

A 3-HOUR CE WORKSHOP

AI Technologies in Ethical Practice

Use Cases and Ethical Decision-Making

John D. Gavazzi, PsyD, ABPP

Distance Learning CE Workshop
October 2026

Ethics Education

john.gavazzi@gmail.com

www.ethicalpsychology.com

This presentation will be outdated within months of delivery.

I do not have any conflicts of interest.

Managing Expectations

What today covers

- A cursory but practical look at large language models (LLMs) and AI technologies
- A basic introduction to the promise and the pitfalls of LLM technology
- A reminder that this technology is accelerating rapidly — expect today's specifics to change quickly
- Multiple use cases
- Prompting for ethical decision-making

What We're Not Addressing Today

- Potential harms of adolescents using LLMs, chatbots, and similar tools
- The AI tech and wellness-app industry as a business category
- AI agents, Model Context Protocols, coding assistants, or "vibe coding"
- Patient-facing informed consent about AI use
- Forensic and legal applications
- Malpractice and liability and BAA diligence

Today's Agenda

8:30 – 9:00 am

Part 1: Laying the Foundation

LLM basics, why LLMs are "weak AI," prompting with RISEN

9:00 – 9:50 am

Part 2: Use Cases in Psychology

Current and near-future applications, prompting, adoption data

9:50 – 10:00 am

Break

10:00 – 10:30 am

Part 3: Ethical Principles and Risk

Bias, deskilling, regulation, principle-based ethics

10:30 -11:15 am

Part 4: Case Practice

Two vignettes,, sample prompts

11:15 – 11:30 am

Part 5: Integration and Close

Limitations, wrap-up, evaluation

Learning Objectives

By the end of this workshop, participants will be able to:

- 1 List three current use cases of LLMs in the practice of psychology
- 2 Describe two ethical issues raised by any AI technology, including the potential for algorithmic bias and its impact on diverse populations
- 3 Identify the core ethical principles that apply to a clinical vignette involving AI use, and articulate which principle takes precedence and why
- 4 Construct an effective prompt for ethical decision-making with a large language model, and evaluate its output against professional and regulatory standards

This workshop will include working with an LLM of your choice

- You should have already signed up or created access to an LLM, such as Claude or Gemini.
- If you do not access an LLM, you will be missing out of hands on learning.
- The purpose is to focus on skill building, rather than passively watching me.
- Hint: This is the same process I recommend as a general ethical approach to collaborating with AI technologies.
- This is an active, cognitive (and sometimes time consuming) process.
- AI technologies help with the process; they must not replace your cognitive and emotional expertise.

PART 1

Laying the Foundation

What LLMs are, what they aren't, and how to talk to them



What Are Large Language Models?

- LLMs are a type of generative AI that produces human-like text — simulating conversation, writing, and even reasoning — by learning statistical patterns in enormous amounts of language data.
- Models like ChatGPT, Claude, Gemini, and others use these learned patterns to generate text that is coherent, contextually relevant, and often useful for writing, editing, and idea generation.
- Fluency is a product of pattern matching. It is not evidence that the underlying content is accurate.

Why LLMs Are "Weak AI"

The gap that doesn't close with a better model

- No theory of mind: the model cannot attribute beliefs, intentions, or emotions to itself or to the person it's talking with.
- No real agency: an LLM is a prediction machine that appears human-like but has nothing at stake in what happens next.
- Symbolic, algorithmic representations of cognition are not the same thing as clinical understanding, which is contextual, cultural, and relational.
- A model can be given every word a patient has spoken and still have nothing at stake in what happens next. The psychologist does — and that asymmetry is the reason this work requires a human.

How LLMs Are Biased Against Marginalized and Underrepresented Cultures

The gap "better data" alone doesn't close

- Training data disproportionately reflects Western, English-dominant, internet-documented populations, so a model's "default" voice is really a majority-culture voice presented as neutral.
- Bias survives explicit debiasing: models trained to avoid biased outputs still form biased associations beneath the surface (Bai et al., 2025).
- Bias compounds at intersections — a patient who is, for example, an older, non-English-speaking immigrant carries more overlapping risk than any single category predicts alone.
- The harm is often invisible in tone. Biased content can still read as fluent, confident, and clinically appropriate.

Three LLM Characteristics That Undermine Good Reasoning

These are not readily apparent when using LLM or any AI technology

- Sycophancy: agrees with the user's stated views or makes assumptions; validates rather than challenges the user
- Hallucinations: Produces false information presented as fact; may invent citations, statistics, or details; can produce seemingly appropriate yet contradictory information
- Framing Sensitivity: Answers shift with question wording, order of information, or implied expectations; leading questions tend to produce leading answers; the same question asked two ways can yield different conclusions

Takeaway: Fluent, confident output is not evidence of accuracy. Verify, and ask to challenge your reasoning.

