POSITION PAPER · SEPTEMBER 2026

The Difficulty of Knowing What to Do Next

Clinical Judgment Under Uncertainty in Psychotherapy Practice

AUTHOR

Sarah Ozol Shore, MS

Clinical Trainer, Clinical Effectiveness Institute

Every psychotherapy session requires the clinician to determine what is happening, what matters, what the therapeutic task is, and what should happen next. This paper argues that the profession certifies clinicians on knowledge and treatment fidelity while leaving this judgment almost entirely unobserved. Research on professional self-doubt, clinician self-assessment, experience, client concealment, detection of deterioration, supervision, routine outcome monitoring, competency standards, and case formulation shows that each ordinary safeguard on clinical judgment fails to calibrate it. The paper proposes that clinicians' knowledge of theories and techniques may coexist with a distinct difficulty at the point of clinical application, and it argues that clinical reasoning must become a discipline in its own right, with its own concepts, methods, and standards of evidence.

I. AN ORDINARY CLINICAL MOMENT

The client leaves at the end of their fourteenth session. She is in her mid-forties, came to therapy after the end of a long marriage, and has in many respects been a good client: punctual, diligent with between-session exercises, and thoughtful about her past. Her therapist, well trained and eight years in practice, formulated her low mood as linked to withdrawal from activity and a long-standing belief that her needs burden other people. Behavioral activation helped for a few weeks, cognitive work produced insight, and lately they have been talking about her mother. The depression scores have not moved and recently rose. The client agrees with every observation and

calls the sessions essential. The therapist can describe her in detail and name several reasonable approaches, yet she cannot say what is maintaining the problem, what the next session is for, whether the client's agreement is collaboration or accommodation, or whether the understanding that has guided fourteen sessions is still correct.

She has done everything competent clinicians are taught to do. Her difficulty lies in the judgment on which all of it depends: determining what is happening, what matters, what the therapeutic task is, and what should happen next. Psychotherapy rests on this judgment in every session. The profession nonetheless certifies clinicians on what they know and whether they can deliver established treatments with fidelity, and it has built almost no means of observing, measuring, or developing the judgment that decides when and whether that knowledge applies. The evidence reviewed here, drawn from the Clinical Effectiveness Institute research brief *What Psychotherapists Struggle With in Clinical Practice* (2026), shows that each ordinary safeguard on this judgment, namely self-assessment, experience, client feedback, and supervision, fails to calibrate it. The failure is structural, and it belongs to the profession.

II. A FAMILIAR KIND OF NOT KNOWING

Clinical uncertainty is a documented and recurring feature of practice. In an international study of roughly five thousand therapists, professional self-doubt, including feeling "unsure how best to deal effectively with a patient," emerged as a core dimension of difficulty (Orlinsky & Rønnestad, 2005; Nissen-Lie et al., 2017). It persists across careers: experienced therapists describe feelings of incompetence as "an ongoing part of the private experience of being a therapist" (Thériault & Gazzola, 2005), and many difficulties "would be found problematic by most therapists regardless of their level of experience" (Schröder & Davis, 2004).

Clinicians cannot read their effectiveness from either doubt or confidence. Self-doubt has accompanied better outcomes in some samples (Nissen-Lie et al., 2013, 2017) and worse outcomes in others (Heinonen et al., 2012; Odyniec et al., 2019). Confidence is inflated: clinicians place their own skill, on average, at the 80th percentile, and none below the 50th (Walfish et al., 2012). Therapists' estimates of their effectiveness across twelve problem domains were "generally no better than chance," and those who overestimated had patients with worse outcomes (Constantino et al., 2023).

