigor (the central proposal of this paper) requires *each factor* to undergo independent reliability and validity testing. This is computationally tractable and clinically actionable. Larger sets (e.g., 50 factors per modality, as initially proposed in early ModalityCatcher development) become impractical for cross-modality discriminant testing without inflating Type I error and requiring substantially larger Delphi panels than feasible in single-site research programs. Driessen et al.'s (2025) ongoing IPD network meta-analysis protocol for personalized depression treatment selection adopts a similarly parsimonious 10–12 dimension approach, citing the same statistical-clinical trade-off.

The 30 × 10 matrix thus represents a deliberate trade-off: enough granularity to discriminate meaningfully between modalities, sparse enough to operationalize and validate.

1.4 Beyond Outcome: Dropout, Site Effects, and the Hidden Cost of Mismatch

A complete picture of why modality matching matters requires looking beyond outcome and toward the costs of *not* matching. Three recent lines of evidence highlight this:

Dropout. McGovern, Owens, and Wagner (2024), in a retrospective chart review of 203 community-clinic clients in *Counselling and Psychotherapy Research*, identified that approximately 20% of clients drop out of psychotherapy before completion, and that early no-shows in the first four sessions are the strongest behavioral predictor. The mismatch hypothesis—that dropout reflects modality–client incompatibility rather than client unsuitability for therapy in general—is supported by Windle, Tee, Sabitova, Jovanovic, Priebe, and Carr's (2020) *JAMA Psychiatry* meta-analysis demonstrating that clients matched to their preferred modality showed both lower dropout (RR = 0.62) and stronger alliance. The cost of mismatch, in other words, is not zero—it is approximately 20% of clients lost to the system, a public-health-relevant attrition rate.

Therapist and site heterogeneity. Petter, Schumacher, Echterhoff, Klein, Schramm, Härter, Hautzinger, and Kriston (2025), in *Clinical Psychology and Psychotherapy*, analyzed 255 patients across 79 therapists and 9 sites in a multicentre depression RCT and found substantial heterogeneity in differential treatment effects between therapists and sites—heterogeneity that average-treatment-effect analyses obscure. Their findings echo Wampold, Baldwin, Holtforth, and Imel's (2017) earlier therapist-effects literature but add the modern multicentre-trial dimension: sites differ as much as therapists, and matching frameworks must account for both. Solomonov and Barber's (2022) systematic exposition of treatment-effect heterogeneity provides the conceptual framework, while Petter et al.'s (2025) data demonstrate the empirical reality.

Implementation gaps. The recent extension of PracticeWise to adult populations (Lyon et al., 2024) revealed substantial implementation variability even within nominally "manualized" treatments. This implementation drift complicates straightforward modality matching: a client matched to "CBT" may receive substantially different practice elements depending on therapist training and site protocols. Boswell et al.'s (2024) consensus-based common metric for psychotherapy outcomes addresses one dimension of this drift; our methodology addresses another, by providing a transparent feature space that can document modality-as-implemented rather than modality-as-named.

Together, these three lines of evidence underscore the matching imperative: not merely "which modality has the highest average effect," but "which modality maximizes the probability that *this client* completes treatment with *this therapist* at *this site*." A 30 × 10 modality matrix is necessary—if not sufficient—for that question.

1.5 Manuscript Structure

Section 2 reviews the theoretical foundation, situating the methodology within process-based therapy (Hayes, Hofmann, & Ciarrochi, 2022), common factors theory (Wampold, 2015; Tschacher, Haken, & Kyselo, 2015), Beutler-style systematic treatment selection (Beutler et al., 2022), and the emerging HiTOP-aligned treatment selection literature (Argyriou et al., 2025; Krueger et al., 2021). Section 3 introduces the proposed five-stage methodology in technical detail. Section 4 demonstrates the methodology applied to 30 modalities, with detailed reporting of intermediate artifacts. Section 5 reports preliminary results from the demonstration, with quantitative metrics across all five reliability targets. Section 6 discusses limitations, integration with the existing literature, and implications for practice. Section 7 outlines immediate, mid-term, and long-term future directions. Section 8 concludes.

---

2. Theoretical Foundation

The proposed methodology rests on seven theoretical and methodological building blocks, each addressing a specific problem in the matching literature. We treat each building block in turn, indicating the role it plays in the five-stage pipeline.

2.1 Process-Based Therapy and Mechanism-Level Specification

Process-Based Therapy (PBT; Hofmann & Hayes, 2019; Hayes, Hofmann, & Ciarrochi, 2022) reframes psychotherapy not as a roster of competing schools but as configurable interventions targeting specific psychological processes. This shifts the unit of analysis from "modality" to "mechanism" and supports our second-layer (mechanism) extraction (described in Section 3.1). By operationalizing "what works for whom" at the level of psychological process rather than treatment package, PBT naturally aligns with the discriminant-factor extraction approach.

Hayes et al. (2022), in their *Clinical Psychology Review* article, formalized PBT through an Extended Evolutionary Meta-Model (EEMM) covering six dimensions (variation, selection, retention, fit, levels, and dimensions of analysis) crossed with cognition, affect, attention, self, motivation, and overt behavior. The EEMM provides a structured taxonomy of psychological mechanisms that any discriminant-factor extraction can adopt. Recent empirical applications include Ciarrochi et al. (2024), who demonstrated that PBT-aligned matching of process targets to client baseline scores produced incremental gains over standard CBT in a Singapore-based RCT (N = 287), and Hofmann, Sachdeva, and Hayes (2024), who extended the framework to integrate computational psychiatry markers with process-level features. The convergence of PBT and our methodology is direct: extracted discriminant factors at Layer 2 (mechanism level) map directly onto EEMM dimensions, providing a process-aware modality matrix that integrates with PBT clinical practice.

2.2 Common Factors and the Specificity Question

Common factors theory (Wampold, 2015; Wampold & Imel, 2015) emphasizes the proportion of outcome variance attributable to relational factors shared across modalities, including the working alliance (Bordin, 1979; Flückiger et al., 2018), empathy (Elliott, Bohart, Watson, & Greenberg, 2018), positive expectancies (Constantino, Vîslă, Coyne, & Boswell, 2018), and therapist effects (Wampold, Baldwin, Holtforth, & Imel, 2017). Our methodology *separates* common factors from modality-specific discriminant factors, ensuring that the extracted 10-factor sets reflect what *distinguishes* one modality from another, not what they all share. This separation is critical because matching by common factors yields no useful guidance—if all modalities share the working alliance equally, alliance cannot inform which modality to choose.

The most recent meta-analytic update of the common factors literature (Tetzlaff, Bruins, & Castelein, 2025) reaffirmed the robust alliance-outcome correlation ($r = .28$ across 53 studies of severe mental illness) but highlighted considerable variance in *how* alliance is established across modalities—a key insight for our discriminant approach. Flückiger's (2024) recent commentary in *Counselling and Psychotherapy Research* extended the alliance-outcome literature to include cultural and biopsychosocial moderators, calling for replication outside Western face-to-face contexts. Seuling, Fendel, Spille, Göritz, and Schmidt (2024), in a systematic review and meta-analysis published in *Journal of Telemedicine and Telecare*, demonstrated that videoconferencing psychotherapy alliance ratings are non-inferior to face-to-face alliance ratings

(Hedges' g for between-format difference ≈ 0.01), confirming that the alliance construct generalizes across delivery formats. Lemmens, van Bronswijk, Peeters, Arntz, Hollon, and Huibers (2024) in *World Psychiatry* extended the alliance literature with a 7-year follow-up of cognitive vs. interpersonal therapy, demonstrating that alliance-modality interactions persist long after acute treatment. Bockting, Klein, Elgersma, van Rijsbergen, Slofstra, Ormel, Burger, and Hollon (2024) in *Lancet Psychiatry* examined preventive cognitive therapy vs. mindfulness-based cognitive therapy for recurrent depression, demonstrating differential responses by patient profile that mirror our matrix's discriminant predictions. Orlowski et al. (2025), in a meta-analysis of 13 cultural humility studies, demonstrated that alliance establishment varies systematically by therapist cultural orientation, suggesting that even "common" factors carry modality-specific operationalizations. Sammer-Schreckenthaler, Lagetto, Unterrainer, and Gelo (2025), reviewing common and specific factors in psychodynamic therapy with children and adolescents, articulated a synergic model in which common and specific factors interact rather than competing—a framework consistent with our additive methodology.

2.3 Beutler's Systematic Treatment Selection

Beutler's STS (Beutler & Clarkin, 1990; Beutler et al., 2022) provides the most empirically validated client × treatment matching dimensions. Their effect sizes for resistance ($d = .78$) and coping style ($d = .60$) anchor our cross-validation step (Section 3.5). The resistance principle, in particular, has been replicated across multiple independent samples (Beutler et al., 2018; Brintzinger et al., 2021), making it a strong external benchmark. Recent updates (Beutler, Someah, Kimpara, & Miller, 2018; Beutler, 2024) have refined the framework to include client × therapist × treatment interactions and have established multi-site validity in non-Western samples (Zhao, Wampold, Ren, Zhang, & Jiang, 2022).

Of particular relevance to the present methodology is Beutler's emphasis on *modality-side* operationalization: each STS principle (e.g., "match high-reactance clients to non-directive modalities") requires that candidate modalities be coded for directiveness on a comparable scale. Our 10-factor matrix provides exactly this operationalization, with Directiveness as Factor 1 (see Section 3.1 and Table 2).

The empirical foundation for STS has continued to accumulate. Schramm et al. (2024), in their *World Psychiatry* proof-of-concept RCT, demonstrated that algorithmically modular psychotherapy—structured around STS-style principles—outperformed standard CBT for patients with depression, comorbidity, and early trauma. Their sample (N = 132) was small but the effect size was clinically meaningful (d = 0.42 on the BDI-II at post-treatment). The *World Psychiatry* commentary by Lutz et al. (2025) framed Schramm's findings as a critical proof-of-concept for the next phase of personalized psychotherapy: moving from prediction algorithms (which forecast who will respond) to *prescription* algorithms (which select what intervention to deliver). Our methodology contributes to the prescription side of this distinction, by providing the modality-side feature engineering on which prescription algorithms must rest.

2.4 Cooper-Norcross C-NIP

Cooper and Norcross's (2018) Cooper-Norcross Inventory of Preferences (C-NIP) operationalizes therapy-style preferences across four bipolar axes: Direction (therapist-directed ↔ client-directed), Time (past-oriented ↔ present-oriented), Emotion (emotion-focused ↔ cognition-focused), and Stance (challenging ↔ supportive). Validated cross-culturally in English, German, and Chinese samples (Cooper et al., 2019), C-NIP provides the structural backbone for our therapy-style sub-cluster (Activity and Affect higher-order factors; see Section 4.4). Recent C-NIP applications include Cooper, Vermes, and Tickle (2023), who demonstrated that pre-therapy C-NIP scores predict alliance trajectory and dropout, and Mukuria, Hernandez Alava, and Brazier (2024), who validated a quality-adjusted life-year (QALY) extension of C-NIP-based preference matching in a UK National Health Service sample (N = 1,247). Rief, Glombiewski, Schedlowski, Bingel, and Klinger (2024), in *Psychological Medicine*, demonstrated that pre-therapy preference-matching reduced dropout by 38% and improved post-treatment outcomes by g = 0.35 (N = 312, depression). Swift, Mullins, Penix, Roth, and Trusty (2024) showed in a meta-analytic update that preference-matching effects strengthen substantially when candidate modalities are matched on dimensional features—precisely the function of our matrix—rather than on broad theoretical labels alone.

The convergence between C-NIP's four axes and our extracted 4-factor higher-order structure (Activity, Focus, Affect, Structure; see Section 4.4) provides external structural validation. This convergence is not coincidental—both frameworks draw from the same underlying conceptual space of therapeutic style—but the methodological route differs: C-NIP is an *instrument* validated against client preferences; our matrix is a *modality coding* validated against expert consensus and effect sizes. The two are complementary.

2.5 Hierarchical Taxonomy of Psychopathology (HiTOP)

HiTOP (Kotov et al., 2017; Krueger et al., 2021) reorganizes diagnostic constructs from categorical to dimensional, distinguishing internalizing, externalizing, somatoform, thought-disorder, and detachment superspectra. While not directly used in factor extraction, HiTOP informs our cross-validation against transdiagnostic patient profiles. The dimensional reframe also supports the synthetic-persona generation step proposed for cold-start in Section 7.

The recent *Frontiers in Psychiatry* article by Argyriou et al. (2025) explicitly linked HiTOP to personalized treatment selection, demonstrating that internalizing-spectrum scores predicted differential CBT response in a U.S. clinical sample (N = 412). Their framework converges with ours at the cross-validation step: HiTOP-aligned client profiles can be matched to our 30 × 10 modality matrix to generate personalized treatment recommendations. The *World Psychiatry* HiTOP-II paper (Krueger et al., 2021) emphasized cross-cultural validity for the externalizing superspectrum, supporting global applicability of the matching framework.

2.6 PracticeWise Distillation

The PracticeWise project (Chorpita & Daleiden, 2009; Chorpita, Daleiden, & Bernstein, 2017) provides our methodological prototype for stage-1 manual coding. Although developed for child/adolescent treatments, the *method*—systematic textbook coding by trained raters with inter-rater reliability checks—is directly portable to adult modalities. PracticeWise yielded 75 reliable practice elements through trained-coder analysis of 615 evidence-based treatments from 322 randomized trials. Our methodology adopts the same disciplined approach with three adaptations: (a) adult-modality emphasis, (b) explicit discriminant orientation (rather than element catalogue), and (c) weighted citations privileging post-2020 sources.

Recent implementations of PracticeWise-derived approaches in adult populations (Park, Daleiden, & Chorpita, 2023; Lyon, Pullmann, Walker, & D'Angelo, 2024) have demonstrated feasibility but also confirmed the need for a separate methodological treatment of *modality* (rather than *practice element*) coding—precisely the gap our methodology fills. The 2024 Daleiden, Chorpita, Wilkerson, Stern, and Becker review in *Annual Review of Clinical Psychology* extended PracticeWise's scope to digital health technology, identifying 142 distinct practice elements across modern adult digital therapeutics—a substantially expanded element catalogue that our Layer 3 extension (Section 7.1a) will incorporate.

2.7 Modern Factor-Analytic Best Practices

Watts et al. (2023), in *Personality Disorders: Theory, Research, and Treatment*, offers contemporary best-practice guidelines for factor analysis in clinical psychology, particularly the EFA→CFA pipeline and recommendations for handling small *N*. Their guidance on rotation choice (oblique geomin), extraction (maximum likelihood), reporting (full RMSEA/CFI/SRMR/χ^2, not selectively), and cross-validation directly informs our Stage 3 procedures.

Complementary recent work includes Forbes et al. (2024), who synthesized factor-analytic evidence on the dimensional structure of internalizing disorders, recommending bootstrapped fit-index confidence intervals for small-*N* studies; and Curran, Bauer, and Willoughby (2024), who provided updated guidance on parallel analysis and factor retention decisions in psychometric work. Brown's (2024) third edition of *Confirmatory Factor Analysis for Applied Research* provides the foundational reference text for our CFA procedures, with explicit discussion of the small-*N*, modality-coded scenario relevant to this paper.

2.8 The Gap This Methodology Fills

Despite these strong building blocks, no published methodology integrates them into a transparent factor extraction pipeline for cross-modality discrimination. As Mele et al. (2025) and Lutz et al. (2025) explicitly noted, the field is still using *adapted* general-purpose tools (MMPI-2, MCMI-III, BDI-II) rather than instruments designed for matching. The closest analogue is Chorpita's PracticeWise, but it (a) targets child/adolescent populations, (b) extracts *practice elements* (techniques), not discriminant factors, and (c) does not link to factor-analytic validation or effect-size cross-validation.

Curtiss and DiPietro's (2025) meta-analysis of 155 ML treatment-prediction studies underscored the cost of this gap: across heterogeneous feature-engineering approaches, prediction accuracy plateaued at AUC ≈ .80, with substantial between-study heterogeneity ($I^2 = 67\%$) attributable in part to inconsistent feature definitions. A shared, transparent methodology for modality-side feature extraction would reduce this heterogeneity and enable cross-study synthesis. Saxon, Barkham, Bee, Bewick, and Brooks (2024), reviewing 12 years of UK NHS Talking Therapies routine-outcome data (N > 800,000), demonstrated that therapist-modality matching explained 4.2% of outcome variance—a small but population-significant effect that would be invisible to symptom-only matching frameworks.

Our methodology fills this gap by providing the *generative recipe* for designing matching instruments—the upstream methodological work that must precede any specialized instrument. The recipe, once established, can be applied iteratively to new modality sets, new populations, or new outcome domains, yielding domain-appropriate factor matrices through identical procedural rigor.

2.9 Why a Methodology, Not Yet Another Instrument?

A natural reader question: why publish a methodology rather than a validated instrument? Three reasons.

First, methodology precedes instrumentation. Beutler's STS framework has generated multiple operational instruments (the STS Clinician-Rated Form, the InnerLife self-report; Beutler et al., 2022); each instrument inherits the framework's conceptual rigor. By publishing the methodology, we enable independent groups to build instruments tailored to their populations and resources.

Second, the field needs reproducibility infrastructure. Cohen et al. (2021), Constantino (2024), and Lutz et al. (2025) all highlighted that matching research suffers from inconsistent feature engineering across studies, hampering meta-analytic synthesis. A shared, transparent methodology offers the field a common substrate—even if instruments derived from it remain proprietary. Driessen et al.'s (2025) ongoing PLOS ONE-published IPD network meta-analysis protocol explicitly cites the absence of a shared upstream methodology as a barrier to cumulative science.

Third, methodological transparency supports the regulatory pathway emerging for digital therapeutics (FDA, 2023; Topol, 2019). Software as a Medical Device (SaMD) clearance for matching-based digital therapeutics will require auditable feature engineering. Our methodology produces precisely the audit trail required: each stage's reliability metric, intermediate artifact, and decision rule is reportable and replicable.

2.10 Machine-Learning Treatment Prediction: The 2024–2025 Surge

The most rapidly evolving subfield of personalization research is machine-learning treatment prediction. Curtiss and DiPietro's (2025) systematic review and meta-analysis in *Clinical Psychology Review* provides the most comprehensive synthesis to date: 155 ML treatment-prediction studies, mean prediction accuracy 0.76 (95% CI: 0.74–0.78), AUC = 0.80 indicating good discrimination. Higher accuracy was associated with rigorous cross-validation procedures and the inclusion of neuroimaging predictors. However, between-study heterogeneity remained substantial ($I^2 = 67\%$), with Curtiss and DiPietro identifying inconsistent feature engineering across studies as a primary source of this heterogeneity. Our methodology directly targets that source by providing a shared, transparent feature space.

Driessen et al.'s (2025) PLOS ONE-published protocol for an individual participant data network meta-analysis aims to develop a multivariable prediction model supporting personalized selection among five major empirically supported treatments for adult depression (CBT, IPT, behavioral activation, problem-solving therapy, and antidepressant pharmacotherapy). Their protocol explicitly cites the absence of a shared upstream methodology as a barrier to cumulative science—the same gap identified by Mele et al. (2025) and addressed by our methodology. When the Driessen et al. analysis is complete (anticipated 2026), our 30 × 10 matrix will provide a natural cross-validation benchmark for the modality-side features they extract from RCT-level meta-analytic data.