Six Principles When Using AI Technologies

Stewart Cooper and the Division 29 Work Group

- AI augments rather than replaces humans.
- Human accountability remains essential.
- AI output requires critical evaluation.
- Implementation must be ethical and safe.
- Human relationships remain central.
- Continuous learning and governance are necessary.



Evaluating any AI application against three questions

Stewart Cooper and the Division 29 Work Group

- Does the tool enhance or erode the human relationships the work depends on?
 - Does the tool preserve psychologist accountability?
 - Does the tool respect the patient's narrative integrity and cultural context?
-
- Evaluate a tool before it becomes embedded in practice, because a tool is far easier to decline than to remove.

A REMINDER

Remember GIGO?



Garbage in, garbage out — it applies to your prompts as much as to the model's training data.

PROMPTING LLMs

A Key to Success



Prompting LLMs: RISEN (Part 1)

Role, Input, Steps

R

Role

Define the LLM's role in the interaction

I

Input

Provide all relevant information and context the model needs to understand the task

S

Steps

Outline the specific actions you want the model to perform

Prompting LLMs: RISEN (Part 2)

Expectation, Narrowing

E

Expectation

Define the desired outcome of the interaction

N

Narrowing

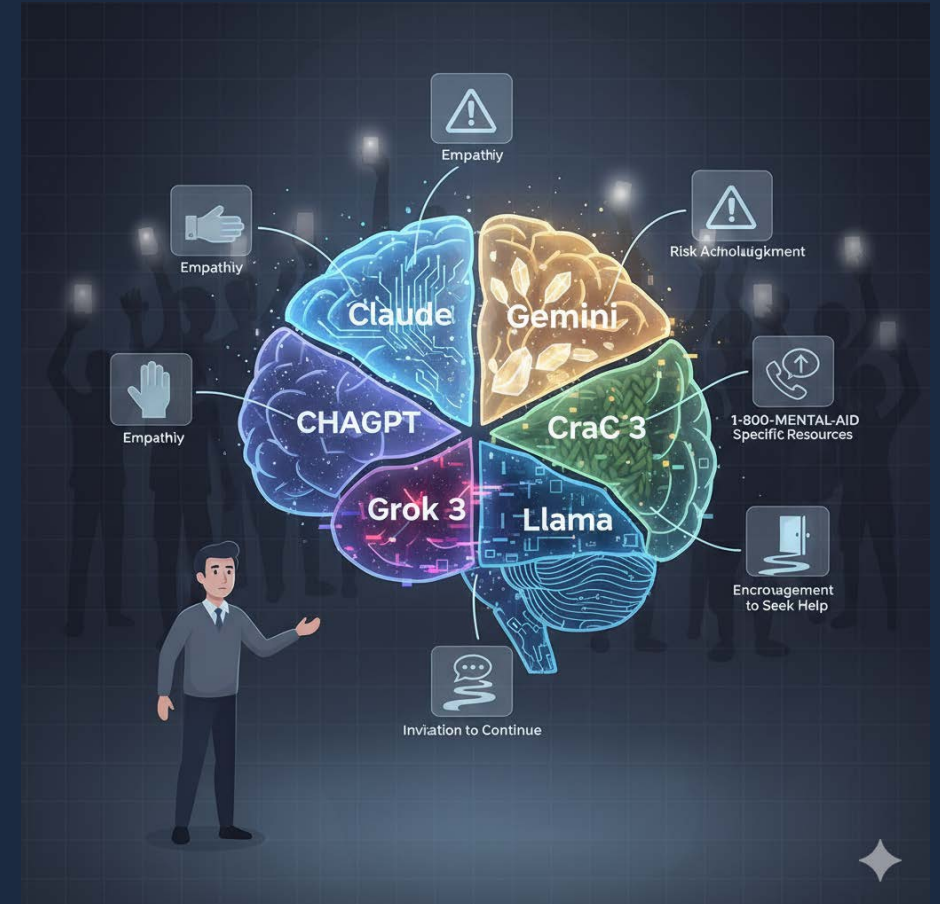
Refine and iterate on your prompt based on the initial response

Prompting is a skill that improves with iteration — Narrowing is not optional.