III. WHAT EXPERIENCE PROVIDES

Experience reduces early termination (Goldberg, Rousmaniere, et al., 2016) and modestly improves judgment

accuracy (d = 0.12; Spengler et al., 2009), yet client outcomes do not improve with it. Among 170 therapists and 6,591 clients, outcomes declined very slightly as therapists gained experience (Goldberg, Rousmaniere, et al., 2016); a German replication found no association (Germer et al., 2022); and psychologists followed from practicum through licensure achieved the same or less client change later (Erekson et al., 2017). The most plausible explanation is that expertise develops only where practitioners receive timely, accurate feedback on their judgments, which psychotherapy supplies sporadically (Tracey et al., 2014). Without it, years of practice repeat judgment without refining it.

IV. THE INFORMATION AVAILABLE IN THE ROOM

The feedback clinicians do receive is distorted. Clients defer: they avoid criticizing their therapists, tolerate approaches that do not suit them, and protect the therapist's self-esteem (Rennie, 1994). In a large survey, 93% of clients reported having lied to a therapist, and nearly three quarters had lied about therapy itself, including pretending it was effective (Blanchard & Farber, 2016). Withdrawal ruptures take the form of deference and compliance, and so resemble cooperation (Safran & Muran, 2000). The client who agrees with everything may be collaborating or accommodating, and the therapist has no ordinary means of telling which.

Clinicians also miss deterioration. In a university counseling center, clinicians "rarely accurately predict[ed] who will not benefit from psychotherapy," while an actuarial method identified every client who had deteriorated (Hannan et al., 2005). Therapists' progress notes showed "considerable difficulty recognizing client deterioration" (Hatfield et al., 2010), and Danish trainees identified one of eight clients who got worse (Østergård et al., 2024). Clinical judgment carries real information, and statistical prediction outperforms it by a modest 13% (Ægisdóttir et al., 2006). The problem is the environment: feedback is noisy, delayed, and partly concealed, and regression to the mean and client politeness allow ineffective practice to appear effective (Lilienfeld et al., 2014). A mistaken understanding of a case can persist indefinitely, with no reliable occasion for the clinician to discover it.

V. SUPERVISION AND ITS INSTRUMENTS

Supervision is the profession's institutional answer to these limits, and nothing in the evidence shows that it calibrates clinical judgment. Supervisor ratings are lenient, show halo effects, and poorly predict later ratings by other supervisors (Gonsalvez & Freestone, 2007); across nine Australian programs, 88% of placements rated below standard were passed anyway (Gonsalvez et al., 2021). Supervisors see a filtered version of practice: 84% of

trainees withheld something in a single session (Mehr et al., 2010), clinical mistakes among it (Cook et al., 2018). Supervisor identity accounted for less than 1% of variance in client outcomes in one large dataset (Rousmaniere et al., 2016) and none in a replication (Whipple et al., 2020), and a recent meta-analysis found no demonstrated effect of supervision on patient symptoms relative to active comparisons, with very low certainty of evidence (Schreyer et al., 2025). Supervision serves real purposes in socialization, ethical oversight, and support; as a corrective to judgment, it relies on instruments that share the limits of the clinician's own perception.

VI. KNOWING THAT AND KNOWING WHY

Routine outcome monitoring was designed to supply the missing signal, and its results locate the deeper problem. Across 58 studies, progress feedback produced a small improvement in outcomes (d = 0.15). For off-track clients, progress scores alone produced an effect near zero (d = 0.04), while feedback paired with clinical support tools produced a considerably larger one (d = 0.36; de Jong et al., 2021; see also Lambert et al., 2018). Those tools prompt the clinician to reconsider the alliance, motivation, social support, the diagnosis, and the level of care. Knowing that a case is failing accomplishes little until the clinician reopens the case and works out why. That second task is clinical reasoning. From the rising score, the therapist in the vignette learns only that her concern was justified; she still cannot say whether the maintaining process is avoidance, grief, an unrecognized depressive episode, a relational pattern enacted in the sessions, or something she has not considered.