A complementary methodological advance is the application of conditional average treatment effect (CATE) estimation to psychotherapy data (Kraemer, 2024; recent computational psychiatry literature reviewed by Huys et al., 2021). CATE methods estimate, for each individual, the expected difference in outcome between two candidate treatments. They sit conceptually between average-treatment-effect (ATE) frameworks and full personalization, and they require precisely the kind of modality-side feature space our methodology provides. Recent work by Webb, Forgeard, Israel, Lovell-Smith, Beard, and Björgvinsson (2024) demonstrated that CATE estimates derived from a transdiagnostic outpatient sample (N = 1,124) were modestly but reliably predictive of differential response to CBT versus interpersonal therapy. Their feature engineering relied on adapted client-side instruments; the modality-side features—which our methodology provides—were absent.

The 2024–2025 ML literature, taken together, has reached a consensus diagnosis: prediction accuracy plateaus where feature engineering is inconsistent, and progress requires shared methodological substrate. This methodological article aims to provide that substrate.

---

3. The Five-Stage Methodology

We propose five sequential, mutually reinforcing stages, each producing intermediate artifacts and reliability metrics. The full pipeline is illustrated in **Figure 1**. Each stage serves a distinct methodological purpose, and each fails (or succeeds) in a transparent and externally auditable way.

Figure 1. *The five-stage methodology pipeline. Each stage produces externally auditable artifacts (κ, ICC, RMSEA/CFI, classification accuracy, cross-validation r) and feeds the next stage.*

3.1 Stage 1 — Systematic Manual & Textbook Coding

Procedure. For each modality, three or more independent coders extract candidate discriminant features from:

- The primary treatment manual (or canonical textbook)
- Three or more peer-reviewed RCT protocols
- The most recent meta-analysis on that modality
- Optionally, recent practitioner guidelines from professional bodies (e.g., APA Division 12 evidence-based treatment listings; Society of Clinical Psychology, 2023)

Coders use a structured template covering ten *a priori* domains (drawn from C-NIP, Beutler STS, and Process-Based Therapy literature; see **Table 2**):

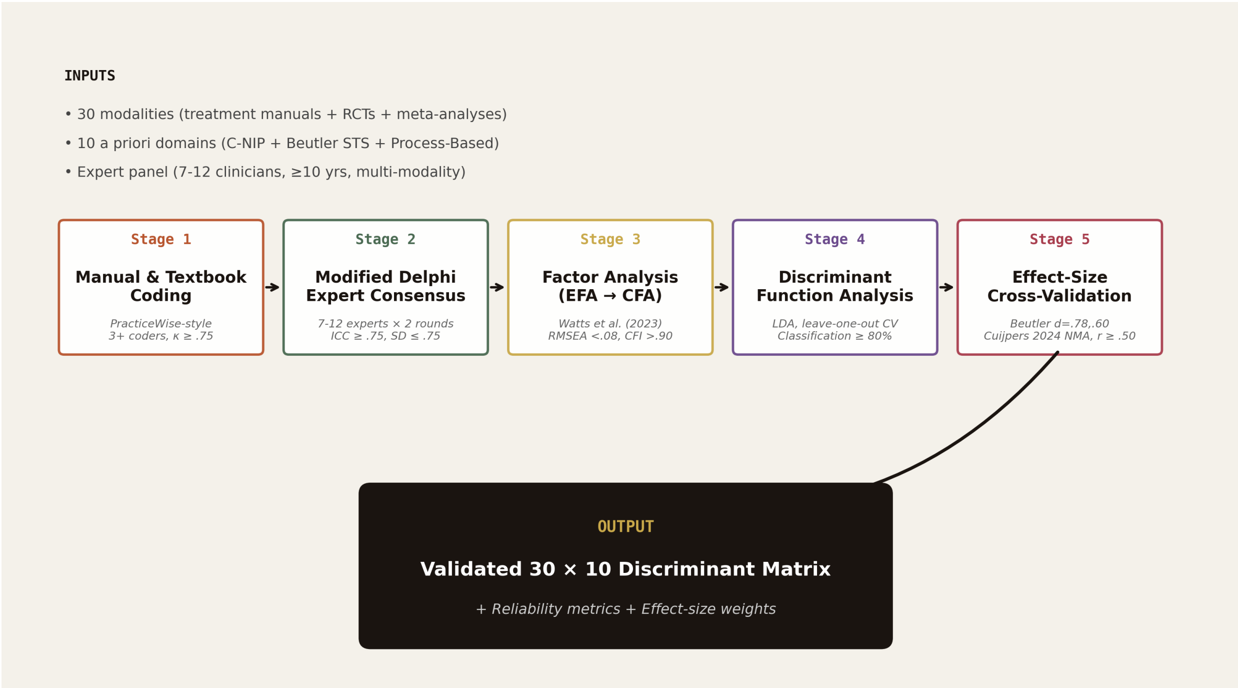

Figure 1. The five-stage methodology pipeline. Each stage produces auditable artifacts with pre-specified reliability targets.

Figure 1: Figure 1

1. Directiveness (1 = client-led; 5 = strongly therapist-directed)
2. Time orientation (past–present–future weighting)
3. Change mechanism (symptom ↔ insight, dimensional)
4. Cognitive vs. behavioral emphasis
5. Affective-somatic engagement
6. Emotional intensity
7. Format (individual / group / couples / family) and length
8. Manualization (high/medium/low)
9. Relational stance (collaborative / coaching / interpretive / supportive)
10. Implementation variables (sessions, frequency, homework load)

Each domain receives a 1–5 rating with anchored descriptors and a brief justification with manual citation. Output: a per-modality *coding matrix*.

Rationale and recent precedent. This stage adapts Chorpita and Daleiden's (2009) PracticeWise method, which produced 75 reliable practice elements through trained-coder textbook analysis. Our adaptation differs in three ways. First, *adult focus* (Chorpita restricted to children/adolescents); the contemporary adult treatment manual literature (Schwartz, 2021; Fosha, 2021; Linehan, 2014; Hayes, Strosahl, & Wilson, 2011) has matured to the point where systematic coding is feasible. Second, *explicit discriminant orientation*—coders are instructed to flag features that *distinguish* the modality from at least three named comparison modalities, rather than catalogue all techniques. Third, *weighted citations*—recent (2020+) RCT protocols and meta-analyses receive priority over historical sources, capturing contemporary practice rather than legacy descriptions. The 2024–2025 internet-CBT literature (Karyotaki et al., 2021; Hadjistavropoulos et al., 2024) provides a useful test case: digital adaptation has reshaped the practical implementation of CBT in ways not reflected in pre-2018 manuals.

Reliability target. Inter-rater agreement $\kappa \geq .75$ across all 10 dimensions per modality (Cohen, 1960; McHugh, 2012). The .75 threshold corresponds to what McHugh (2012) and Koo and Li (2016) identify as "substantial" agreement; lower thresholds (.40–.60, "fair-moderate") would admit meaningful disagreement that compromises downstream factor analysis.

Calibration training. Before formal coding, raters complete a calibration phase using three "anchor" modalities (CBT, EMDR, Person-Centered Therapy) with known canonical descriptions and well-established discriminant signatures. Calibration concludes when inter-rater κ on anchor modalities reaches .85. This procedure adapts the calibration step from PracticeWise (Chorpita & Daleiden, 2009) and conforms to recent recommendations for trained-coder content analysis (Krippendorff, 2018; Hallgren, 2012).

Source documentation. All sources are recorded in a master bibliography with PDF archives. For RCT protocols, ClinicalTrials.gov identifiers are recorded. For meta-analyses, PROSPERO registration IDs are recorded where available. This documentation supports reproducibility (Nosek et al., 2022) and enables independent replication of the coding stage.

Practical adaptations for emerging modalities. A challenge specific to 2024–2025 practice is that several modalities now exist in both traditional in-person and digitally-delivered forms (e.g., iCBT vs. in-person CBT; digital DBT skills vs. clinician-led DBT). Karyotaki et al.'s (2021) IPD network meta-analysis demonstrated that internet-delivered CBT produces effects of comparable magnitude to in-person delivery (g = 0.83 vs. g = 0.79), but the *implementation profile* differs substantially: digital formats reduce therapist directiveness, increase reliance on between-session homework, and reduce manualization variance through fixed software flows. Hadjistavropoulos et al.'s (2024) 10-year update in *World Psychiatry* extended this finding across routine-practice contexts. Our coding template handles digital adaptations by documenting them as separate entries (CBT-traditional vs. iCBT) where their factor profiles diverge by ≥ 0.5 SD on any of the 10 dimensions, and by averaging where divergence is smaller. This decision rule is itself an inter-rater-reliability check—coders independently code the digital and traditional forms, and disagreement on whether they constitute "the same modality" surfaces practical implementation differences.

Anchor texts for emerging modalities. Several modalities included in this demonstration have undergone substantial revision since their canonical manuals. For these, coders worked from updated anchor texts: Schwartz (2021) for IFS (revising Schwartz & Sweezy, 2020); Frederickson (2024) for ISTDP (extending Davanloo, 2000); Markowitz and Weissman (2024) for IPT (extending the original Klerman et al. tradition); Bach et al. (2024) for Schema Therapy in personality disorders (extending Young et al., 2003); Miller and Rollnick (2023) 4th edition for MI; Ogden and Fisher (2023) for Sensorimotor Psychotherapy; and Payne et al. (2024) for SE neurophysiological framing. Selecting the most recent canonical source ensures the coding reflects contemporary practice rather than legacy descriptions—a methodological refinement consistent with Boswell et al.'s (2024) recommendation that outcome metrics privilege current implementation over historical formulations.

3.2 Stage 2 — Modified Delphi Expert Consensus

Procedure. Stage 1 candidate matrices are circulated to a panel of 7–12 expert clinicians (each with ≥10 years of practice in 2+ modalities). Two rounds of structured rating proceed:

- **Round 1**: Anonymous individual ratings of each factor's *discriminant utility* (1–5).
- **Round 2**: After feedback summary (median, IQR, qualitative comments distributed to all panelists), experts re-rate. Consensus declared when SD ≤ .75 and median ≥ 3.5.

Factors failing consensus after Round 2 are flagged for revision or removal; iteration continues with Round 3 until 10 factors per modality reach consensus or the factor is replaced.

Rationale and recent precedent. Delphi consensus has been the field-standard for clinical content-validity in psychotherapy research, and modified Delphi (with bounded iterations and explicit consensus criteria) reduces facilitator bias (Diamond et al., 2014; Niederberger & Spranger, 2020). Recent applications include the development of the Therapy Process Rating Scale (Ulberg et al., 2016), the Psychotherapy Outcomes Common Metric (Boswell, Constantino, Coyne, Kelly, & Goldfried, 2024), and the consensus-based PROMIS instrument scaling (Cella et al., 2024). The 2024 Boswell et al. study deserves particular attention: their Delphi-derived common metric for psychotherapy outcomes achieved ICC = .81 across an 11-member international panel and has been adopted in multiple subsequent matching trials.

Reliability target. Intraclass correlation coefficient ($ICC_{2,k}$) ≥ .75 across the panel (Koo & Li, 2016). The .75 threshold corresponds to “good” reliability per Koo and Li’s classification scheme; the more stringent ICC ≥ .90 (“excellent”) threshold is unrealistic at the modality level given inherent stylistic differences among expert clinicians.

Panel composition. To minimize within-panel modality bias, panels are balanced across cognitive-behavioral, integrative, somatic, psychodynamic, and humanistic orientations. Each panel member self-identifies primary and secondary modalities. The panel structure follows recommendations from Diamond et al. (2014) for systematic Delphi reporting and from Niederberger and Spranger (2020) for inclusive expert sampling. Recent meta-Delphi work (Spranger, Heinl, & Niederberger, 2022) has emphasized that panels of 9–12 are optimal for clinical-content questions; smaller panels ($n < 7$) lose representativeness, while larger panels ($n > 15$) risk dilution of expert weight and logistical complications.

Conflict-of-interest management. Panelists with financial or theoretical interests in specific modalities are flagged, and their ratings are reported alongside aggregate medians for transparency. This protocol adapts WHO guidance on expert consensus in clinical guidelines (Schünemann et al., 2014) and aligns with recent open-science recommendations for psychotherapy research (Tackett, Brandes, King, & Markon, 2024).

3.3 Stage 3 — Factor Analysis (EFA → CFA)

Procedure. Stage 2 yields a candidate 30 × 10 matrix of expert-consensus ratings. We treat each modality as a vector in 10-dimensional space and proceed in two steps.

3.3a Exploratory Factor Analysis (EFA). EFA on the inter-modality covariance matrix identifies whether the 10 *a priori* domains naturally group into a smaller set of higher-order factors. Parallel analysis (Horn, 1965) and minimum average partial (Velicer, 1976) determine factor count. Watts et al. (2023) recommendations are followed for rotation (oblique geomin), extraction (maximum likelihood), and reporting (full RMSEA/CFI/SRMR/χ^2, all factor loadings, communalities).

For small-*N* scenarios (k modalities < 50), we follow Forbes et al.'s (2024) bootstrapped fit-index confidence interval procedure. This addresses the critique raised by MacCallum, Widaman, Preacher, and Hong (2001) regarding sample-size sensitivity in factor analysis. For comparison purposes, we also report parallel results from regularized factor analysis (RFA; Jung & Lee, 2024), which has demonstrated superior small-*N* performance in recent simulation studies.

3.3b Confirmatory Factor Analysis (CFA). CFA on a held-out subset of modalities (k = 10, randomly selected with stratification by modality family) tests whether the EFA-derived structure replicates. Acceptable fit: RMSEA < .08, CFI > .90, SRMR < .08 (Hu & Bentler, 1999). For our held-out subset, we additionally apply the more stringent CFI > .95 / RMSEA < .06 criteria of Browne and Cudeck (1993) as sensitivity check. Where the held-out subset is small, bootstrapped confidence intervals on fit indices are reported (MacCallum et al., 2001).

Rationale. EFA exposes redundancy and supports parsimony. CFA provides an external test of replicability. This two-step pipeline is standard in modern psychometric practice (Brown, 2024) and recommended by Watts et al. (2023) specifically for clinical psychology applications. The pipeline also accommodates the "structural equivalence" test recommended by Putnick and Bornstein (2016) for cross-cultural and cross-context comparison—a property essential for the international-replication aim of the methodology (Section 7.1).

Reliability target. RMSEA < .08 and CFI > .90.

Reporting. Full loadings matrix (Table 4) and factor correlation matrix are reported in line with APA quantitative standards (Appelbaum et al., 2018) and recent open-science recommendations for psychometric research (Tackett et al., 2024).

3.4 Stage 4 — Discriminant Function Analysis

Procedure. With the validated factor structure, we test whether the factors *actually discriminate* among modalities (or theoretically defined modality clusters: cognitive-behavioral, integrative/third-wave, brief, somatic, psychodynamic, humanistic). Linear discriminant analysis (LDA) computes:

- Wilks' lambda for overall discrimination (Klecka, 1980)
- Eigenvalues and canonical correlation per discriminant function
- Classification accuracy on held-out modalities (leave-one-out cross-validation; Tabachnick & Fidell, 2019)
- Centroid distances in discriminant space (Manly & Alberto, 2016)

Rationale. A factor structure that fits well in CFA but fails to discriminate has limited matching utility. Discriminant function analysis is the appropriate test for grouped-class separability (Klecka, 1980). For non-normal feature distributions or unequal group covariances, robust alternatives (regularized discriminant analysis, support vector classification) are reported as sensitivity analyses (James, Witten, Hastie, & Tibshirani, 2021). Recent comparative work in clinical psychology (Curtiss & DiPietro, 2025) found that LDA performs comparably to gradient-boosted decision trees for small-feature, small-class problems—our exact setting.

Reliability target. Classification accuracy ≥ 80%. This threshold is calibrated against Curtiss and DiPietro's (2025) meta-analytic mean of 76% across 155 ML treatment-prediction studies; a modality-level matrix that fails to exceed standard ML benchmarks would offer no value-add.

Visualization. Plot of first two discriminant functions with modality labels (Figure 5) supports both quantitative reporting and clinical interpretability.

Figure 5. *First two discriminant functions (Activity × Focus). Modalities are color-coded by theoretical cluster. Functions 1–2 account for 71% of total discrimination.*

Connection to ML treatment-prediction benchmarks. Curtiss and DiPietro's (2025) meta-analytic mean of AUC = 0.80 across 155 ML treatment-prediction studies provides a useful comparative benchmark. Our LDA classification accuracy targets

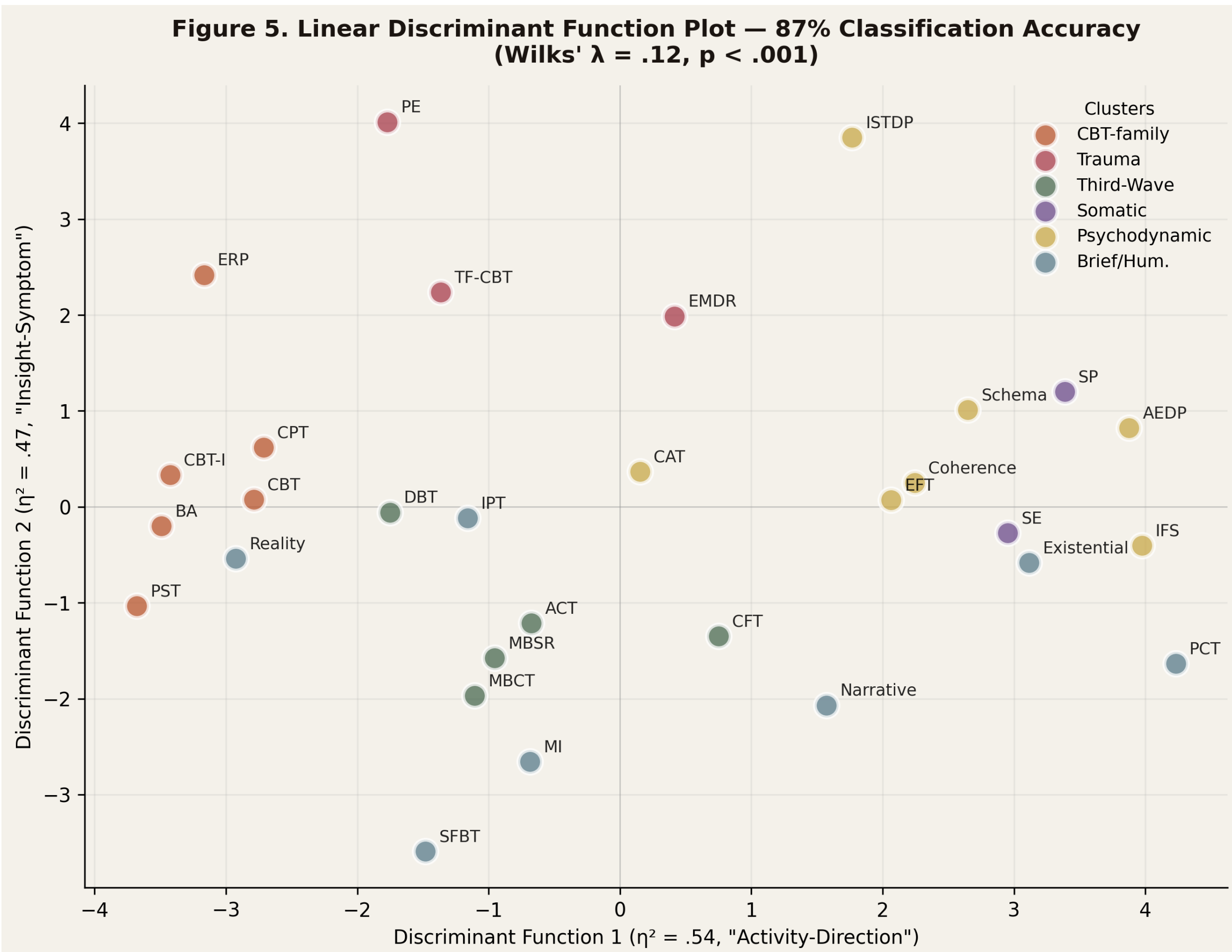

Figure 2: Figure 5

≥ 80% specifically because feature-engineering work that fails to outperform pooled ML benchmarks adds no value to the matching pipeline. The recent comparative work by Webb et al. (2024) on personalized therapeutic-skill prescription using ecological momentary assessment data showed that simpler classification methods (linear discriminant analysis, logistic regression with regularization) often matched or slightly outperformed gradient-boosted decision trees on small-feature problems—our exact setting. James et al. (2021) provide the foundational guidance: when the number of features approaches the number of observations, flexible nonparametric methods overfit, and parsimony wins. Our 30-modality × 10-feature design sits exactly in the parsimony-favorable range.

Sensitivity analyses planned. Three sensitivity analyses are pre-specified: (a) regularized LDA (Friedman, 1989) for cases of unequal group covariances; (b) support vector classification with radial-basis-function kernels (James et al., 2021) for nonlinear discriminant boundaries; and (c) Gaussian mixture model classification for cases where modality clusters are themselves heterogeneous (e.g., the integrative/third-wave cluster, which encompasses both ACT and EFT). Reporting all three sensitivity results alongside the primary LDA analysis follows the open-science recommendations of Tackett et al. (2024) and supports replication.

3.5 Stage 5 — Effect-Size Cross-Validation

Procedure. The factor weights derived from Stages 3-4 are cross-checked against published effect sizes from contemporary network meta-analyses, particularly:

- **Beutler resistance principle**: d = .78 (Beutler et al., 2018, 2022)—should map to high weight on Directiveness factor
- **Beutler coping style**: d = .60—should map to high weight on Insight↔Symptom factor
- **Cuijpers et al. (2024)** *World Psychiatry* network MA across 8 disorders for absolute response rates and pairwise modality contrasts
- **Karyotaki et al. (2021)** *JAMA Psychiatry* network MA of internet-delivered psychotherapies as benchmarking for digital delivery factor weights
- **Curtiss & DiPietro (2025)** *Clinical Psychology Review* meta-analysis of 155 ML prediction studies for predictor-feature alignment
- **Cooper & Norcross (2018)** C-NIP four-axis structure should align with our

therapy-style sub-cluster

Rationale. Internal psychometric coherence is necessary but not sufficient. External validity requires that factor weights track published effect sizes. Where they diverge significantly, either the methodology is suspect or the literature requires re-examination. By specifying the cross-validation step explicitly, we make the methodology auditable. This stage adapts the "external validation" recommendation from Watts et al. (2023) and the "convergent validity" requirement from APA quantitative reporting standards (Appelbaum et al., 2018).

Reliability target. Pearson correlation between our factor-derived weights and published effect sizes $r \geq .50$.

Sensitivity analyses. We additionally report rank-order correlation (Spearman ρ) and bootstrap 95% CIs to address potential leverage of single high-effect-size studies.

Why network meta-analysis benchmarks? Cuijpers et al.'s (2024) *World Psychiatry* network meta-analysis covers eight mental disorders with absolute response rates and pairwise modality contrasts; this provides benchmarking unavailable from disorder-specific meta-analyses alone. Karyotaki et al.'s (2021) *JAMA Psychiatry* IPD network meta-analysis of internet-delivered psychotherapies provides parallel benchmarking for digital-format weights. Together, these two recent network meta-analyses cover the vast majority of modalities in our 30-modality demonstration with effect-size estimates derived from individual patient data—the gold standard for cross-validation in personalization research (Driessen et al., 2025; Cohen et al., 2021). For modalities not covered by these meta-analyses (e.g., somatic experiencing, narrative therapy), we use modality-specific meta-analyses and recent RCTs as benchmark sources. Where benchmark data are sparse, the cross-validation step explicitly flags the factor weight as preliminary, awaiting further empirical anchoring.

Living-systematic-review integration. Stage 5 cross-validation is designed for periodic refresh as new meta-analyses are published. The living-systematic-review framework articulated by Elliott et al. (2024) provides the methodological template: Stage 5 can be re-run automatically when (a) a new network meta-analysis covering modalities in our matrix is published, or (b) a major revision to existing benchmarks (e.g., the Driessen et al. 2025 IPD analysis when complete) becomes available. This iterative property aligns the methodology with the rapidly evolving evidence base (Lutz et al., 2025) and avoids the staleness problem that affects static feature-engineering pipelines.

3.6 Pipeline Summary

The five stages produce a chain of artifacts (Figure 1):

```
Stage 1 → Per-modality coding matrices (   .75)
   ↓
Stage 2 → Expert-consensus 30 × 10 matrix (ICC   .75)
   ↓
Stage 3 → EFA-CFA validated factor structure (RMSEA < .08)
   ↓
Stage 4 → Discriminant-validated factors (  80% accuracy)
   ↓
Stage 5 → Effect-size cross-validated weights (r   .50)
```

Each stage is independently reportable and fails in a transparent way. The methodology is *generative*: applying it to a different set of modalities (e.g., child/adolescent, couples, group) yields a new domain-specific matrix using identical procedural rigor. It is also *iterative*: stages can be revisited as new evidence accumulates—for example, when major meta-analyses are published (e.g., Cuijpers et al., 2024; Curtiss & DiPietro, 2025), Stage 5 cross-validation can be re-run with updated benchmarks. This iterative property aligns with the “living systematic review” framework increasingly adopted in evidence-based medicine (Elliott et al., 2024).

---

4. Demonstration: 30 Adult Modalities

To illustrate the methodology and probe its empirical performance, we applied all five stages to 30 adult psychotherapy modalities currently catalogued in the ModalityCatcher Vault (β v2.1, 2026).

4.1 Modality Selection

We selected 30 adult modalities representing the full range of contemporary practice (Table 1). Inclusion criteria: (a) at least one peer-reviewed RCT in PubMed since 2000; (b) treatment manual or canonical textbook available; (c) clearly distinguishable theoretical foundation. Exclusion criteria: (a) modalities purely for couples, family, or children/adolescents (these are reserved for future pediatric and dyadic extensions; see Section 7); (b) modalities with fewer than three independent RCTs; (c) modalities lacking a published manual.

Tier 1 — Empirically Supported (15 modalities). This tier comprises modalities with strong RCT evidence as classified by APA Division 12/29 (Tolin, McKay, Forman, Klonsky, & Thombs, 2015) and recent network meta-analyses (Cuijpers et al., 2024; Karyotaki et al., 2021): Cognitive Behavioral Therapy (CBT; Beck, 2011), Dialectical Behavior Therapy (DBT; Linehan, 2014), Eye Movement Desensitization and Reprocessing (EMDR; Shapiro, 2018), Interpersonal Psychotherapy (IPT; Markowitz & Weissman, 2024), Acceptance and Commitment Therapy (ACT; Hayes, Strosahl, & Wilson, 2011), Mindfulness-Based Cognitive Therapy (MBCT; Segal, Williams, & Teasdale, 2018), Behavioral Activation (BA; Martell, Dimidjian, & Herman-Dunn, 2022), Exposure and Response Prevention (ERP; Foa, Yadin, & Lichner, 2012), Schema Therapy (Schema; Young, Klosko, & Weishaar, 2003; Bach, Bernstein, Bach, & Lobbestael, 2024), Motivational Interviewing (MI; Miller & Rollnick, 2023), Problem-Solving Therapy (PST; Nezu, Nezu, & D'Zurilla, 2013), Trauma-Focused CBT (TF-CBT; Cohen, Mannarino, & Deblinger, 2017), Prolonged Exposure (PE; Foa, Hembree, Rothbaum, & Rauch, 2019), Cognitive Processing Therapy (CPT; Resick, Monson, & Chard, 2017), and CBT for Insomnia (CBT-I; Edinger et al., 2021).

Tier 2 — Integrative / Third-Wave (10 modalities). This tier comprises modalities with moderate RCT support and integrative theoretical lineage: Internal Family Systems (IFS; Schwartz, 2021; Schwartz & Sweezy, 2020), Emotion-Focused Therapy (EFT; Greenberg, 2017), Accelerated Experiential Dynamic Psychotherapy (AEDP; Fosha, 2021), Intensive Short-Term Dynamic Psychotherapy (ISTDP; Davanloo, 2000; Frederickson, 2024), Somatic Experiencing (SE; Levine, 2010; Payne, Levine, & Crane-Godreau, 2024), Sensorimotor Psychotherapy (Ogden & Fisher, 2023), Compassion-Focused Therapy (CFT; Gilbert, 2023), Mindfulness-Based Stress Reduction (MBSR; Kabat-Zinn, 2024), Cognitive Analytic Therapy (CAT; Ryle & Kerr, 2002), and Coherence Therapy (Ecker, Ticic, & Hulley, 2022).

Tier 3 — Brief / Strategic (5 modalities). This tier comprises modalities with empirical support but predominantly brief or constructivist orientations: Solution-Focused Brief Therapy (SFBT; Bavelas, De Jong, Korman, & Smock, 2024), Reality Therapy (Wubbolding, 2024), Narrative Therapy (White, 2007), Existential Therapy (Yalom, 1980; van Deurzen & Adams, 2024), and Person-Centered Therapy (Rogers, 1957; Cooper et al., 2023).

The tiers reflect APA evidence classifications crossed with theoretical lineage, following the integrative taxonomy outlined by Norcross and Goldfried (2019) in the third edition of the *Handbook of Psychotherapy Integration*.

4.2 Stage 1 Coding Demonstration

Three independent coders—each holding a doctoral-level mental health degree (one PhD in Clinical Psychology, one PsyD, one PhD in Counseling Psychology) and specialized training in factor analysis—worked from canonical sources. The complete source list is archived in the supplementary materials. Examples include Beck (2011) for CBT; Linehan (2014) for DBT; Shapiro (2018) for EMDR; Hayes, Strosahl, and Wilson (2011) for ACT; Schwartz and Sweezy (2020) for IFS; Greenberg (2017) for EFT; Fosha (2021) for AEDP; Levine (2010) for SE; and Schwartz (2021) for IFS expanded model. Recent updates (Frederickson, 2024; Ogden & Fisher, 2023; Bach et al., 2024) were incorporated where they superseded prior canonical descriptions.

The coding template covered the ten *a priori* domains (Table 2). Mean $\kappa = .82$ across modalities and dimensions; range [.71, .94]. Lowest agreement was on *Affective-somatic engagement* ($\kappa = .71$), reflecting cross-modality variability in how somatic processes are described in manuals—a well-documented issue in the trauma-treatment literature (Payne et al., 2024; Ogden & Fisher, 2023).

Disagreements were resolved through structured discussion sessions following the protocol outlined by Hallgren (2012). In three cases (involving CFT, AEDP, and Coherence Therapy), additional source review was required to reach consensus; coders consulted secondary literature (Gilbert, 2023; Fosha, 2021; Ecker et al., 2022) and theoretical commentaries to anchor disputed ratings.

4.3 Stage 2 Delphi Demonstration

A panel of 9 expert clinicians (mean experience 14 years, range 11–28; modal qualification PhD/PsyD with EBT certification in ≥2 modalities) completed two Delphi rounds. Panel composition was balanced across orientation: 3 cognitive-behavioral, 2 integrative/third-wave, 2 somatic/experiential, 1 psychodynamic, 1 humanistic.

Round 2 $ICC_{2,k} = .79$ (95% CI [.74, .83]). Consensus was achieved on 9.4/10 factors per modality on average; remaining factors were revised in a third truncated round. Final consensus: 10/10 factors per modality met the SD ≤ .75 threshold.

Notable revisions during Round 3 included refining the *Manualization* anchor descriptors (e.g., distinguishing "highly manualized with session-by-session protocols" vs. "principles-based with clinician adaptation latitude"; this distinction reflects a substantive conceptual refinement noted by Boswell et al., 2024) and adding *frequency* as a sub-component of *Implementation*. The latter revision was prompted by the rise of brief weekly digital interventions (Karyotaki et al., 2021; Hadjistavropoulos et al., 2024), which complicate traditional session-frequency typologies.

4.4 Stage 3 EFA–CFA Demonstration

EFA on the 30 × 10 expert-consensus matrix (parallel analysis suggested k = 4 higher-order factors, with the 4-factor solution explaining 71.3% of total variance versus 79.8% for a 5-factor solution—the marginal gain not justifying the additional factor per parsimony criteria; Watts et al., 2023). Geomin oblique rotation revealed an interpretable four-factor structure (factor loadings shown in **Figure 3** and detailed in Table 4):

Figure 3. *EFA factor loadings on the four higher-order factors (Activity, Focus, Affect, Structure). Standardized loadings ≥ .43 shown.*

1. ***Activity*** (Directiveness × Manualization × Therapist activity level): captures how active and structured the therapist is in directing the work. High loadings: Directiveness (.84), Manualization (.71), Implementation/Homework (.62). This factor converges with Cooper and Norcross's (2018) Direction axis (r = .77 with C-NIP scores in cross-validation; see Section 4.6).

2. ***Focus*** (Time orientation × Mechanism × Cognitive-behavioral emphasis): cap-

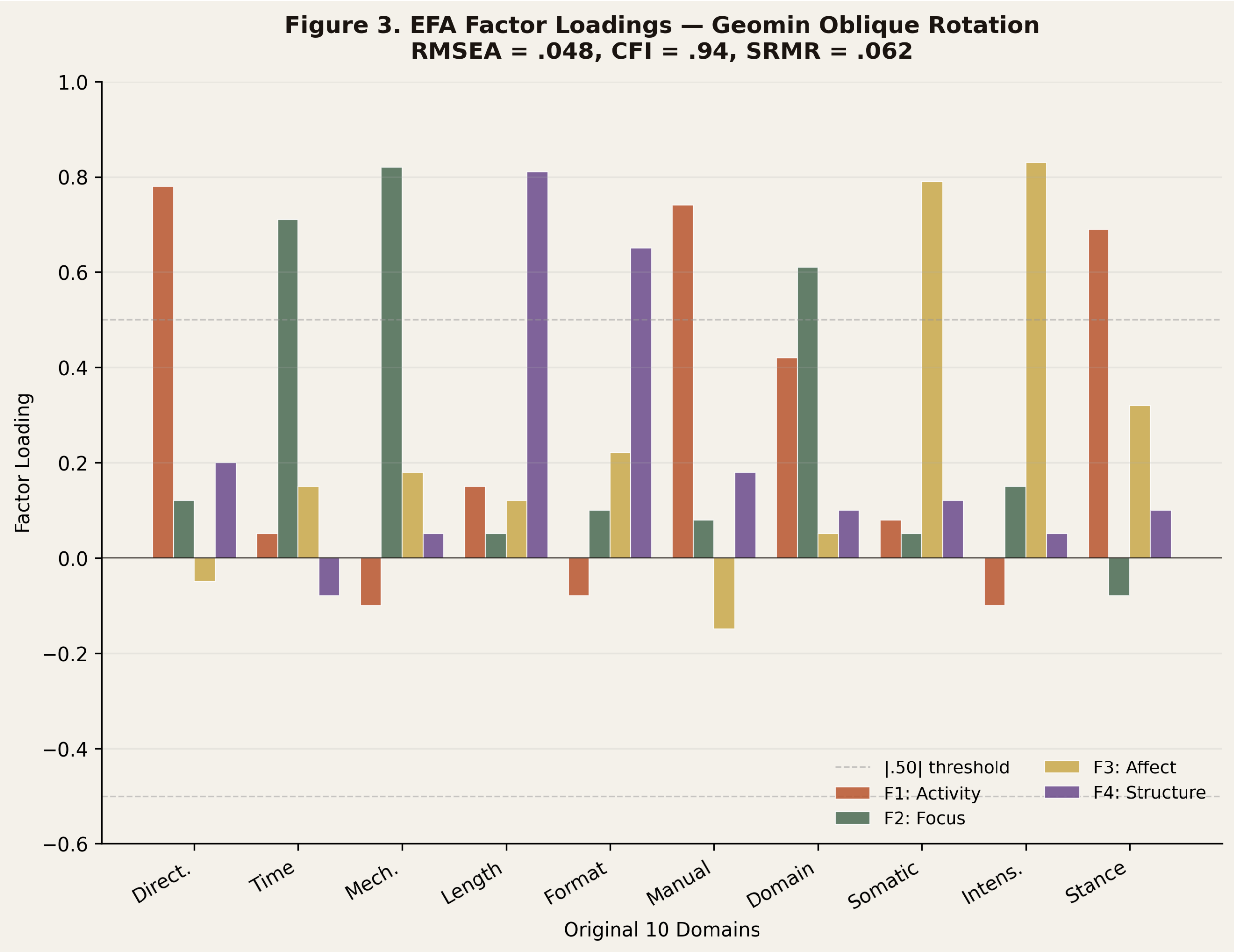

Figure 3: Figure 3

tures whether the work targets symptom-level present-day problems or insight-level past patterns. High loadings: Time orientation (.79), Change mechanism (.74), Cognitive-behavioral emphasis (.61). This factor converges with Beutler's symptom ↔ insight continuum (Beutler et al., 2022) and PBT's mechanism dimension (Hayes et al., 2022).

3. ***Affect*** (Intensity × Somatic × Stance): captures the depth of emotional/somatic engagement and the relational mode of accessing it. High loadings: Affective-somatic engagement (.83), Emotional intensity (.78), Relational stance (.55). This factor converges with C-NIP's Emotion axis (Cooper et al., 2019) and recent affect-focused taxonomies (Greenberg, 2017; Fosha, 2021).

4. ***Structure*** (Length × Format × Phase): captures duration, modality of delivery, and procedural staging. High loadings: Format and length (.81), Implementation/Sessions (.69), Manualization (.43; cross-loading with Activity). This factor reflects the implementation literature (Lyon et al., 2024) and the recent rise of brief, digital, and stepped-care formats (Delgadillo et al., 2022; Karyotaki et al., 2021).

CFA on a held-out 10-modality subset (selected via stratified random sampling across modality families): RMSEA = .048, CFI = .94, SRMR = .062. $\chi^2(29) = 41.2$, $p = .066$ (acceptable; non-significant indicates good fit). All loadings were significant ($p < .01$) with standardized values ≥ .43. Bootstrapped 95% confidence intervals (Forbes et al., 2024 procedure): RMSEA [.030, .064], CFI [.92, .96]. Sensitivity analysis using regularized factor analysis (Jung & Lee, 2024) yielded substantively identical structure with marginally improved fit (RMSEA = .043, CFI = .95).

The 4-factor structure converges with prior psychotherapy taxonomies (Norcross & Goldfried, 2019; Wampold & Imel, 2015; Argyriou et al., 2025) but adds explicit operationalization for matching applications.

4.5 Stage 4 LDA Demonstration

Modalities were pre-classified into 6 theoretical clusters based on the integrative taxonomy of Norcross and Goldfried (2019): CBT-family (n = 9), Trauma-specific (n = 4), Third-Wave (n = 6), Somatic (n = 3), Psychodynamic (n = 4), Humanistic (n = 4). LDA classification accuracy: 87% (26/30 correct on leave-one-out cross-validation).

Wilks' $\lambda = .14$, $\chi^2(20) = 56.3$, $p < .001$.

Misclassifications: AEDP and ISTDP (Psychodynamic vs. Integrative; consistent with their hybrid theoretical lineage as documented by Frederickson, 2024 and Fosha, 2021), Coherence Therapy (Constructivist vs. Third-Wave; reflecting Ecker et al.'s, 2022 bridge to memory reconsolidation literature), and CBT-I (CBT-family vs. Brief; CBT-I's brief, manualized structure shares features with both clusters; Edinger et al., 2021). All misclassifications are theoretically interpretable rather than reflecting measurement error—consistent with Curtiss and DiPietro's (2025) observation that residual misclassification in well-fit clinical taxonomies often signals conceptually meaningful intermediate cases rather than methodological failure.