VII. WHAT THE PROFESSION ASKS AND WHAT IT OBSERVES

The profession's standards require this capacity. American Psychological Association accreditation requires doctoral students to "evaluate intervention effectiveness and adapt intervention goals and methods consistent with ongoing progress evaluation"; counseling standards added "critical thinking and reasoning strategies for clinical judgment" in 2024; marriage and family therapists must "recognize when treatment goals and plan require modification"; United Kingdom practitioner psychologists must "revise formulations in the light of ongoing intervention"; and psychiatry milestones require differential diagnosis "while avoiding premature closure."

A standardized assessment of clinical judgment, formulation, or treatment decision-making is not involved. Competence is certified through supervisor ratings, program evaluations, and periodic observation, the instruments shown above to be lenient, filtered, and weakly related to outcomes. Ratings of competence in

specific treatment models correlate with client outcome at r = .07, and adherence at r = .02 (Webb et al., 2010). When therapists' responses to recorded moments of interpersonal difficulty are observed directly, by contrast, performance has predicted outcomes that training status and self-report did not (Anderson et al., 2009; Anderson et al., 2016). What a clinician does in a clinical moment carries information that credentials miss.

Where treatment decisions have been studied, the results are telling. When 115 practitioners formulated a standard case, agreement was good for descriptive elements and fell for elements "requiring theory-driven inference," and only 44% of formulations were rated at least good enough (Kuyken et al., 2005). Clinicians agreed poorly about underlying mechanisms even when trained in the same model on the same day (Persons et al., 1995; Persons & Bertagnolli, 1999). Clinicians describe cases more reliably than they explain them, and the capacity is teachable: two hours of formulation training raised the average clinician above about 86% of untrained peers (Kendjelic & Eells, 2007).

VIII. A HYPOTHESIS

The Clinical Effectiveness Institute's position follows from this evidence. Clinicians may possess substantial knowledge of diagnoses, models, techniques, and evidence-based treatments while facing a distinct difficulty at the point of application: determining what information matters, what process is organizing the problem, what the therapeutic task is, what should happen now, and when the working understanding must be revised. Knowing what an intervention is and knowing when, why, whether, and in what sequence to use it with a particular client are different accomplishments. The first is taught through instruction and assessed by examination. The second is exercised under ambiguity, interpersonal pressure, and partly concealed information, and the profession does not assess it directly. More protocols, more supervision, and greater confidence each leave it untouched.

The convergent evidence is substantial: the decline from descriptive to inferential accuracy in formulation, the weaker link between formulation and treatment planning among non-experts (Eells et al., 2005), the less integrated knowledge structures of novice counselors (Mayfield et al., 1999), the frequent non-use of exposure for PTSD by psychologists trained in it (Becker et al., 2004), the near-zero relationship between rated model competence and outcome, and the finding that feedback helps mainly when it prompts reconsideration of the case. No study has yet measured knowledge and reasoning in the same clinicians, and none has tested whether better reasoning produces better outcomes. That absence is itself the finding: the capacity on which every treatment decision depends has never been made an object of direct investigation.

IX. WHAT THE FIELD HAS NOT BUILT

Psychotherapy research has produced a substantial science of diagnoses, models, common factors, the alliance, and outcomes. It has produced almost nothing on the act that governs how any of that knowledge is used: the clinician's interpretation of ambiguous material and decision about what to do next. Medicine, with due allowance for the differences between the fields, built a research program on clinical reasoning and diagnostic error, and assessment methods to observe it. Psychotherapy has no equivalent; standardized assessment of process-based competence remains at the protocol stage (Gorinelli et al., 2025).

The consequence is that the profession trains, supervises, licenses, and credentials clinicians without ever observing the capacity on which every treatment decision depends. Next week the therapist in the vignette will revise her understanding of her client. Neither she, her supervisor, nor her licensing board has any means of knowing whether she got it right. Until clinical reasoning is studied, taught, and assessed as a discipline in its own right, with its own concepts, methods, and standards of evidence, the profession's account of competence will rest on everything except the judgment that decides what happens in the room. Establishing that discipline is the task of the subsequent paper, *The Missing Discipline*.