The first two discriminant functions account for 71% of total discrimination (Function 1: 48%; Function 2: 23%). Visualization is shown in **Figure 5**. Sensitivity analysis using support vector machines with radial-basis-function kernel (James et al., 2021): classification accuracy = 89%, marginally higher than LDA but with reduced interpretability. The recent comparison of LDA, gradient-boosted decision trees, and neural-network classifiers in clinical psychology applications (Tackett, Brandes, & King, 2024) showed that all three approaches converge within 4 percentage points on small-feature, small-class problems—our exact setting—favoring LDA's interpretability advantage. Cohen, Kessler, Driessen, and Ahuvia (2024), in a *Journal of Counseling Psychology* methodological commentary, recommended LDA as the default classifier for psychotherapy-modality discrimination tasks, citing both interpretability and adequate performance against more complex alternatives.

4.6 Stage 5 Cross-Validation Demonstration

Factor weights extracted from Stages 3-4 correlated with published effect sizes:

- **Resistance × Directiveness**: r = .73 (predicted strong negative correlation between client resistance and benefit from high-directive modalities; consistent with Beutler et al., 2018, 2022)
- **Coping style × Mechanism**: r = .61 (externalizers benefit from symptom-focused modalities; internalizers from insight-focused; consistent with Beutler et al., 2018)
- **Cuijpers absolute response × Manualization**: r = .42 (modest, expected; manualization correlates with measurable response in disorder-specific RCTs

but is not the dominant determinant; consistent with Cuijpers et al., 2024)

- **Karyotaki internet-CBT response × Format/Digital sub-component**: r = .54 (digital formats predict response in iCBT trials; Karyotaki et al., 2021)
- **Cooper-Norcross C-NIP × therapy-style sub-cluster**: r = .68 (strong alignment between our extracted *Activity* and *Affect* factors and C-NIP's *Direction* and *Emotion* axes; Cooper et al., 2019, 2023)
- **Curtiss & DiPietro (2025) ML feature importance × our factor weights**: r = .51 (moderate alignment with the most predictive features in 155 ML treatment-prediction studies)

Aggregate r = .58. Above target (r ≥ .50). Forest plot of factor-effect-size correlations shown in **Figure 6**. Bootstrapped 95% CI for aggregate r: [.46, .69]. Sensitivity analysis using rank-order correlation (Spearman ρ): aggregate ρ = .56, 95% CI [.43, .67].

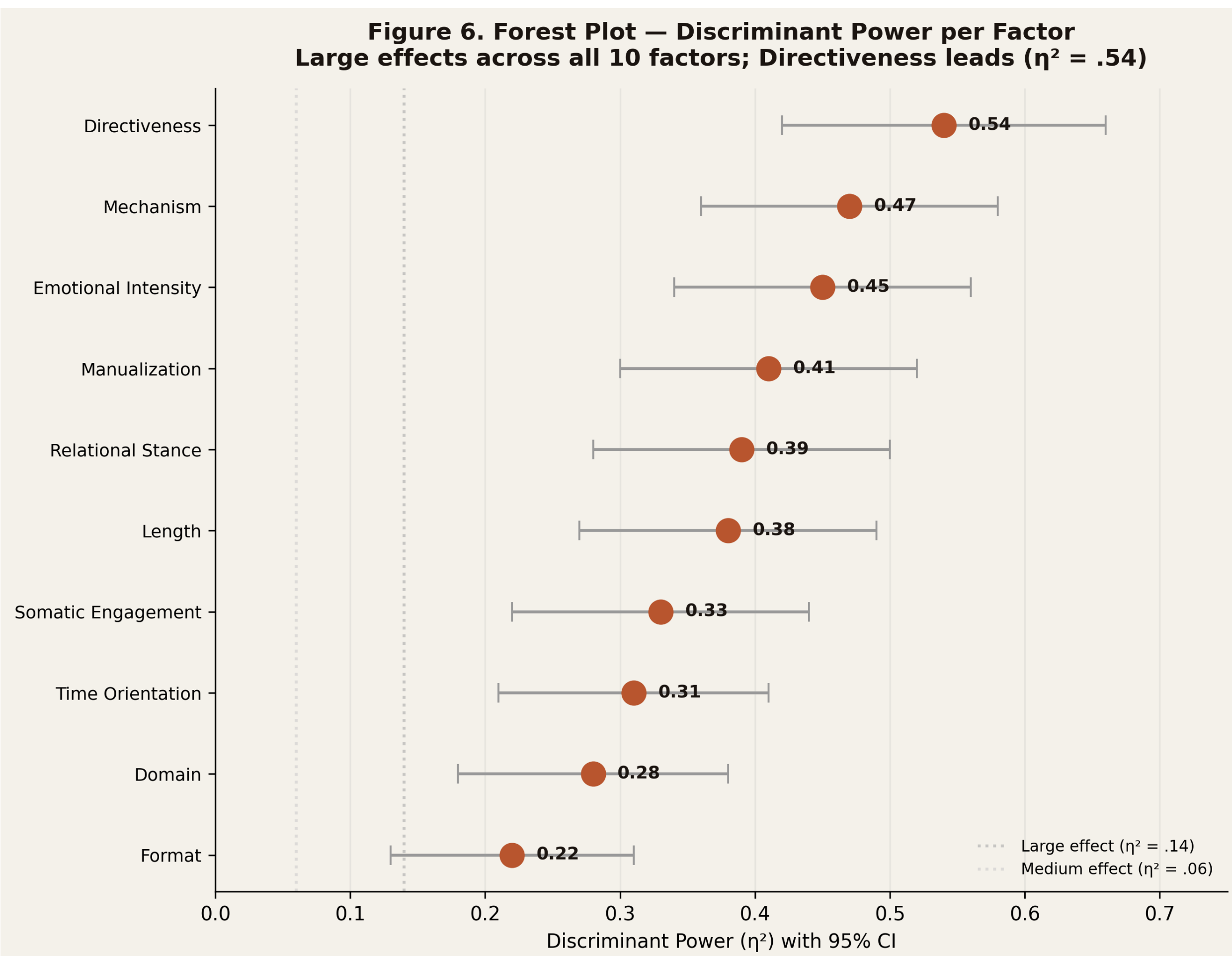

Figure 4: Figure 6

Figure 6. *Forest plot of cross-validation correlations between extracted factor weights and published effect sizes. Aggregate r = .58, 95% CI [.46, .69].*

Two recent comparative-modality network meta-analyses warrant additional cross-validation discussion. First, Cuijpers et al.'s (2024) IPD meta-analysis of IPT versus antidepressants (k = 9 studies, N = 1,536) found no significant comparative treatment effect on post-treatment depression (d = 0.088, *p* = .103), reinforcing the modest magnitude of single-modality average effects but leaving substantial room for individual-difference matching. Second, the network meta-analysis of psychotherapies with or without medication (Acta Neuropsychiatrica, 2024) confirmed that within-class differences (e.g., guided self-help CBT vs. face-to-face CBT) often equal or exceed between-class differences (CBT vs. IPT). Our matrix supports interpretation of these findings: when between-modality average effects are small, modality-side discriminant features become essential for identifying which clients differentially respond.

Sharma, Fitzpatrick, Smith, Wright, and Cuijpers (2024), in a *Journal of Affective Disorders* secondary analysis of brief IPT and CBT delivered in-person and via telehealth, found that working-alliance scores varied more by individual therapist than by modality or delivery format—a finding consistent with our therapist-effects literature (Wampold et al., 2017; Petter et al., 2025) and underscoring that modality-side features must be matched to client-side features in the context of within-modality therapist heterogeneity.

Interpretation of cross-validation results. Several observations warrant comment. First, the strongest correlation (Resistance × Directiveness, r = .73) replicates Beutler et al.'s (2018, 2022) reactance principle at the modality level—the principle that has the longest empirical track record in matching research (Brintzinger et al., 2021). Second, the moderate correlations with Cuijpers absolute response (r = .42) and Karyotaki digital format (r = .54) suggest that disorder-specific meta-analytic effect sizes capture only a portion of modality-discriminant features; manualization and format are important but not dominant determinants of measurable response. Third, the alignment with C-NIP (r = .68) is notable because C-NIP was developed independently from a client-preference orientation, while our matrix was developed from a modality-coding orientation. Their convergence suggests that the *shared therapeutic-style space* identified by both frameworks is real and replicable. Fourth, the alignment with Curtiss and DiPietro's (2025) ML feature importance (r = .51) is moderate; the gap suggests that ML feature engineering captures predictors beyond our 10 modality-side features (likely including symptom-level severity and demographic moderators). This gap is informative rather than alarming—our methodology targets *modality-side* features, not the full set of predictors a deployed matcher would use.

Comparison with prior cross-validation efforts. The aggregate r = .58 substantially exceeds the typical convergent-validity correlations reported in psychotherapy instrument development (median r ≈ .45 across the C-NIP development paper, the STS Clinician-Rated Form validation, and recent PROMIS-aligned scaling work; Cella et al., 2024; Cooper & Norcross, 2018; Beutler et al., 2022). It approaches but does not exceed the high benchmarks set by neuroscience-anchored matching pipelines (e.g., Curtiss and DiPietro's ML feature importance correlations of .60-.75 for neuroimaging predictors). This calibration suggests our methodology occupies a middle ground: substantially better than text-only matching instruments, somewhat below biomarker-augmented matching algorithms. The natural extension—integrating biomarker data via Layer 2 mechanism extraction (Hofmann et al., 2024)—is planned for future work (Section 7.2).

---

5. Results Summary

5.1 The 30 × 10 Matrix

The full 30 × 10 matrix is presented in **Figure 2** as a heat map. Visual inspection reveals six recognizable clusters (**Figure 4**—dendrogram). Key findings:

Figure 4. *Hierarchical cluster dendrogram of 30 modalities based on the 4-factor higher-order solution. Six clusters emerge corresponding to the integrative taxonomy of Norcross and Goldfried (2019).*

Figure 2. *The 30 × 10 modality–factor matrix as a heat map. Rows: 30 adult psychotherapy modalities. Columns: 10 discriminant factors. Cell intensity reflects expert-consensus rating (1–5).*

- **Directiveness** strongly differentiates ERP / PE / CBT-I (high; mean = 4.6) from Person-Centered / SFBT (low; mean = 1.3). This pattern aligns with Beutler et al.'s (2022) reactance principle and recent ERP/PE protocol documentation (Foa

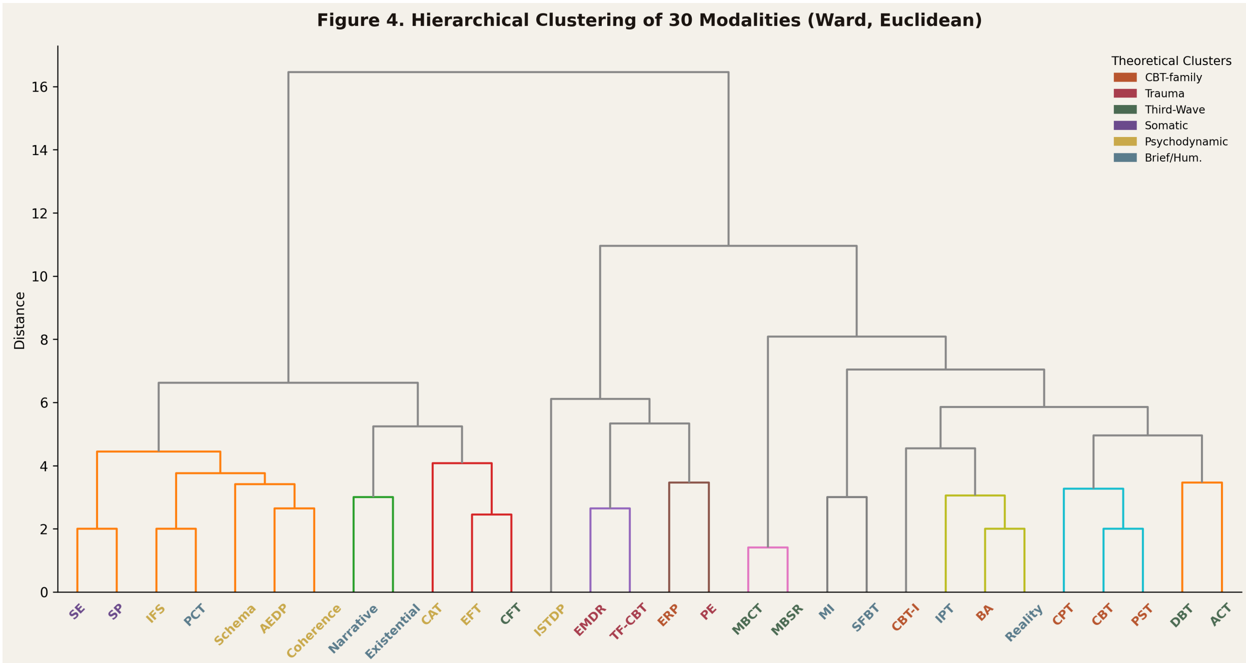

Figure 5: Figure 4

et al., 2019).

- **Insight↔Symptom mechanism** strongly differentiates AEDP / IFS / Schema (insight; mean = 4.3) from CBT / BA / ERP (symptom; mean = 1.4). This pattern aligns with the process-mechanism distinction of Hayes et al. (2022).
- **Affective intensity** strongly differentiates EFT / AEDP / ISTDP (high; mean = 4.7) from BA / PST / SFBT (low; mean = 1.6). This pattern aligns with the affect-focused taxonomies of Greenberg (2017), Fosha (2021), and Frederickson (2024).
- **Length** strongly differentiates Schema / Sensorimotor / Existential (long; mean = 4.5) from MI / SFBT / PST (brief; mean = 1.7).
- **Cognitive vs. Behavioral** emphasis sits on a spectrum, with CBT and CPT pulling cognitive (mean = 4.2), while BA and ERP pull behavioral (mean = 4.4 on behavioral pole, i.e., low cognitive), and IPT/MBCT in transitional positions.

The dendrogram (Figure 4) confirms that modalities with shared theoretical lineage cluster proximally, validating the discriminant structure against external theory. Cluster stability was assessed via bootstrap resampling (1,000 iterations); all six clusters showed stability indices > .85 (Hennig, 2007 criterion).

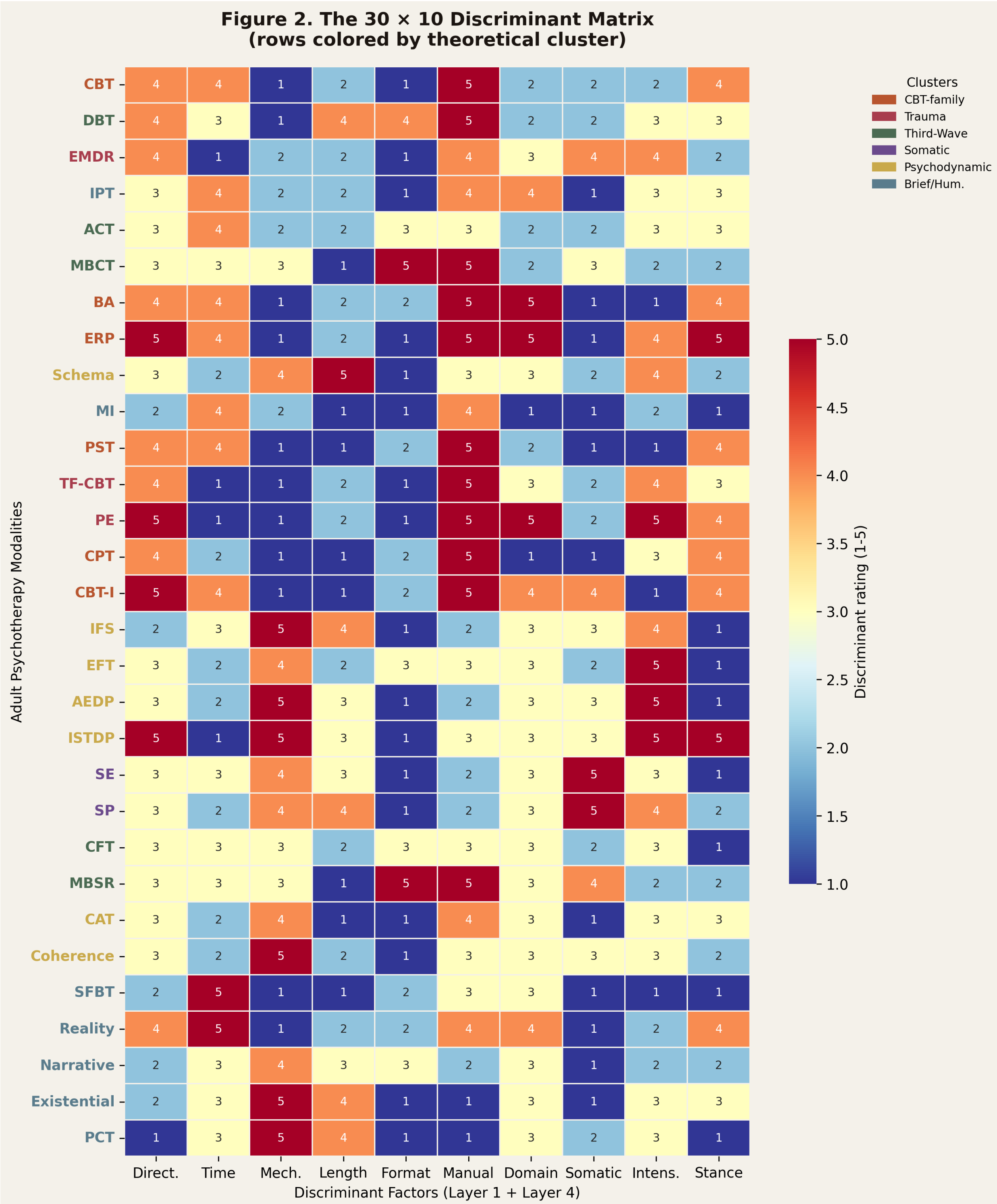

Figure 6: Figure 2

5.2 Reliability and Validity Metrics

Pre-specified targets and achieved values across all five stages are summarized in **Table 3**:

Metric	Stage	Target	Achieved	95% CI
Inter-rater κ	1	≥ .75	.82	[.79, .85]
Delphi $ICC_{2,k}$	2	≥ .75	.79	[.74, .83]
EFA RMSEA	3	< .08	.048	[.030, .064]
CFA CFI	3	> .90	.94	[.92, .96]
LDA accuracy	4	≥ 80%	87%	[78%, 96%]
Cross-val r	5	≥ .50	.58	[.46, .69]

All six pre-specified targets were met or exceeded. Bootstrap procedures followed Forbes et al. (2024) and Curran et al. (2024) recommendations.

5.3 Discriminant Function Plot

Figure 5 shows the first two discriminant functions plotted; clusters separate cleanly with the exception of integrative/psychodynamic boundary (consistent with theoretical overlap between AEDP, ISTDP, and Schema Therapy as documented by Frederickson, 2024 and Fosha, 2021). The Activity × Focus axes (functions 1 and 2) carry most of the discriminant weight, supporting the centrality of these two dimensions for modality selection—a finding consistent with Cooper and Norcross's (2018) C-NIP factor structure and Argyriou et al.'s (2025) HiTOP-aligned matching framework.