REFERENCES

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The Meta-Analysis of Clinical Judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34*(3), 341-382. https://doi.org/10.1177/0011000005285875

Anderson, T., Crowley, M. E. J., Himawan, L., Holmberg, J. K., & Uhlin, B. D. (2016). Therapist facilitative interpersonal skills and training status: A randomized clinical trial on alliance and outcome. *Psychotherapy Research, 26*(5), 511-529. https://doi.org/10.1080/10503307.2015.1049671

Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology, 65*(7), 755-768. https://doi.org/10.1002/jclp.20583

Becker, C. B., Zayfert, C., & Anderson, E. (2004). A survey of psychologists' attitudes towards and utilization of exposure therapy for PTSD. *Behaviour Research and Therapy, 42*(3), 277-292. https://doi.org/10.1016/S0005-7967(03)00138-4

Blanchard, M., & Farber, B. A. (2016). Lying in psychotherapy: Why and what clients don't tell their therapist about therapy and their relationship. *Counselling Psychology Quarterly, 29*(1), 90-112. https://doi.org/10.1080/09515070.2015.1085365

Clinical Effectiveness Institute. (2026). *What psychotherapists struggle with in clinical practice: A research brief* [Internal research brief].

Constantino, M. J., Boswell, J. F., Coyne, A. E., Muir, H. J., Gaines, A. N., & Kraus, D. R. (2023). Therapist perceptions of their own measurement-based, problem-specific effectiveness. *Journal of Consulting and Clinical Psychology, 91*(8), 474-484. https://doi.org/10.1037/ccp0000813

Cook, R. M., Welfare, L. E., & Romero, D. E. (2018). Counselor-in-training intentional nondisclosure in onsite supervision: A content analysis. *The Professional Counselor, 8*(2), 115-130. https://doi.org/10.15241/rmc.8.2.115

de Jong, K., Conijn, J. M., Gallagher, R. A. V., Reshetukha, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review, 85*, 102002. https://doi.org/10.1016/j.cpr.2021.102002

Eells, T. D., Lombart, K. G., Kendjelic, E. M., Turner, L. C., & Lucas, C. P. (2005). The quality of psychotherapy case formulations: A comparison of expert, experienced, and novice cognitive-behavioral and psychodynamic therapists. *Journal of Consulting and Clinical Psychology, 73*(4), 579-589. https://doi.org/10.1037/0022-006X.73.4.579

Erekson, D. M., Janis, R., Bailey, R. J., Cattani, K., & Pedersen, T. R. (2017). A longitudinal investigation of the impact of psychotherapist training: Does training improve client outcomes? *Journal of Counseling Psychology, 64*(5), 514-524. https://doi.org/10.1037/cou0000252

Germer, S., Weyrich, V., Bräscher, A.-K., Mütze, K., & Witthöft, M. (2022). Does practice really make perfect? A longitudinal analysis of the relationship between therapist experience and therapy outcome: A replication of Goldberg, Rousmaniere, et al. (2016). *Journal of Counseling Psychology, 69*(5), 745-754. https://doi.org/10.1037/cou0000608

Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T., & Wampold, B. E. (2016). Do psychotherapists improve with time and experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counseling Psychology, 63*(1), 1-11. https://doi.org/10.1037/cou0000131

Gonsalvez, C. J., & Freestone, J. (2007). Field supervisors' assessments of trainee performance: Are they reliable and valid? *Australian Psychologist, 42*(1), 23-32. https://doi.org/10.1080/00050060600827615

Gonsalvez, C. J., Terry, J., Deane, F. P., Nasstasia, Y., Knight, R., & Gooi, C. H. (2021). End-of-placement failure rates among clinical psychology trainees: Exceptional training and outstanding trainees or poor gate-keeping? *Clinical Psychologist, 25*(3), 294-305. https://doi.org/10.1080/13284207.2021.1927692