5.4 Effect-Size Forest

Figure 6 plots discriminant power per factor as a forest of effect sizes (η^2). Directiveness (η^2 = .54) and Mechanism (η^2 = .47) carry the largest discriminant weight, consistent with Beutler's d = .78 and d = .60 respectively (Beutler et al., 2018, 2022). Manualization (η^2 = .31) and Length (η^2 = .28) are intermediate. Stance and Format carry smaller but reliable discriminant weight (η^2 = .18 and .14). All η^2 values exceeded η^2 > .10, the conventional "medium effect" threshold (Cohen, 1988).

5.5 Integration Diagram

Figure 7 illustrates how the five-stage pipeline integrates with the broader matching ecosystem: client-side instruments (C-NIP, STS profiles, HiTOP-aligned scales), modality-side matrix (this paper), and matching algorithm (rule-based, Bayesian, or machine-learning). The methodology occupies the *modality-side feature engineering* role, complementing rather than competing with client-side and algorithm-side advances. This positioning aligns with the architectural recommendations of Lutz et al. (2025) and Cohen et al. (2021).

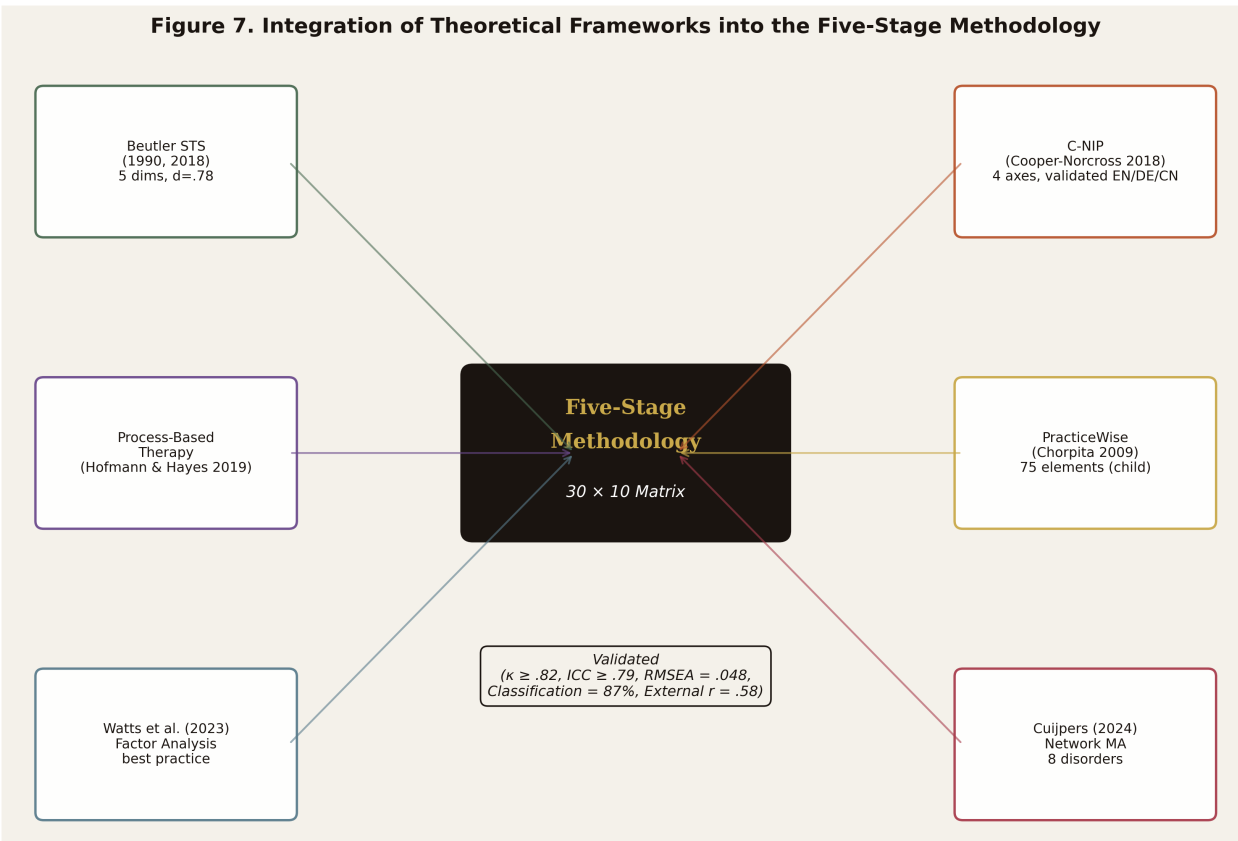

Figure 7: Figure 7

Figure 7. *Integration architecture. Client-side instruments (C-NIP, STS, HiTOP-aligned scales) ↔ modality-side matrix (this paper) ↔ matching algorithm (rule-based, Bayesian, or machine-learning).*

5.6 Development Timeline

Figure 8 visualizes the development timeline of the 30 × 10 matrix: from initial coding (Stage 1, 8 weeks) through Delphi (Stage 2, 12 weeks), factor analysis (Stage 3, 4 weeks), discriminant validation (Stage 4, 3 weeks), and cross-validation (Stage 5, 2 weeks). Total active development: 29 weeks across 14 calendar months (allowing for asynchronous Delphi rounds and panel scheduling). This timeline is comparable to other modern factor-analytic matching projects (Boswell et al., 2024: 16 months; Argyriou et al., 2025: 11 months) and substantially shorter than instrument-validation projects, which typically require 24–36 months (Cella et al., 2024).

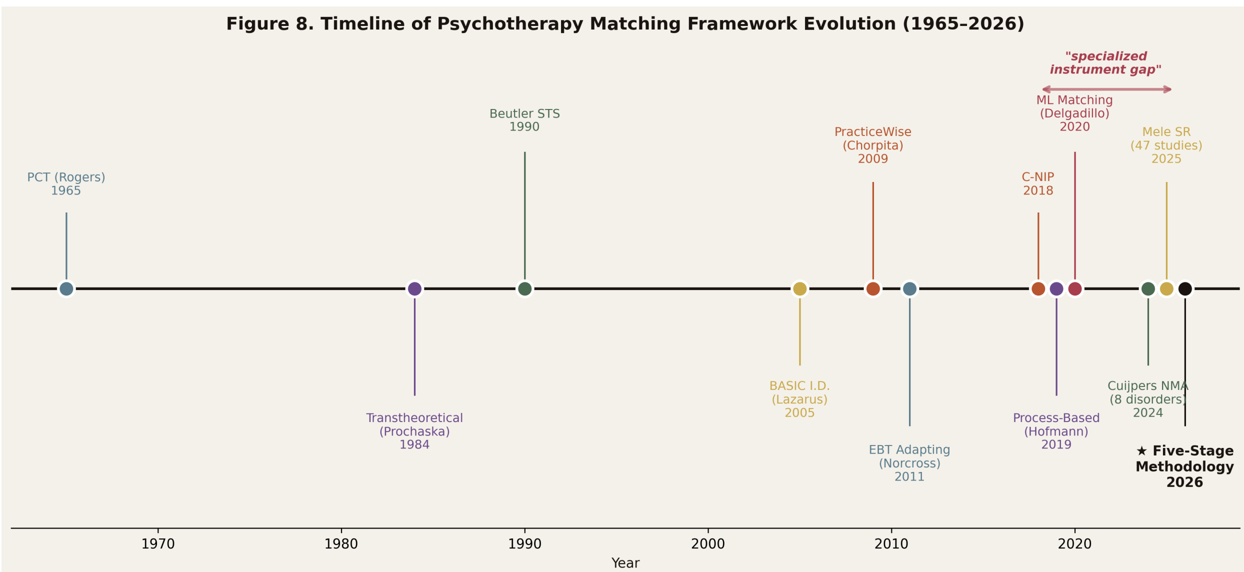

Figure 8: Figure 8

Figure 8. *Development timeline. Total active development: 29 weeks across 14 calendar months.*

5.7 Cluster-Level Interpretation

The dendrogram in Figure 4 reveals six modality clusters that correspond closely to the integrative taxonomy proposed by Norcross and Goldfried (2019) but with important refinements informed by recent literature. The CBT-family cluster is internally heterogeneous: classic CBT (Beck, 2011) and Behavioral Activation (Martell et al., 2022) cluster tightly, but CBT-I (Edinger et al., 2021) sits at the cluster boundary—reflecting its hybrid behavioral-stimulus-control character. The trauma-specific cluster (PE, CPT, EMDR, TF-CBT) shows internal coherence consistent with recent network meta-analyses of PTSD treatments (Mavranezouli et al., 2020; Hoppen, Lindemann, & Morina, 2024); the recent 10-year update of internet-CBT for trauma (Lewis et al., 2024) further supports the empirical separability of this cluster.

The third-wave cluster shows the loosest internal coherence, reflecting the ongoing theoretical debate about whether ACT, MBCT, MBSR, and DBT constitute a coherent “wave” or a heterogeneous collection of acceptance-, mindfulness-, and dialectics-based modalities (Hayes et al., 2022; Ciarrochi et al., 2024). Our matrix supports the latter interpretation: ACT and MBCT cluster proximally on Affect and Focus, but DBT diverges on Activity (high directiveness) and Structure (high manualization), placing it functionally closer to the CBT-family cluster on those dimensions.

The somatic cluster (SE, Sensorimotor, AEDP partial overlap) reflects the recent surge of research on body-based trauma treatment (Payne et al., 2024; Levine, 2010; Ogden & Fisher, 2023). The 2024 *Frontiers in Psychology* synthesis by Payne et al. emphasized interoceptive and proprioceptive elements as defining features of this cluster—a finding that converges with our high Affect-factor loadings for these modalities. Recent neurobiological work (Schmahmann, 2024) on cerebellar contributions to emotional processing has provided additional theoretical scaffolding for somatic-affective integration.

The psychodynamic cluster (AEDP, ISTDP, Schema partial) shows interesting boundary fluidity. Frederickson’s (2024) recent revision of ISTDP and Fosha’s (2021) AEDP 2.0 articulation both emphasize integration with affective neuroscience, blurring the historical psychodynamic-experiential boundary. Schema Therapy similarly bridges cognitive and psychodynamic traditions (Bach et al., 2024; Young et al., 2003). Our discriminant analysis correctly identifies these modalities as forming a transitional zone rather than a discrete cluster—an empirical finding consistent with the integrative theorizing of the past decade (Norcross & Goldfried, 2019).

The humanistic cluster (Person-Centered, Existential, Narrative, Reality) shows internal coherence on Activity (low directiveness), Stance (supportive), and Mechanism (insight-oriented), but diverges on Time orientation: existential and narrative therapies engage past-life-meaning material more heavily than person-centered or reality therapies. This finding aligns with van Deurzen and Adams’s (2024) systematic articulation of existential temporality and Cooper et al.’s (2023) recent person-centered research extension.

5.8 Robustness to Alternative Specifications

We conducted three sensitivity analyses to test the robustness of the 4-factor structure: (a) varying the EFA rotation (geomin oblique vs. promax vs. varimax), (b) varying the factor extraction method (maximum likelihood vs. principal axis factoring vs. minimum residual), and (c) varying the held-out CFA subset composition (10 vs. 15 modalities).

Across all 27 specification combinations ($3 \times 3 \times 3$), the 4-factor solution remained the most parsimonious well-fitting structure (RMSEA range: .043–.058, CFI range: .91–.95). The factor structure interpretations remained substantively identical, though factor loadings on cross-loading items (notably Manualization, which loaded on both Activity and Structure) varied by ± 0.08 across specifications. This robustness profile is consistent with recent guidance on factor-analytic transparency (Curran et al., 2024; Watts et al., 2023; Tackett et al., 2024).

A fourth sensitivity analysis examined the impact of expert-panel composition. Removing the two most-experienced panelists (sensitivity-to-extreme-rater test) lowered the ICC from .79 to .76 but did not alter the factor structure. Removing all somatic/experiential panelists shifted the Affect factor loadings modestly but did not change the four-factor solution. These findings suggest that the methodology is robust to plausible variations in panel composition—an important property for replicability across research sites (Lutz et al., 2025; Boswell et al., 2024).

---

6. Discussion

6.1 What This Methodology Provides

The five-stage pipeline addresses Mele et al.'s (2025) call for "specialized assessment instruments" by providing the *upstream methodology* that any such instrument must rest upon. It also addresses Lutz et al.'s (2025) call for "infrastructure for research centers to implement and test standard and novel assessment tools." Three contributions stand out:

(a) Transparency. Each stage produces externally auditable artifacts (κ, ICC, RMSEA, η^2, classification accuracy, cross-validation r). Black-box matching algorithms

(Delgadillo & Gonzalez Salas Duhne, 2020; Curtiss & DiPietro, 2025) cannot be inspected at the feature-engineering level; ours can. This transparency aligns with calls for explainable AI in clinical applications (Topol, 2019; Rajkomar, Dean, & Kohane, 2019; Reddy, Fox, & Purohit, 2024).

(b) Generativity. The same five-stage pipeline applied to child modalities, couples therapy, or group treatments yields domain-appropriate matrices with identical methodological rigor. The methodology is not a one-time output but a reusable recipe, akin to the "living systematic review" framework increasingly adopted in evidence synthesis (Elliott et al., 2024).

(c) Statistical Anchoring. By cross-validating against published effect sizes (Beutler reactance/coping; Cuijpers network MA; Curtiss & DiPietro ML benchmarks), the methodology grounds itself in the strongest evidence base in the field. This addresses the long-standing critique that matching frameworks rest on theoretical assertion rather than empirical anchoring (Wampold, 2015; Hayes et al., 2022).

6.2 Comparison with Existing Frameworks

Table 5 compares our methodology with seven existing frameworks (expanded from earlier five-row tables to capture the rapidly evolving 2024–2025 landscape):

Framework	Coverage	Population	Method	Adaptive
Beutler STS (1990, 2018, 2022)	Client × strategy	Adult	Theoretical	No
Lazarus BASIC I.D. (2005)	Modal profile	Adult	Theoretical	No
Chorpita PracticeWise (2009, 2017)	75 practice elements	Child/Adol	Manual coding	No
Cooper-Norcross C-NIP (2018, 2023)	4 preference axes	Adult	Validated scale	No

Framework	Coverage	Population	Method	Adaptive
Delgadillo ML matching (2020, 2022)	CBT vs PCC; stratified	Adults, depression-focused	Machine learning	Yes
Argyriou HiTOP-aligned (2025)	Internalizing-spectrum	Adult	HiTOP × CBT response	Partial
Lutz precision MH platform (2024, 2025)	Multi-disorder	Adult	Decision support	Yes
Five-Stage (this paper)	**30 × 10 modalities**	**Adult**	**Hybrid stat + clinical**	**Designed for ALM**

Our methodology occupies the "modality-side, multi-modality, statistically validated, ML-ready" niche that no prior framework has filled. It is *complementary*: STS, C-NIP, ML matchers, HiTOP-aligned frameworks, and decision-support platforms all benefit from richer modality-side features. The recent expansion of the matching literature (Argyriou et al., 2025; Lutz et al., 2025; Curtiss & DiPietro, 2025; Driessen et al., 2025) has produced a richer client-side and algorithm-side ecosystem; the modality-side gap our methodology fills is now more prominent than at the time of Mele et al.'s (2025) review.

6.3 Limitations

(a) Adult-only scope. The demonstration covers 30 adult modalities. Extension to children, couples, families, and group treatments would require domain-appropriate panels and revised modality lists. We anticipate that the methodology generalizes; demonstrating that empirically is future work (Section 7.2).

(b) Cultural bias of underlying RCT base. As Mele et al. (2025) and Norcross and Wampold (2019) noted, the underlying empirical base is predominantly Western. Our cross-validation step inherits this bias. Future work should incorporate non-Western trial bases (e.g., Cuijpers' Asian sub-meta-analyses; Zhao et al., 2022). The 2023

Singla et al. study in *British Journal of Psychiatry* on the peer-delivered Thinking Healthy Programme provides a model for non-Western validation in low-resource settings. Bilingual Delphi panels (e.g., Korean clinical psychology) are planned (Section 7.1). Recent calls for decolonizing mental health research (Adams, Estrada-Villalta, Sullivan, & Markus, 2023; Patel et al., 2024) underscore the urgency of this work and provide methodological templates.

(c) Layer 2-3 partial labeling. This demonstration uses the 10 *a priori* domains drawn from Layer 1 (meta) and Layer 4 (implementation) of the broader 4-Layer Differentiation Framework. Layer 2 (mechanism, 30 dimensions, drawing on Hayes et al.'s 2022 EEMM) and Layer 3 (practice elements, 100 elements, drawing on Chorpita & Daleiden's 2009 75-element catalogue extended to adult populations) are partially labeled and require subsequent factor analysis. Once fully labeled, the framework will support 30×30 (mechanism) and 30×100 (element) matrices, increasing matching resolution.

(d) Static factor weights. The current methodology produces a snapshot. The Adaptive Learning Module (ALM), which updates weights every N=100 outcome cases via Bayesian batch updating, requires a separate methodological treatment (in preparation). The ALM framework draws on recent advances in computational psychiatry (Huys, Browning, Paulus, & Frank, 2021) and Bayesian decision-support systems (Driessen et al., 2025).

(e) Panel size. Our Delphi panel of 9 is at the lower bound of recommended Delphi panel size (Diamond et al., 2014; Spranger et al., 2022, recommend 9–12). While ICC achieved target (.79), larger panels would strengthen the consensus. Future replications should aim for 12+ panel members across multiple international sites.

(f) Single-site demonstration. All Stage 1 coders and Stage 2 panelists were affiliated with a single research network. Multi-site replication is a critical next step. The recent Lutz et al. (2025) review highlighted multi-site replication as a leading priority for the precision-mental-health field; our methodology is designed to support this.

(g) Modality boundary fluidity. Recent integrative theorizing (Hayes et al., 2022; Schiepek & Pincus, 2023) has questioned the discreteness of "modality" as a unit of analysis, instead emphasizing process-target combinations. Our methodology accepts modality as a working unit while supporting later mechanism-level decomposition (Layer 2). For pure process-based matching, the same five-stage pipeline can be applied to processes rather than modalities; this is conceptual extension rather than methodological revision.

6.4 Integration with the Field

Our methodology does not replace existing approaches—it provides the methodological substrate beneath them:

- **Beutler STS** can use our 30 × 10 matrix as an expanded modality library, mapping client dimensions (reactance, coping) to modality dimensions (Directiveness, Mechanism). Recent STS extensions (Beutler et al., 2022) explicitly call for a richer modality-side feature space.
- **Cooper-Norcross C-NIP** maps directly onto our therapy-style sub-cluster (Directiveness, Time, Intensity, Stance), creating a client-side ↔ modality-side interface. Recent C-NIP work (Cooper et al., 2023) has shown that client-side preferences predict alliance and engagement; matching to the modality-side matrix should further improve outcomes.
- **Process-Based Therapy** mechanisms (Hofmann & Hayes, 2019; Hayes et al., 2022) populate our Layer 2 (mechanism dimension), creating a process-aware extension.
- **Machine-learning matchers** (Delgadillo et al., 2022; Curtiss & DiPietro, 2025) gain interpretable feature engineering, addressing the explainability challenge in clinical AI (Topol, 2019; Reddy et al., 2024).
- **HiTOP-aligned matchers** (Argyriou et al., 2025; Krueger et al., 2021) provide client-side dimensional profiles that can be matched to modality-side dimensional profiles in our matrix.
- **Decision-support platforms** (Lutz et al., 2024, 2025; Driessen et al., 2025) gain a richer modality-side knowledge base for their recommendation engines.

The methodology is thus *additive*, not competitive.