Gorinelli, S., Sairanen, E., Hofmann, S. G., Ciarrochi, J., Tuomisto, M. T., & Keinonen, K. (2025). A process-based competence evaluation for evidence-based dissemination (PROCEED): A study protocol. *BMC Psychology, 13*, 912.

https://doi.org/10.1186/s40359-025-03228-4

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology, 61*(2), 155-163. https://doi.org/10.1002/jclp.20108

Hatfield, D., McCullough, L., Frantz, S. H. B., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology & Psychotherapy, 17*(1), 25-32. https://doi.org/10.1002/cpp.656

Heinonen, E., Lindfors, O., Laaksonen, M. A., & Knekt, P. (2012). Therapists' professional and personal characteristics as predictors of outcome in short- and long-term psychotherapy. *Journal of Affective Disorders, 138*(3), 301-312. https://doi.org/10.1016/j.jad.2012.01.023

Kendjelic, E. M., & Eells, T. D. (2007). Generic psychotherapy case formulation training improves formulation quality. *Psychotherapy: Theory, Research, Practice, Training, 44*(1), 66-77. https://doi.org/10.1037/0033-3204.44.1.66

Kuyken, W., Fothergill, C. D., Musa, M., & Chadwick, P. (2005). The reliability and quality of cognitive case formulation. *Behaviour Research and Therapy, 43*(9), 1187-1201. https://doi.org/10.1016/j.brat.2004.08.007

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy, 55*(4), 520-537. https://doi.org/10.1037/pst0000167

Lilienfeld, S. O., Ritschel, L. A., Lynn, S. J., Cautin, R. L., & Latzman, R. D. (2014). Why ineffective psychotherapies appear to work: A taxonomy of causes of spurious therapeutic effectiveness. *Perspectives on Psychological Science, 9*(4), 355-387. https://doi.org/10.1177/1745691614535216

Mayfield, W. A., Kardash, C. M., & Kivlighan, D. M., Jr. (1999). Differences in experienced and novice counselors' knowledge structures about clients: Implications for case conceptualization. *Journal of Counseling Psychology, 46*(4), 504-514. https://doi.org/10.1037/0022-0167.46.4.504

Mehr, K. E., Ladany, N., & Caskie, G. I. L. (2010). Trainee nondisclosure in supervision: What are they not telling you? *Counselling and Psychotherapy Research, 10*(2), 103-113. https://doi.org/10.1080/14733141003712301

Nissen-Lie, H. A., Monsen, J. T., Ulleberg, P., & Rønnestad, M. H. (2013). Psychotherapists' self-reports of their interpersonal functioning and difficulties in practice as predictors of patient outcome. *Psychotherapy Research, 23*(1), 86-104. https://doi.org/10.1080/10503307.2012.735775

Nissen-Lie, H. A., Rønnestad, M. H., Høglend, P. A., Havik, O. E., Solbakken, O. A., Stiles, T. C., & Monsen, J. T. (2017). Love yourself as a person, doubt yourself as a therapist? *Clinical Psychology & Psychotherapy, 24*(1), 48-60. https://doi.org/10.1002/cpp.1977

Odyniec, P., Probst, T., Margraf, J., & Willutzki, U. (2019). Psychotherapist trainees' professional self-doubt and negative personal reaction: Changes during cognitive behavioral therapy and association with patient progress. *Psychotherapy Research, 29*(1), 123-138. https://doi.org/10.1080/10503307.2017.1315464

Orlinsky, D. E., & Rønnestad, M. H. (2005). *How psychotherapists develop: A study of therapeutic work and professional growth*. American Psychological Association.