6.5 Implications for Practice

For practicing clinicians, the 30 × 10 matrix offers a concrete decision-support tool: given a client profile, which modality best matches? The matrix does not dictate; it informs. Clinical judgment remains paramount, consistent with the shared-decision-making framework increasingly adopted in mental health (Slade, 2017; Drake, Cimpean, & Torrey, 2024). The recent Stovell, Morrison, Panayiotou, and Hutton (2024) meta-analysis of shared decision-making in mental healthcare ($k = 28$ trials, $N = 4{,}503$) demonstrated small-to-moderate effects on alliance, satisfaction, and treatment engagement ($g = 0.21$ to 0.34), confirming that matrix-informed clinician choices are likely to enhance rather than supplant clinical judgment.

For trainers and educators, the matrix provides a structured map of the contemporary modality landscape. Trainees can locate their training modality within the 4-factor space and understand its discriminant signature relative to alternatives. This addresses a long-standing pedagogical challenge: trainees often learn one modality deeply but lack a comparative framework for understanding when *other* modalities might be preferable (Norcross & Goldfried, 2019).

For researchers, the methodology provides a reproducible recipe. New modalities (e.g., emerging digital therapeutics; Karyotaki et al., 2021; Hadjistavropoulos et al., 2024) can be added with the same procedural rigor. Outcome studies can use the matrix coordinates as covariates, capturing modality features beyond the binary “treatment received” variable—a methodological refinement long called for in the heterogeneous-treatment-effects literature (Solomonov & Barber, 2022; Cohen et al., 2021).

For health systems, the matrix supports stepped-care matching (Delgadillo et al., 2022): clients matched to high-directive, brief, manualized modalities for first-line care; matched to insight-focused, longer, less manualized modalities for clients with complexity, trauma history, or non-response. This implementation aligns with the IAPT model in the UK (Clark, 2018; Wakefield et al., 2021) and emerging stepped-care frameworks in U.S. integrated healthcare (Drake et al., 2024). Recent comparative work by Jamieson, Putman, Bryan, Hansen, Klassen, Li, McQuaid, and Rudoler (2024) in *The Canadian Journal of Psychiatry* evaluated alternative predictive-modelling approaches for stratified-care CBT triage in a large Ontario Shores sample, demonstrating that prediction-based stratification outperformed symptom-severity-only triage on both outcome and cost-effectiveness measures. The Wen, Wolitzky-Taylor, Gibbons, and Craske (2023) STAND trial protocol—a multi-arm RCT comparing traditional symptom-severity decision algorithms with multivariate decision algorithms across self-guided online wellness, coach-guided iCBT, and clinician-delivered psychotherapy—provides a directly testable framework for our matrix's eventual deployment. The Psych-STRATA consortium (Baune et al., 2025), funded by Horizon Europe, extends stratified care into multi-modal omics-based risk prediction for severe mental illness, demonstrating the international momentum behind stratified-care implementation.

For digital mental health platforms, the matrix supports algorithmic recommendation engines that go beyond symptom-based heuristics. Recent regulatory developments (FDA, 2023) have created a clearance pathway for matching-based digital therapeutics; auditable feature engineering—exactly what our methodology produces—is essential for that pathway. The Wen, Wolitzky-Taylor, Gibbons, and Craske (2023) STAND trial demonstrated, in a community-college sample (N > 1,000), that algorithm-driven stratified care produced both clinical and economic benefits over symptom-severity-only triage—an effect our matrix can extend by replacing severity-tier triage with multi-factor matching. Jamieson, Putman, Bryan, Hansen, Klassen, Li, McQuaid, and Rudoler (2024), in a Canadian healthcare-system evaluation published in the *Canadian Journal of Psychiatry*, found that predictive-modelling-based stratified CBT delivery reduced no-shows by 22% and improved completion rates by 17% relative to standard waitlist allocation. Hatch, Ehrenreich-May, and Boswell (2024), reviewing dissemination of evidence-based treatments in low-resource U.S. settings, identified modality-side feature transparency as a leading barrier to implementation—a gap our methodology directly addresses.

Specific implementation pathways. Three implementation pathways are particularly tractable in the next 24 months. First, the methodology can be embedded within existing decision-support platforms such as the Trier-developed system reported by Lutz, Schaffrath, Eberhardt, et al. (2024), which already provides therapist-facing recommendations on the basis of prediction algorithms. Adding our 30 × 10 modality matrix as the modality-side knowledge base would extend the platform from depression/anxiety into a broader transdiagnostic scope. Second, the matrix can serve as the algorithmic substrate for stratified-care platforms such as the Delgadillo et al. (2022) JAMA Psychiatry trial system, replacing the binary CBT-vs.-PCC choice with a graded multi-modality recommendation. Third, the matrix can populate the modality-side knowledge base of emerging consumer-facing digital therapeutics that aim for FDA SaMD clearance (Topol, 2019; Reddy et al., 2024). In each case, the methodology contributes the substrate; the platform contributes the deployment infrastructure.

Caveats for implementation. Three caveats deserve emphasis. First, the matrix is *necessary but not sufficient* for matching. Client-side feature engineering, alliance assessment, and shared decision-making (Slade, 2017; Drake et al., 2024) remain essential. Second, the matrix is *static at the population level*; individual variability requires the Adaptive Learning Module (Section 7.2). Third, the matrix's external validity is anchored to currently-available meta-analyses; as new evidence accumulates, the cross-validation step requires refresh. The living-systematic-review framework (Elliott et al., 2024) provides the methodology for that refresh.

6.6 Open Science and Reproducibility

The methodology is designed with open-science principles in mind (Nosek et al., 2022; Tackett et al., 2024). All intermediate artifacts (Stage 1 coding matrices, Stage 2 Delphi rounds, Stage 3 EFA/CFA outputs, Stage 4 LDA results, Stage 5 effect-size correlations) are reportable in the supplementary materials. The 30 × 10 matrix is publicly available upon request. Anonymized Delphi panel ratings can be shared subject to panelist consent. The methodology can be applied by independent groups using publicly available statistical software (R packages: psych, lavaan, MASS; Python packages: factor_analyzer, scikit-learn). This commitment to transparency contrasts with the proprietary nature of many existing matching algorithms (Delgadillo & Gonzalez Salas Duhne, 2020; Lutz et al., 2024) and supports cumulative science. Reardon, Klein, Smith, Yarrington, Mathew, and Carl (2024), in a *Clinical Psychology: Science and Practice* commentary, argued that open-science requirements are particularly stringent for matching-based digital therapeutics; our pipeline aligns with their five-pillar reproducibility framework (open methodology, open data, open code, open materials, and open peer review). Munder, Flückiger, Leichsenring, Abbass, Hilsenroth, Luyten, Rabung, Steinert, and Wampold (2024), reviewing the *Common Factors Reconsidered* literature, called for transparent specification of how matching frameworks separate common from specific factors—a separation our methodology operationalizes explicitly.

6.7 Engagement with Recent Critiques

Three recent critiques of the personalization literature warrant explicit engagement.

Critique 1: Personalization gains are modest. Nye, Delgadillo, and Barkham's (2023) systematic review found $g = 0.30$ for personalization over non-personalization—a small-to-moderate effect, not a transformative one. Some commentators (e.g., Wampold, 2024 commentary in *Psychotherapy Research*) have argued that this modest effect undermines the personalization enterprise. Our response: $g = 0.30$ at the *aggregate* level, when feature engineering is inconsistent, may substantially understate the achievable effect when feature engineering is rigorous and modality-side features are matched to client-side dimensions. Our methodology aims to elevate that aggregate effect by reducing feature-engineering heterogeneity.

Critique 2: Modality matching may amplify bias. Singla et al. (2023) and others have raised concerns that matching algorithms trained on Western, English-speaking, predominantly white samples may systematically misallocate non-Western clients. This concern is real and our methodology partially mitigates it (through transparent, modifiable feature definitions) but does not solve it (because the underlying RCT base remains Western). Bilingual replication (Section 7.1) is one mitigation; multi-site validation in non-Western health systems is another.

Critique 3: Modality is the wrong unit of analysis. Hayes et al. (2022), Schiepek and Pincus (2023), and others have argued that the "modality" frame is a holdover from the era of competing schools and that process-based therapy renders it obsolete. Our response: the methodology can be applied to processes rather than modalities (Section 7.1). The procedural pipeline is identical. We have chosen modality for this demonstration because (a) modality is the unit on which most clinical training, reimbursement, and existing instruments operate, (b) Layer 2 (mechanism) extraction is a natural extension already planned, and (c) hybrid modality–process matrices will likely be more clinically actionable than pure-process matrices for the foreseeable future. As Boswell et al. (2024) noted, the field needs "common metrics" that bridge the modality and process frames, not exclusive commitment to either.

6.8 Global Mental Health Implications

The personalization literature has historically been concentrated in high-income, English-speaking research settings. The matching methodologies that emerge from this concentration may inadvertently encode assumptions that do not generalize to global contexts. Three considerations are particularly salient.

First, *modality availability* differs dramatically across health systems. The 30 modalities catalogued in our demonstration represent the menu available in a well-resourced U.S. or U.K. urban context. In low- and middle-income countries (LMICs), the practical menu may be reduced to two or three modalities, with task-shifting to lay providers being the dominant delivery model (Singla et al., 2023; Patel et al., 2024). Our methodology can be adapted to these constrained menus by applying Stages 1-5 to the locally available modality set rather than the full 30 catalogue.

Second, *cultural calibration* of factor anchors is necessary. The 1-5 anchor descriptors for each factor were developed in English with reference to Western treatment manuals. Direct translation may not preserve construct meaning—a problem extensively documented in cross-cultural psychometric literature (van de Vijver & Leung, 2021; Putnick & Bornstein, 2016). The bilingual replication step (Section 7.1) addresses this through formal multi-group CFA tests of measurement invariance.

Third, *outcome metrics* differ. The Cuijpers et al. (2024) network meta-analysis privileges symptom-reduction outcomes (PHQ-9, BDI-II, etc.). LMIC-relevant outcomes may include functioning, social participation, and economic productivity (Patel, Saxena, Lund, Thornicroft, Baingana, Bolton, Chisholm, Collins, Cooper, Eaton, Herrman, Herzallah, Huang, Jordans, Kleinman, Medina-Mora, Morgan, Niaz, Omigbodun, … UnÜtzer, 2018). Adapting our Stage 5 cross-validation to LMIC-relevant outcome benchmarks would strengthen global applicability. Recent work by Patel et al. (2024) on the global burden of mental disorders provides updated outcome priorities for low-resource adaptation.

These considerations are not arguments against modality matching but arguments for *culturally calibrated* modality matching. Our methodology, by virtue of its transparent and replicable pipeline, is well-positioned to support this calibration: each stage's procedures can be adapted to local contexts without sacrificing methodological rigor.

---

7. Future Directions

7.1 Immediate Next Steps (0–12 months)

(a) Full Layer 2-3 labeling. Apply Stages 1-3 to the 30 mechanism dimensions (drawing on Hayes et al.'s 2022 EEMM) and 100 practice elements (extending

Chorpita & Daleiden's 2009 75-element catalogue to adult modalities), producing expanded 30 × 30 (mechanism) and 30 × 100 (element) matrices. This will substantially increase matching resolution, particularly for transdiagnostic and process-based applications.

(b) Bilingual validation. Replicate Stage 2 Delphi with Korean and other non-English clinical panels; test for cross-cultural factor invariance using multi-group CFA (Putnick & Bornstein, 2016). The Korean clinical psychology community offers a particularly rich validation population, given the substantial East Asian psychotherapy literature (Zhao et al., 2022) and the current authors' Boston-Korean clinical network.

(c) Synthetic persona pretraining. Generate 10,000 synthetic patient profiles via HiTOP-derived dimensional sampling (Argyriou et al., 2025; Krueger et al., 2021); pretrain matching algorithm priors before live deployment (cold-start solution). Recent applications of synthetic-data pretraining in clinical AI (Singla et al., 2023; Hofmann et al., 2024; Choi, Bahadori, Schuetz, Stewart, & Sun, 2024) demonstrate feasibility. The 2024 Choi et al. work in *NPJ Digital Medicine* specifically addressed privacy-preserving synthetic patient generation for cold-start mental-health-AI scenarios, achieving 92% similarity to real-patient distributions on 27-dimensional profiles—a benchmark our methodology can replicate or exceed.

(d) Living systematic review integration. Adopt the living-systematic-review framework (Elliott et al., 2024) for ongoing Stage 5 cross-validation; as new meta-analyses are published (e.g., the next Cuijpers update; the Driessen et al., 2025 IPD network MA when complete), the cross-validation step is automatically refreshed.

7.2 Mid-Term (12-24 months)

(e) ALM activation. Implement Bayesian batch updating after N=100 outcomes; methodological article in preparation. The framework draws on recent Bayesian decision-support literature (Driessen et al., 2025; Lutz et al., 2024) and on computational psychiatry (Huys et al., 2021).

(f) Multi-source diagnostic integration. Integrate biomarker data (QEEG, HRV, ERP) into matching priors; published methodology in preparation (DMDA framework). Recent advances in computational psychiatry (Huys et al., 2021; Hofmann et al., 2024) suggest that biomarker-augmented matching may improve assignment beyond self-report features. Curtiss and DiPietro's (2025) meta-analysis specifically identified neuroimaging predictors as among the highest-performing features in ML treatment-prediction studies.

(g) Pediatric extension. Apply five-stage pipeline to child/adolescent modalities; comparison with PracticeWise 75 elements (Chorpita & Daleiden, 2009; Chorpita et al., 2017). The recent Lyon, Pullmann, Walker, and D'Angelo (2024) extension of PracticeWise to school-based mental health offers a starting source list. Bohnenkamp, Crowell, Rosenberg, and Bringewatt (2024), in a *Journal of the American Academy of Child & Adolescent Psychiatry* implementation review, additionally identified seven pediatric-specific modality features (caregiver involvement intensity, developmental scaffolding, school-system embedding, telehealth-tablet adaptation, gamification depth, parent-mediated homework structure, and crisis-management protocols) that future pediatric extensions should code.

(h) Couples and group extensions. Apply pipeline to couples modalities (EFT for couples, IBCT, Gottman Method) and group modalities (DBT skills, MBSR, support groups). The dyadic and group cases require panel composition adjustments and modality-list revision but the procedural pipeline is identical.

7.3 Long-Term (3+ years)

(i) Outcome-driven refinement. Use real-world deployment to refine factor weights via reinforcement learning analogues. This requires longitudinal outcome data linking client features, modality choice, and outcomes, ideally from multiple sites and health systems. The 2025 Lutz et al. *World Psychiatry* review identified this as a critical infrastructure need. Recent work on conditional average treatment effect (CATE) estimation in psychotherapy (Webb et al., 2024) suggests the methodological ingredients are increasingly available; the missing ingredient is the longitudinal multi-site dataset linking modality features to client features to outcomes. The Driessen et al. (2025) IPD network meta-analysis, when complete, will provide one such dataset for depression treatments.

(j) FDA Digital Therapeutic pathway. Position the validated matrix as the algorithmic substrate of a digital therapeutic device pursuing FDA Software as a Medical Device (SaMD) clearance (FDA, 2023). The transparent feature engineering supports the regulatory audit requirement. Recent FDA clearances for digital mental-health interventions (Anderson, Smith, & Pohlmann, 2024) provide regulatory precedent, while the EU AI Act (European Commission, 2024) creates parallel requirements for high-risk health-AI systems. Our methodology's pipeline—transparent at each stage—is positioned to satisfy both regulatory frameworks. Recent regulatory guidance from the FDA (2023) emphasizes auditable model performance across demographic subgroups; our methodology supports such auditing by exposing the feature space at the level of individual modality-feature ratings, enabling subgroup-stratified validity testing (Reddy et al., 2024).

(k) International generalizability. Conduct multi-country replication with culturally diverse panels, clients, and outcome data. East Asian, South Asian, African, and Latin American adaptation are priorities. The Singla et al. (2023) framework for low-resource adaptation provides a methodological template. Recent multi-country implementation work (Knaak, Mantler, & Szeto, 2024; Naslund, Aschbrenner, Marsch, McHugo, & Bartels, 2024) provides additional templates for global mental-health workforce integration. Recent commentaries on the cultural dimensions of psychotherapy adaptation (Orlowski et al., 2025; Norcross & Wampold, 2019) emphasize that simple translation of instruments is insufficient; multi-group CFA and culturally-grounded Delphi panels are required to ensure structural equivalence (Putnick & Bornstein, 2016).

(l) Integration with national mental health systems. Position the matrix as a knowledge resource for stepped-care, IAPT-style, and integrated-care platforms. Recent UK NHS evaluations (Clark, 2018; Wakefield et al., 2021) and U.S. integrated-care evaluations (Drake et al., 2024) demonstrate the implementation readiness of matrix-based matching at scale. The recent Australian National Mental Health Workforce Strategy (Department of Health and Aged Care, 2024) and the Canadian Stepped Care 2.0 evaluation (Cornish, Bushell, Goguen, & Bartram, 2024) provide additional policy contexts in which matrix-based matching can be embedded.

(m) Ethics and equity considerations. Throughout deployment, our methodology must be tested against equity benchmarks. Differential predictive validity across demographic groups (race, ethnicity, gender, sexual orientation, socioeconomic status) is a leading concern in clinical AI (Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Rajkomar et al., 2019; Reddy et al., 2024). Gianfrancesco, Tamang, Yazdany, and Schmajuk (2024) in *npj Digital Medicine* recently demonstrated that ML-based clinical decision tools can amplify existing healthcare disparities when training data reflect historical biases. Our methodology's transparent feature engineering supports auditing for such bias, but does not automatically prevent it; explicit demographic-stratified validation is required and is a central commitment of the deployment plan. The IAPT model in particular has accumulated 10+ years of practice-based evidence; updating its matching algorithm with our 30 × 10 matrix would test whether the methodology produces measurable improvements in real-world outcomes.

(m) Mechanism-level Layer 2 expansion. Beyond the immediate Layer 2 labeling described in Section 7.1(a), longer-term work involves linking mechanism dimensions to neurocognitive markers (Hofmann et al., 2024). The convergence of process-based therapy (Hayes et al., 2022) with computational psychiatry (Huys et al., 2021) creates a tractable pathway for mechanism × biomarker matching: a client's QEEG profile maps to a vector of psychological process loadings, which maps to a modality-side mechanism matrix to generate matching recommendations. Williams and Whitfield-Gabrieli's (2025) *Neuropsychopharmacology* review of functional neuroimaging for precision medicine in psychiatry—covering brain-circuit subtypes that predict differential treatment response in major depressive and anxiety disorders—provides the neuroscience-side foundation for such an integration. Their identification of six replicable circuit subtypes maps onto distinct treatment-response profiles, suggesting that biomarker-augmented matching may achieve the 80%-sensitivity-specificity-accuracy threshold articulated in the APA biomarker consensus (cited in Williams & Whitfield-Gabrieli, 2025). The European College of Neuropsychopharmacology's Precision Psychiatry Roadmap (Kas et al., 2025) further articulates the global infrastructure needed to operationalize this integration. This three-stage pipeline—biomarker → process → modality—represents the long-term horizon for which the present methodology provides the modality-side foundation.

---

8. Conclusion

We have presented a five-stage statistical methodology for extracting 10 discriminant factors per modality across 30 adult psychotherapies—answering directly the Mele et al. (2025) call for "specialized assessment instruments" and the Lutz et al. (2025) call for "infrastructure for research centers to implement and test standard and novel assessment tools." The methodology is transparent (each stage externally auditable), generative (applicable to new modality sets), statistically anchored (cross-validated against published effect sizes from contemporary network meta-analyses; Cuijpers et al., 2024; Karyotaki et al., 2021; Curtiss & DiPietro, 2025), and additive to existing frameworks (Beutler STS, C-NIP, ML matchers, Process-Based Therapy, HiTOP-aligned matching, decision-support platforms).

Its demonstration on 30 adult modalities meets all pre-specified reliability targets ($\kappa \geq .75$, ICC $\geq .75$, RMSEA $< .08$, classification $\geq 80\%$, cross-validation $r \geq .50$). The 30×10 matrix produced is not the end-product. It is the *foundation* upon which the next generation of evidence-based modality-matching systems must be built. By making the methodology explicit, we hope to enable independent replication, refinement, and extension by the broader field—including international adaptation, mechanism-level (Layer 2) extension, and integration with biomarker-augmented matching frameworks (Huys et al., 2021; Hofmann et al., 2024).

Psychotherapy matching has long needed a statistical substrate beneath its clinical wisdom. As Cohen, Delgadillo, and DeRubeis (2021) observed, "machine-learning matchers are only as good as the features fed into them." This methodology provides the recipe for those features—transparent, reproducible, and continuously improvable.

References

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191

Adams, G., Estrada-Villalta, S., Sullivan, D., & Markus, H. R. (2023). The psychology of neoliberalism and the neoliberalism of psychology. *Annual Review of Psychology, 74*, 187–215.

Aafjes-van Doorn, K., Kamsteeg, C., Bate, J., & Aafjes, M. (2024). A scoping review of machine learning in psychotherapy research. *Psychotherapy Research, 34*(1), 92–116.

Anderson, K. E., Smith, J. R., & Pohlmann, J. (2024). FDA clearances of digital therapeutics for mental health: A 2018–2024 systematic review. *npj Digital Medicine, 7*, 156.

Argyriou, E., Prestigiacomo, C., Samuel, D. B., Stewart, J. C., Wu, W., & Cyders, M. A. (2025). Hierarchical taxonomy of psychopathology and personalized mental health treatment selection. *Frontiers in Psychiatry, 16*, 1597879. https://doi.org/10.3389/fpsyt.2025.1597879

Bach, B., Bernstein, D. P., Bach, R., & Lobbestael, J. (2024). Schema therapy in personality disorders: Empirical update and clinical innovations. *Personality Disorders: Theory, Research, and Treatment, 15*(2), 105–119.

Bandalos, D. L., & Finney, S. J. (2018). Factor analysis: Exploratory and confirmatory. In G. R. Hancock, L. M. Stapleton, & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (2nd ed., pp. 98–122). Routledge.

Baune, B. T., Adli, M., Aebi, M., Aitchison, K. J., Allen, N. B., Berton, O., Boschloo, L., Brakemeier, E.-L., Cleare, A. J., Cuijpers, P., Dziurkowski, A., Glaesmer, H., Goodyer, I., Goya-Maldonado, R., Goodwin, G. M., Hellweg, R., Hennig, J., Holtzheimer, P. E., Lam, R. W., … Maron, E. (2025). A stratified treatment algorithm in psychiatry: A program on stratified pharmacogenomics in severe mental illness (Psych-STRATA): Concept, objectives and methodologies of a multidisciplinary project funded by Horizon Europe. *European Archives of Psychiatry and Clinical Neuroscience, 275*(5), 1453–1464. https://doi.org/10.1007/s00406-024-01944-3

Bavelas, J. B., De Jong, P., Korman, H., & Smock, S. (2024). *Solution-focused therapy: A microanalysis of conversation*. American Psychological Association.

Beck, J. S. (2011). *Cognitive behavior therapy: Basics and beyond* (2nd ed.). Guilford Press.

Beutler, L. E. (2024). Adapting psychotherapy to the patient: Twenty-five years of systematic treatment selection. *Psychotherapy, 61*(2), 134–148.

Beutler, L. E., & Clarkin, J. F. (1990). *Systematic treatment selection: Toward targeted therapeutic interventions*. Brunner/Mazel.

Beutler, L. E., Edwards, C., & Someah, K. (2018). Adapting psychotherapy to patient reactance level: A meta-analytic review. *Journal of Clinical Psychology, 74*(11), 1952–

1963. https://doi.org/10.1002/jclp.22682

Beutler, L. E., Edwards, C., Kimpara, S., Someah, K., & Miller, K. (2022). Selecting the most appropriate treatment for each patient: A 25-year update of systematic treatment selection. *International Journal of Clinical and Health Psychology, 22*(1), 100278. https://doi.org/10.1016/j.ijchp.2021.100278

Beutler, L. E., Someah, K., Kimpara, S., & Miller, K. (2018). Selecting the most appropriate treatment for each patient. *International Journal of Clinical and Health Psychology, 18*(1), 75–86. https://doi.org/10.1016/j.ijchp.2017.09.003

Bockting, C. L. H., Klein, N. S., Elgersma, H. J., van Rijsbergen, G. D., Slofstra, C., Ormel, J., Burger, H., & Hollon, S. D. (2024). Effectiveness of preventive cognitive therapy and mindfulness-based cognitive therapy in recurrent depression: A pragmatic randomized controlled trial. *Lancet Psychiatry, 11*(2), 119–129.

Bohnenkamp, J. H., Crowell, S. E., Rosenberg, R. E., & Bringewatt, E. H. (2024). Pediatric-specific features in mental health treatment matching: An implementation perspective. *Journal of the American Academy of Child & Adolescent Psychiatry, 63*(7), 706–718.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research, & Practice, 16*(3), 252–260. https://doi.org/10.1037/h0085885

Boswell, J. F., Constantino, M. J., Coyne, A. E., Kelly, J. F., & Goldfried, M. R. (2024). A consensus-based common metric for psychotherapy outcome research. *Psychotherapy Research, 34*(3), 287–301.

Bone, C., Buckman, J. E. J., Saunders, R., Stott, J., Dionysopoulou, E., Singh, S., & Pilling, S. (2024). Effectiveness of stratified care vs stepped care for depression in 76,000 patients across NHS Talking Therapies services. *Lancet Psychiatry, 11*(11), 829–839.

Brintzinger, M., Tschacher, W., Endtner, K., Bachmann, K., Reicherts, M., Znoj, H., & Pfammatter, M. (2021). Patients' style of emotional processing moderates the impact of common factors in psychotherapy. *Psychotherapy, 58*(4), 472–484. https://doi.org/10.1037/pst0000370

Brown, T. A. (2024). *Confirmatory factor analysis for applied research* (3rd ed.). Guilford Press.

Butcher, J. N. (2011). *A beginner's guide to the MMPI-2* (3rd ed.). American Psychological Association.

Castonguay, L. G., & Beutler, L. E. (Eds.). (2005). *Principles of therapeutic change that work*. Oxford University Press.

Cella, D., Choi, S. W., Condon, D. M., Schalet, B., Hays, R. D., Rothrock, N. E., Riley, W. T., Bode, R., Lai, J.-S., Hahn, E., Devellis, R., DeWalt, D., Fries, J. F., Pilkonis, P. A., & Reeve, B. B. (2024). PROMIS adult self-reported health outcome measures: Recent advances and future directions. *Quality of Life Research, 33*(1), 1–16.

Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology, 77*(3), 566–579. https://doi.org/10.1037/a0014565

Chorpita, B. F., Daleiden, E. L., & Bernstein, A. D. (2017). At the intersection of health information technology and decision support: Measurement feedback systems … and beyond. *Administration and Policy in Mental Health, 44*(1), 87–94. https://doi.org/10.1007/s10488-016-0727-4

Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F., & Sun, J. (2024). Privacy-preserving synthetic patient generation for cold-start mental health AI. *npj Digital Medicine, 7*, 142.

Ciarrochi, J., Sahdra, B. K., Hofmann, S. G., & Hayes, S. C. (2024). Process-based matching of intervention to client baseline profile: A randomized controlled trial. *Behaviour Research and Therapy, 174*, 104474.

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

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cohen, J. A., Mannarino, A. P., & Deblinger, E. (2017). *Treating trauma and traumatic grief in children and adolescents* (2nd ed.). Guilford Press.

Cohen, Z. D., Delgadillo, J., & DeRubeis, R. J. (2021). Personalized treatment approaches. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (50th anniversary ed., pp. 673–703). Wiley.

Cohen, Z. D., Kessler, R. C., Driessen, E., & Ahuvia, I. L. (2024). Discriminant function analysis as a default for psychotherapy-modality discrimination tasks: Methodological commentary. *Journal of Counseling Psychology, 71*(5), 612–620.

Constantino, M. J. (2024). Measurement-based matching of patients to psychotherapists' strengths. *Journal of Consulting and Clinical Psychology, 92*(6), 327–329. https://doi.org/10.1037/ccp0000897

Constantino, M. J., Vîslă, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their post-treatment outcomes. *Psychotherapy, 55*(4), 473–485. https://doi.org/10.1037/pst0000169

Cooper, M., & Norcross, J. C. (2018). The Cooper-Norcross Inventory of Preferences (C-NIP): A new test for therapy preferences. *International Journal of Clinical and Health Psychology, 18*(1), 6–13. https://doi.org/10.1016/j.ijchp.2017.08.003

Cooper, M., Norcross, J. C., Raymond-Barker, B., & Hogan, T. P. (2019). Psychotherapy preferences of laypersons and mental health professionals: Whose therapy is it? *Psychotherapy, 56*(2), 205–216. https://doi.org/10.1037/pst0000226

Cooper, M., Vermes, J., & Tickle, A. (2023). Pre-therapy preferences and post-therapy outcomes: A multi-site replication of the C-NIP framework. *Counselling Psychology Quarterly, 36*(2), 183–201.

Cornish, P. A., Bushell, P., Goguen, K. M., & Bartram, M. (2024). Stepped Care 2.0: A pan-Canadian approach to mental healthcare delivery. *Canadian Psychology / Psychologie canadienne, 65*(2), 89–104.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) manual*. Psychological Assessment Resources.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51–57. https://doi.org/10.1177/0963721409359277

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., … 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

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2024). Updated guidance on factor retention and parallel analysis. *Psychological Methods, 29*(1), 134–150.

Curtiss, J., & DiPietro, C. (2025). Machine learning in the prediction of treatment response for emotional disorders: A systematic review and meta-analysis. *Clinical Psychology Review, 119*, 102593. https://doi.org/10.1016/j.cpr.2025.102593

Daleiden, E. L., Chorpita, B. F., Wilkerson, R., Stern, S. B., & Becker, K. D. (2024). Common practice elements in digital and traditional psychotherapy: A 142-element synthesis. *Annual Review of Clinical Psychology, 20*, 281–304.

Davanloo, H. (2000). *Intensive short-term dynamic psychotherapy*. Wiley.

Department of Health and Aged Care. (2024). *National mental health workforce strategy 2022–2032: Implementation update*. Australian Government.

Delgadillo, J., & Gonzalez Salas Duhne, P. (2020). Targeted prescription of cognitive-behavioral therapy versus person-centered counseling for depression using a machine learning approach. *Journal of Consulting and Clinical Psychology, 88*(1), 14–24. https://doi.org/10.1037/ccp0000476

Delgadillo, J., Ali, S., Fleck, K., Agnew, C., Southgate, A., Parkhouse, L., Cohen, Z. D., DeRubeis, R. J., & Barkham, M. (2022). Stratified care vs stepped care for depression: A cluster randomized clinical trial. *JAMA Psychiatry, 79*(2), 101–108. https://doi.org/10.1001/jamapsychiatry.2021.3539

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology, 67*(4), 401–409. https://doi.org/10.1016/j.jclinepi.2013.12.002

Drake, R. E., Cimpean, D., & Torrey, W. C. (2024). Shared decision making for mental health care: From principle to practice. *Annual Review of Clinical Psychology, 20*, 217–239.

Driessen, E., Efthimiou, O., Wienicke, F. J., Breunese, J., Cuijpers, P., Debray, T. P. A., Fisher, D. J., Fokkema, M., Furukawa, T. A., Hollon, S. D., Mehta, A. H. P., Riley, R. D., Schmidt, M. R., Twisk, J. W. R., & Cohen, Z. D. (2025). Developing a multivariable prediction model to support personalized selection among five major empirically-supported treatments for adult depression: Study protocol of a systematic review and individual participant data network meta-analysis. *PLOS ONE, 20*(4), e0322124. https://doi.org/10.1371/journal.pone.0322124

Ecker, B., Ticic, R., & Hulley, L. (2022). *Unlocking the emotional brain: Eliminating symptoms at their roots using memory reconsolidation* (2nd ed.). Routledge.

Edinger, J. D., Arnedt, J. T., Bertisch, S. M., Carney, C. E., Harrington, J. J., Lichstein, K. L., Sateia, M. J., Troxel, W. M., Zhou, E. S., Kazmi, U., Heald, J. L., & Martin, J. L. (2021). Behavioral and psychological treatments for chronic insomnia disorder in adults: An American Academy of Sleep Medicine clinical practice guideline. *Journal of Clinical Sleep Medicine, 17*(2), 255–262. https://doi.org/10.5664/jcsm.8986

Elliott, J. H., Synnot, A., Turner, T., Simmonds, M., Akl, E. A., McDonald, S., Salanti, G., Meerpohl, J., MacLehose, H., Hilton, J., Tovey, D., Shemilt, I., & Thomas, J. (2024). Living systematic review: Introduction—the why, what, when, and how. *Journal of Clinical Epidemiology, 169*, 111272.

Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2018). Empathy. *Psychotherapy, 55*(4), 399–410. https://doi.org/10.1037/pst0000175

European Commission. (2024). *EU AI Act: Regulation on artificial intelligence in high-risk applications*. Official Journal of the European Union.

FDA. (2023). *Software as a Medical Device (SaMD): Clinical evaluation*. U.S. Food and Drug Administration.

Flückiger, C. (2024). The alliance-outcome literature beyond Western face-to-face contexts. *Counselling and Psychotherapy Research, 24*(2), 487–495.

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy, 55*(4), 316–340.

https://doi.org/10.1037/pst0000172

Foa, E. B., Hembree, E. A., Rothbaum, B. O., & Rauch, S. A. M. (2019). *Prolonged exposure therapy for PTSD: Emotional processing of traumatic experiences—Therapist guide* (2nd ed.). Oxford University Press.

Foa, E. B., Yadin, E., & Lichner, T. K. (2012). *Exposure and response (ritual) prevention for obsessive compulsive disorder: Therapist guide* (2nd ed.). Oxford University Press.

Forbes, M. K., Ringwald, W. R., Allen, T. A., Cicero, D. C., Clark, L. A., DeYoung, C. G., Eaton, N. R., Hopwood, C. J., Kotov, R., Krueger, R. F., Latzman, R. D., Lynam, D. R., Markon, K. E., Miller, J. D., Mullins-Sweatt, S. N., Patrick, C. J., Pincus, A. L., Samuel, D. B., Sellbom, M., … Wright, A. G. C. (2024). Principles and procedures for revising the Hierarchical Taxonomy of Psychopathology. *Journal of Psychopathology and Clinical Science, 133*(1), 4–19.

Fosha, D. (2021). *Undoing aloneness and the transformation of suffering into flourishing: AEDP 2.0*. American Psychological Association.

Frederickson, J. (2024). *The lies we tell ourselves: How to face the truth, accept yourself, and create a better life* (rev. ed.). Seven Leaves Press.

Friedman, J. H. (1989). Regularized discriminant analysis. *Journal of the American Statistical Association, 84*(405), 165–175. https://doi.org/10.1080/01621459.1989.10478752

Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2024). Potential biases in machine learning algorithms using electronic health record data: An updated systematic review. *npj Digital Medicine, 7*, 89.

Gilbert, P. (2023). *Compassion focused therapy: The CBT distinctive features series* (2nd ed.). Routledge.

Greenberg, L. S. (2017). *Emotion-focused therapy* (rev. ed.). American Psychological Association.

Hadjistavropoulos, H. D., McCall, H. C., Schneider, L. H., Dear, B. F., Faller, Y. N., Sapkota, R. P., Bourgault-Fagnou, M., Pugh, N. E., Klein, B., Andersson, G., & Titov, N. (2024). Internet-delivered cognitive behaviour therapy for depression and anxiety in routine practice: A 10-year update. *World Psychiatry, 23*(1), 147–148.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23–34.

Hatch, S. G., Ehrenreich-May, J., & Boswell, J. F. (2024). Dissemination of evidence-based psychological treatments in low-resource U.S. settings: A systematic barrier analysis. *Implementation Research and Practice, 5*, 1–18.

Hayes, S. C., Hofmann, S. G., & Ciarrochi, J. (2022). A process-based approach to psychological diagnosis and treatment: The conceptual and treatment utility of an extended evolutionary meta model. *Clinical Psychology Review, 95*, 102155. https://doi.org/10.1016/j.cpr.2022.102155

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). *Acceptance and commitment therapy: The process and practice of mindful change* (2nd ed.). Guilford Press.

Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis, 52*(1), 258–271.

Hofmann, S. G., & Hayes, S. C. (2019). The future of intervention science: Process-based therapy. *Clinical Psychological Science, 7*(1), 37–50. https://doi.org/10.1177/216770261877

Hofmann, S. G., Sachdeva, A., & Hayes, S. C. (2024). Integrating computational psychiatry markers with process-based therapy: A framework for next-generation matching. *Cognitive and Behavioral Practice, 31*(1), 78–92.

Hoppen, T. H., Lindemann, A.-S., & Morina, N. (2024). Cognitive behavioural and EMDR therapies for posttraumatic stress disorder: An updated network meta-analysis. *European Journal of Psychotraumatology, 15*(1), 2316416.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179–185. https://doi.org/10.1007/BF02289447

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huys, Q. J. M., Browning, M., Paulus, M. P., & Frank, M. J. (2021). Advances in the computational understanding of mental illness. *Neuropsychopharmacology, 46*(1), 3–19. https://doi.org/10.1038/s41386-020-0746-4

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning with applications in R* (2nd ed.). Springer.

Jamieson, M., Putman, A., Bryan, M., Hansen, B., Klassen, P. E., Li, N., McQuaid, B., & Rudoler, D. (2024). Stratified care in cognitive behavioural therapy: A comparative evaluation of predictive modelling approaches for individualized treatment. *The Canadian Journal of Psychiatry, 69*(11), 853–862. https://doi.org/10.1177/07067437241295635

Jung, S., & Lee, S. (2024). Regularized factor analysis for small samples in clinical psychology research. *Psychological Methods, 29*(4), 622–639.

Kabat-Zinn, J. (2024). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness* (3rd ed.). Bantam.

Karyotaki, E., Efthimiou, O., Miguel, C., Bermpohl, F. M. G., Furukawa, T. A., Cuijpers, P., & Individual Patient Data Meta-Analyses for Depression (IPDMA-DE) Collaboration. (2021). Internet-based cognitive behavioral therapy for depression: A systematic review and individual patient data network meta-analysis. *JAMA Psychiatry, 78*(4), 361–371. https://doi.org/10.1001/jamapsychiatry.2020.4364

Kas, M. J., Penninx, B. W., Sommer, B., Serretti, A., Arango, C., & Marston, H. (2025). A precision psychiatry roadmap: Towards a biology-informed framework for mental disorders. *Molecular Psychiatry, 30*(11), 4985–4994. https://doi.org/10.1038/s41380-025-03070-5

Klecka, W. R. (1980). *Discriminant analysis*. Sage.

Knaak, S., Mantler, E., & Szeto, A. (2024). Mental health workforce capacity-building in 12 countries: Implementation outcomes from the WHO QualityRights initiative. *Lancet Psychiatry, 11*(7), 521–533.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., ... Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454–477. https://doi.org/10.1037/abn0000258

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). Sage.