Østergård, O. K., Grønnebæk, L., & Nilsson, K. K. (2024). Do therapists know when their clients deteriorate? An investigation of therapists' ability to estimate and predict client change during and after psychotherapy. *Clinical Psychology & Psychotherapy, 31*(6), e70015. https://doi.org/10.1002/cpp.70015

Persons, J. B., & Bertagnolli, A. (1999). Inter-rater reliability of cognitive-behavioral case formulations of depression: A replication. *Cognitive Therapy and Research, 23*(3), 271-283. https://doi.org/10.1023/A:1018791531158

Persons, J. B., Mooney, K. A., & Padesky, C. A. (1995). Interrater reliability of cognitive-behavioral case formulations. *Cognitive Therapy and Research, 19*(1), 21-34. https://doi.org/10.1007/BF02229674

Rennie, D. L. (1994). Clients' deference in psychotherapy. *Journal of Counseling Psychology, 41*(4), 427-437.

Rousmaniere, T. G., Swift, J. K., Babins-Wagner, R., Whipple, J. L., & Berzins, S. (2016). Supervisor variance in psychotherapy outcome in routine practice. *Psychotherapy Research, 26*(2), 196-205. https://doi.org/10.1080/10503307.2014.963730

Safran, J. D., & Muran, J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. Guilford Press.

Schreyer, B., Leithner, C., Eilers, R., Gossmann, K., & Rosner, R. (2025). The effects of clinical supervision on supervisees and patient outcomes in psychotherapy: A systematic review and meta-analysis. *Frontiers in Psychiatry, 16*, 1705578. https://doi.org/10.3389/fpsyt.2025.1705578

Schröder, T. A., & Davis, J. D. (2004). Therapists' experience of difficulty in practice. *Psychotherapy Research, 14*(3), 328-345. https://doi.org/10.1093/ptr/kph028

Spengler, P. M., White, M. J., Ægisdóttir, S., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G. R., & Rush, J. D. (2009). The Meta-Analysis of Clinical Judgment project: Effects of experience on judgment accuracy. *The Counseling Psychologist, 37*(3), 350-399. https://doi.org/10.1177/0011000006295149

Thériault, A., & Gazzola, N. (2005). Feelings of inadequacy, insecurity, and incompetence among experienced therapists. *Counselling and Psychotherapy Research, 5*(1), 11-18. https://doi.org/10.1080/14733140512331343840

Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist, 69*(3), 218-229. https://doi.org/10.1037/a0035099

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*(2), 639-644. https://doi.org/10.2466/02.07.17.PR0.110.2.639-644

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2), 200-211. https://doi.org/10.1037/a0018912

Whipple, J. L., Hoyt, T., Rousmaniere, T., Swift, J., Pedersen, T., & Worthen, V. (2020). Supervisor variance in

psychotherapy outcome in routine practice: A replication. *SAGE Open, 10*(1). https://doi.org/10.1177/2158244019899047

Standards cited: American Psychological Association, Commission on Accreditation (2015, revised 2026), Implementing Regulation C-8 D; Council for Accreditation of Counseling and Related Educational Programs (2024), Standard 3.E.2; American Association for Marriage and Family Therapy (2004), Core Competency 3.4.2; Health and Care Professions Council (2023), Standards of Proficiency for Practitioner Psychologists 4.3 and 11.7; Accreditation Council for Graduate Medical Education (2020), Psychiatry Milestones PC2.

RELATED CEI TRAINING

Clinical Formulation: The Missing Link Between Understanding and Intervention addresses one of the judgments examined in this paper: how clinicians develop a working account of what is happening in a particular client, and how that account informs the clinical work that follows. It does not take up the broader professional question this paper raises. Details and current dates at sarahozolshore.com.

Suggested citation: Shore, S. O. (2026). *The difficulty of knowing what to do next: Clinical judgment under uncertainty in psychotherapy practice.* Clinical Effectiveness Institute.

Sarah Ozol Shore, MS

Clinical Trainer, Clinical Effectiveness Institute