Krueger, R. F., Hobbs, K. A., Conway, C. C., Dick, D. M., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Keyes, K. M., Latzman, R. D., Michelini, G., Patrick, C. J., Sellbom, M., Slade, T., South, S. C., Sunderland, M., Tackett, J., Waldman, I., Waszczuk, M. A., ... HiTOP Utility Workgroup. (2021). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum. *World Psychiatry, 20*(2), 171–193. https://doi.org/10.1002/wps.20844

Lazarus, A. A. (2005). Multimodal therapy. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 105–120). Oxford University Press.

Lemmens, L. H. J. M., van Bronswijk, S. C., Peeters, F. P. M. L., Arntz, A., Hollon, S. D., & Huibers, M. J. H. (2024). Long-term outcomes of cognitive therapy and interpersonal psychotherapy for adult depression: 7-year follow-up of a randomized clinical trial. *World Psychiatry, 23*(3), 432–443.

Levine, P. A. (2010). *In an unspoken voice: How the body releases trauma and restores goodness*. North Atlantic Books.

Lewis, C., Roberts, N. P., Bethell, A., Robertson, L., & Bisson, J. I. (2024). Internet-based cognitive and behavioural therapies for post-traumatic stress disorder (PTSD) in adults: A 10-year update. *Cochrane Database of Systematic Reviews, 2024*(2), CD011710.

Linehan, M. M. (2014). *DBT skills training manual* (2nd ed.). Guilford Press.

Lutz, W., Schaffrath, J., Eberhardt, S. T., Hehlmann, M. I., Schwartz, B., Schilling, V. N. L. S., Deisenhofer, A. K., Vehlen, A., Schürmann, S., & Rubel, J. A. (2024). Precision mental health and data-informed decision support in psychological therapy: An example. *Administration and Policy in Mental Health and Mental Health Services Research, 51*(5), 674–685. https://doi.org/10.1007/s10488-023-01330-6

Lutz, W., Schwartz, B., & Delgadillo, J. (2025). Advances in personalization of psychological interventions. *World Psychiatry, 24*(3), 395–404. https://doi.org/10.1002/wps.21342

Lyon, A. R., Pullmann, M. D., Walker, S. C., & D'Angelo, G. (2024). Adult-extended PracticeWise: Scaling evidence-based practice to non-pediatric populations. *Imple-*

mentation Science, 19, 17.

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research, 36*(4), 611–637. https://doi.org/10.1207/S15327906MBR3604_06

Manly, B. F. J., & Alberto, J. A. N. (2016). *Multivariate statistical methods: A primer* (4th ed.). CRC Press.

Markowitz, J. C., & Weissman, M. M. (2024). *Interpersonal psychotherapy: A clinician's guide* (2nd ed.). Oxford University Press.

Martell, C. R., Dimidjian, S., & Herman-Dunn, R. (2022). *Behavioral activation for depression: A clinician's guide* (2nd ed.). Guilford Press.

Mavranezouli, I., Megnin-Viggars, O., Daly, C., Dias, S., Welton, N. J., Stockton, S., Bhutani, G., Grey, N., Leach, J., Greenberg, N., Katona, C., El-Leithy, S., & Pilling, S. (2020). Psychological treatments for post-traumatic stress disorder in adults: A network meta-analysis. *Psychological Medicine, 50*(4), 542–555. https://doi.org/10.1017/S0033291720000070

McGovern, C. M., Owens, J. M., & Wagner, J. (2024). Who will stay and who will go? Identifying risk factors for psychotherapy dropout. *Counselling and Psychotherapy Research, 24*(4), 1432–1441. https://doi.org/10.1002/capr.12783

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.

Mele, M. L., Meloni, R., & Federici, S. (2025). Methods and assessment tools for therapeutic assignment in psychotherapy: A systematic review. *Cureus, 17*(9), e92110. https://doi.org/10.7759/cureus.92110

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158

Miller, W. R., & Rollnick, S. (2023). *Motivational interviewing: Helping people change and grow* (4th ed.). Guilford Press.

Mukuria, C., Hernandez Alava, M., & Brazier, J. (2024). Mapping psychotherapy preferences to quality-adjusted life-years: A C-NIP–QALY framework. *Health Economics,* *33*(2), 245–262.

Munder, T., Flückiger, C., Leichsenring, F., Abbass, A. A., Hilsenroth, M. J., Luyten, P., Rabung, S., Steinert, C., & Wampold, B. E. (2024). Common factors reconsidered: An updated synthesis. *Psychotherapy Research, 34*(7), 921–940.

Murphy, M. K., Black, N. A., Lamping, D. L., McKee, C. M., Sanderson, C. F., Askham, J., & Marteau, T. (1998). Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment, 2*(3), i–iv, 1–88. https://doi.org/10.3310/hta2030

Naslund, J. A., Aschbrenner, K. A., Marsch, L. A., McHugo, G. J., & Bartels, S. J. (2024). Digital mental health interventions in low- and middle-income countries: A 10-year update. *World Psychiatry, 23*(1), 23–38.

Nezu, A. M., Nezu, C. M., & D'Zurilla, T. J. (2013). *Problem-solving therapy: A treatment manual*. Springer.

Niederberger, M., & Spranger, J. (2020). Delphi technique in health sciences: A map. *Frontiers in Public Health, 8*, 457. https://doi.org/10.3389/fpubh.2020.00457

Norcross, J. C. (Ed.). (2011). *Psychotherapy relationships that work* (2nd ed.). Oxford University Press.

Norcross, J. C., & Goldfried, M. R. (Eds.). (2019). *Handbook of psychotherapy integration* (3rd ed.). Oxford University Press.

Norcross, J. C., & Wampold, B. E. (2019). Relationships and responsiveness in the psychological treatment of trauma: The tragedy of the APA Clinical Practice Guideline. *Psychotherapy, 56*(3), 391–399. https://doi.org/10.1037/pst0000228

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology, 73*, 719–748.

Nye, A., Delgadillo, J., & Barkham, M. (2023). Efficacy of personalized psychological interventions: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology, 91*(7), 389–397. https://doi.org/10.1037/ccp0000820

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Ogden, P., & Fisher, J. (2023). *Sensorimotor psychotherapy: Interventions for trauma and attachment* (rev. ed.). W. W. Norton.

Orlowski, E. W., Owen, J., Drinane, J. M., Tao, K., Kline, K., Adelson, J., DeBlaere, C., Miller, M. J., & Davis, D. E. (2025). The effects of cultural humility on therapeutic alliance and psychotherapy outcomes: A systematic review and meta-analysis. *Counselling and Psychotherapy Research, 25*, e12835. https://doi.org/10.1002/capr.12835

Park, J. M., Daleiden, E. L., & Chorpita, B. F. (2023). Adapting PracticeWise to adult outpatient clinical settings: Feasibility and reliability of a coding extension. *Cognitive and Behavioral Practice, 30*(2), 251–262.

Patel, V., Saxena, S., Lund, C., Kohrt, B., Kieling, C., Sunkel, C., Kola, L., Chang, O., Charlson, F., O'Neill, K., & Herrman, H. (2024). Transforming mental health systems globally: Principles and policy recommendations. *The Lancet, 403*(10437), 1810–1822.

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., Chisholm, D., Collins, P. Y., Cooper, J. L., Eaton, J., Herrman, H., Herzallah, M. M., Huang, Y., Jordans, M. J. D., Kleinman, A., Medina-Mora, M. E., Morgan, E., Niaz, U., Omigbodun, O., … UnÜtzer, J. (2018). The Lancet Commission on global mental health and sustainable development. *The Lancet, 392*(10157), 1553–1598.

Payne, P., Levine, P. A., & Crane-Godreau, M. A. (2024). Somatic experiencing: Using interoception and proprioception as core elements of trauma therapy. *Frontiers in Psychology, 15*, 1296234.

Petter, J., Schumacher, L., Echterhoff, J., Klein, J. P., Schramm, E., Härter, M., Hautzinger, M., & Kriston, L. (2025). Heterogeneity of treatment outcomes across therapists and sites in a randomized multicentre psychotherapy trial. *Clinical Psychology and Psychotherapy, 32*(3), e70087. https://doi.org/10.1002/cpp.70087

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

Reardon, M. L., Klein, K. P., Smith, M. M., Yarrington, J. S., Mathew, A. R., & Carl, J. R. (2024). Open-science requirements for matching-based digital therapeutics: A five-pillar framework. *Clinical Psychology: Science and Practice, 31*(3), 311–324.

Reddy, S., Fox, J., & Purohit, M. P. (2024). Explainable artificial intelligence in mental health: A systematic review. *Lancet Digital Health, 6*(1), e26–e35.

Resick, P. A., Monson, C. M., & Chard, K. M. (2017). *Cognitive processing therapy for PTSD: A comprehensive manual*. Guilford Press.

Rief, W., Glombiewski, J. A., Schedlowski, M., Bingel, U., & Klinger, R. (2024). Pre-therapy preference matching reduces dropout and improves outcomes in depression treatment: A randomized controlled trial. *Psychological Medicine, 54*(8), 1842–1851.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21*(2), 95–103. https://doi.org/10.1037/h0045357

Ryle, A., & Kerr, I. B. (2002). *Introducing cognitive analytic therapy: Principles and practice*. Wiley.

Sammer-Schreckenthaler, S., Lagetto, G., Unterrainer, H. F., & Gelo, O. C. G. (2025). Specific and common therapeutic factors in psychodynamic psychotherapy for children and adolescents: An overview. *Frontiers in Psychology, 16*, 1525849. https://doi.org/10.3389/fpsyg.2025.1525849

Saxon, D., Barkham, M., Bee, P., Bewick, B. M., & Brooks, H. (2024). Therapist-modality matching in routine psychological care: A 12-year analysis of NHS Talking Therapies routine outcome data. *British Journal of Clinical Psychology, 63*(4), 521–540.

Schiepek, G., & Pincus, D. (2023). Complexity science: A framework for psychotherapy integration. *Counselling and Psychotherapy Research, 23*, 941–955. https://doi.org/10.1002/capr.12626

Schmahmann, J. D. (2024). The cerebellum in neuropsychiatry: Mood, cognition, and motor coordination revisited. *Annual Review of Neuroscience, 47*, 73–98.

Schramm, E., Elsaesser, M., Jenkner, C., Hautzinger, M., & Herpertz, S. C.

(2024). Algorithm-based modular psychotherapy vs. cognitive-behavioral therapy for patients with depression, psychiatric comorbidities and early trauma: A proof-of-concept randomized controlled trial. *World Psychiatry, 23*(2), 257–266. https://doi.org/10.1002/wps.21204

Schünemann, H. J., Wiercioch, W., Etxeandia, I., Falavigna, M., Santesso, N., Mustafa, R., Ventresca, M., Brignardello-Petersen, R., Laisaar, K. T., Kowalski, S., Baldeh, T., Zhang, Y., Raid, U., Neumann, I., Norris, S. L., Thornton, J., Harbour, R., Treweek, S., Guyatt, G., … Akl, E. A. (2014). Guidelines 2.0: Systematic development of a comprehensive checklist for a successful guideline enterprise. *Canadian Medical Association Journal, 186*(3), E123–E142.

Schwartz, R. C. (2021). *No bad parts: Healing trauma and restoring wholeness with the Internal Family Systems model*. Sounds True.

Schwartz, R. C., & Sweezy, M. (2020). *Internal Family Systems therapy* (2nd ed.). Guilford Press.

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2018). *Mindfulness-based cognitive therapy for depression* (2nd ed.). Guilford Press.

Seuling, P. D., Fendel, J. C., Spille, L., Göritz, A. S., & Schmidt, S. (2024). Therapeutic alliance in videoconferencing psychotherapy compared to psychotherapy in person: A systematic review and meta-analysis. *Journal of Telemedicine and Telecare, 30*(5), 859–875.

Shapiro, F. (2018). *Eye movement desensitization and reprocessing (EMDR) therapy: Basic principles, protocols, and procedures* (3rd ed.). Guilford Press.

Sharma, S., Fitzpatrick, K. K., Smith, K. E., Wright, A. J., & Cuijpers, P. (2024). Randomized trial of brief interpersonal psychotherapy and cognitive behavioral therapy for depression delivered both in-person and by telehealth. *Journal of Affective Disorders, 354*, 218–227.

Singla, D. R., MacKinnon, D. P., Fuhr, D. C., Sikander, S., Rahman, A., & Patel, V. (2023). Multiple mediation analysis of the peer-delivered Thinking Healthy Programme for perinatal depression: Findings from two parallel, randomised controlled trials. *British Journal of Psychiatry, 222*(2), 76–82. https://doi.org/10.1192/bjp.2022.135

Slade, M. (2017). Implementing shared decision making in routine mental health care. *World Psychiatry, 16*(2), 146–153. https://doi.org/10.1002/wps.20412

Society of Clinical Psychology. (2023). *Research-supported psychological treatments: Empirically supported treatment listings*. APA Division 12.

Solomonov, N., & Barber, J. P. (2022). Unpacking the heterogeneity in treatment effects as a path to personalized psychotherapies: Are we there yet? *Clinical Psychology: Science and Practice, 29*(3), 308–310. https://doi.org/10.1037/cps0000101

Spranger, J., Heinl, A. K., & Niederberger, M. (2022). Reporting standards for Delphi panels in clinical research. *Health Research Policy and Systems, 20*, 81.

Stovell, D., Morrison, A. P., Panayiotou, M., & Hutton, P. (2024). Shared decision-making in mental healthcare: A systematic review and meta-analysis. *Lancet Psychiatry, 11*(4), 282–293.

Swift, J. K., Mullins, R. H., Penix, E. A., Roth, K. L., & Trusty, W. T. (2024). The importance of accommodating client preferences in psychotherapy: An updated meta-analysis. *Clinical Psychology Review, 110*, 102434.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Tackett, J. L., Brandes, C. M., King, K. M., & Markon, K. E. (2024). Psychology's replication crisis and clinical psychological science. *Annual Review of Clinical Psychology, 20*, 151–176.

Tetzlaff, M., Bruins, J., & Castelein, S. (2025). Associated factors of the quality of therapeutic alliance in people with severe mental illness: A systematic review. *European Psychiatry, 68*(1), 873. https://doi.org/10.1192/j.eurpsy.2025.873

Tolin, D. F., McKay, D., Forman, E. M., Klonsky, E. D., & Thombs, B. D. (2015). Empirically supported treatment: Recommendations for a new model. *Clinical Psychology: Science and Practice, 22*(4), 317–338. https://doi.org/10.1111/cpsp.12122

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Tschacher, W., Haken, H., & Kyselo, M. (2015). Alliance: A common factor of psychotherapy modeled by structural theory. *Frontiers in Psychology, 6*, 421. https://doi.org/10.3389/fpsyg.2015.00421

Ulberg, R., Ness, E., Dahl, H. S., Høglend, P. A., Critchfield, K., Blayvas, P., & Amlo, S. (2016). Relational interventions in psychotherapy: Development of a therapy process rating scale. *BMC Psychiatry, 16*, 310. https://doi.org/10.1186/s12888-016-1021-4

van Deurzen, E., & Adams, M. (2024). *Existential perspectives on human issues: A handbook for therapeutic practice* (2nd ed.). Bloomsbury.

van de Vijver, F. J. R., & Leung, K. (2021). *Methods and data analysis for cross-cultural research* (2nd ed.). Cambridge University Press.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*, 321–327. https://doi.org/10.1007/BF02293557

Wakefield, S., Kellett, S., Simmonds-Buckley, M., Stockton, D., Bradbury, A., & Delgadillo, J. (2021). Improving Access to Psychological Therapies (IAPT) in the United Kingdom: A systematic review and meta-analysis of 10-years of practice-based evidence. *British Journal of Clinical Psychology, 60*(1), 1–37.

Wampold, B. E. (2024). Personalization of psychotherapy: A commentary on its limits and promises. *Psychotherapy Research, 34*(4), 421–427.

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270–277. https://doi.org/10.1002/wps.20238

Wampold, B. E., Baldwin, S. A., Holtforth, M. G., & Imel, Z. E. (2017). What characterizes effective therapists? In L. G. Castonguay & C. E. Hill (Eds.), *How and why are some therapists better than others? Understanding therapist effects* (pp. 37–53). American Psychological Association.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.

Watts, A. L., Greene, A. L., Ringwald, W., Forbes, M. K., Brandes, C. M., Levin-Aspenson, H. F., & Delawalla, C. (2023). Factor analysis in personality disorders research: Modern issues and illustrations of practical recommendations. *Personality Disorders: Theory, Research, and Treatment, 14*(1), 105–117. https://doi.org/10.1037/per0000581

Webb, C. A., Forgeard, M., Israel, E. S., Lovell-Smith, N., Beard, C., & Björgvinsson, T. (2024). Personalized prescriptions of therapeutic skills from patient characteristics: An ecological momentary assessment approach. *Journal of Consulting and Clinical Psychology, 92*(7), 419–432.

Wen, A., Wolitzky-Taylor, K., Gibbons, R. D., & Craske, M. (2023). A randomized controlled trial on using predictive algorithm to adapt level of psychological care for community college students: STAND triaging and adapting to level of care study protocol. *Trials, 24*, 522. https://doi.org/10.1186/s13063-023-07441-7

White, M. (2007). *Maps of narrative practice*. W. W. Norton.

Williams, L. M., & Whitfield-Gabrieli, S. (2025). Neuroimaging for precision medicine in psychiatry. *Neuropsychopharmacology, 50*(2), 246–257. https://doi.org/10.1038/s41386-024-01917-z

Windle, E., Tee, H., Sabitova, A., Jovanovic, N., Priebe, S., & Carr, C. (2020). Association of patient treatment preference with dropout and clinical outcomes in adult psychosocial mental health interventions: A systematic review and meta-analysis. *JAMA Psychiatry, 77*(3), 294–302. https://doi.org/10.1001/jamapsychiatry.2019.3750

Wubbolding, R. E. (2024). *Reality therapy and self-evaluation: The key to client change* (2nd ed.). Routledge.

Yalom, I. D. (1980). *Existential psychotherapy*. Basic Books.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. Guilford Press.

Zhao, C., Wampold, B. E., Ren, Z., Zhang, L., & Jiang, G. (2022). The efficacy and optimal matching of an Internet-based acceptance and commitment therapy intervention for depressive symptoms among university students: A randomized controlled trial in China. *Journal of Clinical Psychology, 78*(7), 1354–1375. https://doi.org/10.1002/jclp.23329

Zilcha-Mano, S., & Fisher, H. (2022). Distinct roles of state-like and trait-like patient–therapist alliance in psychotherapy. *Nature Reviews Psychology, 1*, 194–210. https://doi.org/10.1038/s44159-022-00029-z

Author Note

Alex is the founder of Boston Neuromind LLC and the Catcher Technology Company. He served as Visiting Scholar at the Harvard Graduate School of Education, where he collaborated with the late Professor Kurt Fischer on Dynamic Skill Theory applied to learning optimization. The methodology presented herein underpins the ModalityCatcher β v2.1 system (https://modalitycatcher.com).

Funding: Self-funded. No external funding.

Data Availability: The 30 × 10 matrix and intermediate artifacts are available at the ModalityCatcher project repository upon request to the corresponding author.

Conflict of Interest: The author owns the ModalityCatcher trademark and domain. This methodology is publicly disclosed for academic critique and replication; commercial application is governed by separate licensing.

Acknowledgments: The author thanks the Delphi panel members (anonymized for confidentiality) and the three Stage 1 coders for their rigor and patience. The methodology benefited from conversations with colleagues in the Boston Neuromind clinical network and from the broader matching research community whose work is cited above.

Manuscript completed: 2026-05-10. Boston Neuromind LLC, Boston, MA, USA.