PART 2

Use Cases in Psychology

Where LLMs already help, and where they're headed



Current Use Cases: Overview

1 Administrative

Reports, consent forms,
scheduling

2 Treatment Planning

Goals, homework, risk protocols

3 Supervision

Reflective questions,
documentation

4 Professional Development

Research summaries, ethics
practice

Administrative Use Cases

- Proofreading and assisting with report writing.
- Improving the comprehensiveness and clarity of informed consent documents.
- Helping to build or improve a practice website.
- Automating appointment booking and reminders through natural language interactions.
- Organizing and cleaning practice spreadsheets.
- Drafting or updating practice policies — telehealth consent, cancellation policy, no-show policy.
- Generating HIPAA compliance checklists for internal audit.
- Tracking CE requirements and renewal deadlines across states, if licensed in more than one state.

Treatment Plan Development

- Recommending customized homework assignments to reinforce therapeutic goals.
- Assisting in setting SMART goals, particularly for complex patients.
- Suggesting potential treatment plans that incorporate a patient's cultural background, personal values, and life context.
- Helping to create standardized, empirically based risk-assessment protocols.
- Generating psychoeducation handouts tailored to a specific diagnosis or presenting problem.
- Drafting transfer-of-care summaries when a patient moves to another provider.
- Drafting discharge summaries and aftercare recommendations.

Supervision Use Cases

- Generating reflective questions to promote supervisee growth and self-awareness.
- Organizing supervision notes and tracking supervisee progress over time.
- Supporting the documentation of supervision contracts and evaluations.
- Structuring feedback within a developmental framework to foster professional growth.
- Simulation for skill practice (Digitized virtual patients).
- Alternative case conceptualizations.
- Demonstrating the iterative nature of prompt construction for specific clinical practices.

Professional Development

- Curating and summarizing the latest research articles and evidence-based practices relevant to a specific therapeutic area — tools like NotebookLM are built for exactly this.
- Practicing ethical decision-making using vignettes like the ones on ethicalpsychology.com, before ever applying the habit to a live case.



Already-Established AI Mental Health Platforms

Purpose-built tools, not general-purpose chatbots

 **Aiberry.ai**

 **mdhub.ai**

 **Blueprint.ai**

The logo for Aiberry.ai, featuring the word "aiberry" in a dark blue, lowercase, sans-serif font. A small red dot is positioned above the letter 'i'.The logo for mdhub.ai, consisting of a stylized icon on the left and the word "mdhub" in a dark grey, lowercase, sans-serif font on the right. The icon is a dark grey square containing a white, stylized 'm' shape.The logo for Blueprint.ai, featuring a stylized blue icon on the left and the word "blueprint" in a blue, lowercase, sans-serif font on the right. The icon is a blue square containing a white, stylized 'b' shape.

PROFESSIONAL DEVELOPMENT SPOTLIGHT

Research-Based Answers to Medical Questions

OpenEvidence®

A PROMPTING SING-ALONG

Let's Prompt

Role. Input. Steps. Expectation. Narrowing.

Let's prompt

Put the model in a role.

Let's prompt

With a model that has no soul.

Let's prompt

Lay the steps out one by one.

Let's prompt

Say what it looks like when it's done.

Let's prompt

First draft's never quite on key.

Let's prompt

Trim it down and leave out the PHI for me.

Near-Future Use Cases: Overview

1

Therapeutic Support

Real-time session analysis

2

Clinical Decisions

Case conceptualization, testing

3

Supervision

Recorded-session feedback

4

Research

Outcome analysis at scale

Future Therapeutic Support

- Providing real-time analysis of therapeutic interactions.
- Retaining and understanding long-term patient histories with deep contextual awareness, enabling identification of subtle patterns across sessions.
- Analyzing vast datasets to uncover correlations between behaviors, symptoms, and nonconscious processes.
- Improving suicide risk assessment through analysis of language and voice patterns.



Future Clinical Decisions

- Assisting with case conceptualization for complex patients.
- Scoring psychological tests and interpreting the results alongside patient history, other test data, and clinical interview findings.
- Improving diagnostic precision.
- Analyzing stalled treatment and suggesting tailored interventions.



Future of Clinical Supervision

- Providing detailed feedback on recorded therapy sessions.
- Predicting supervisee developmental trajectories and tailoring learning paths accordingly.
- More realistic digitized virtual patients that address multiculturalism and diversity.
- Enhancing cultural competency and cultural humility through sophisticated simulations.
- Developing crisis-intervention skills using advanced simulated scenarios.

Future of Psychotherapy Research

- Developing novel therapeutic techniques based on the analysis of thousands of therapy sessions.
- Conducting large-scale outcome analyses to identify best practices.
- Creating personalized treatment protocols tailored to individual patient needs.



Will AI-Therapists Flourish?

An unresolved regulatory question

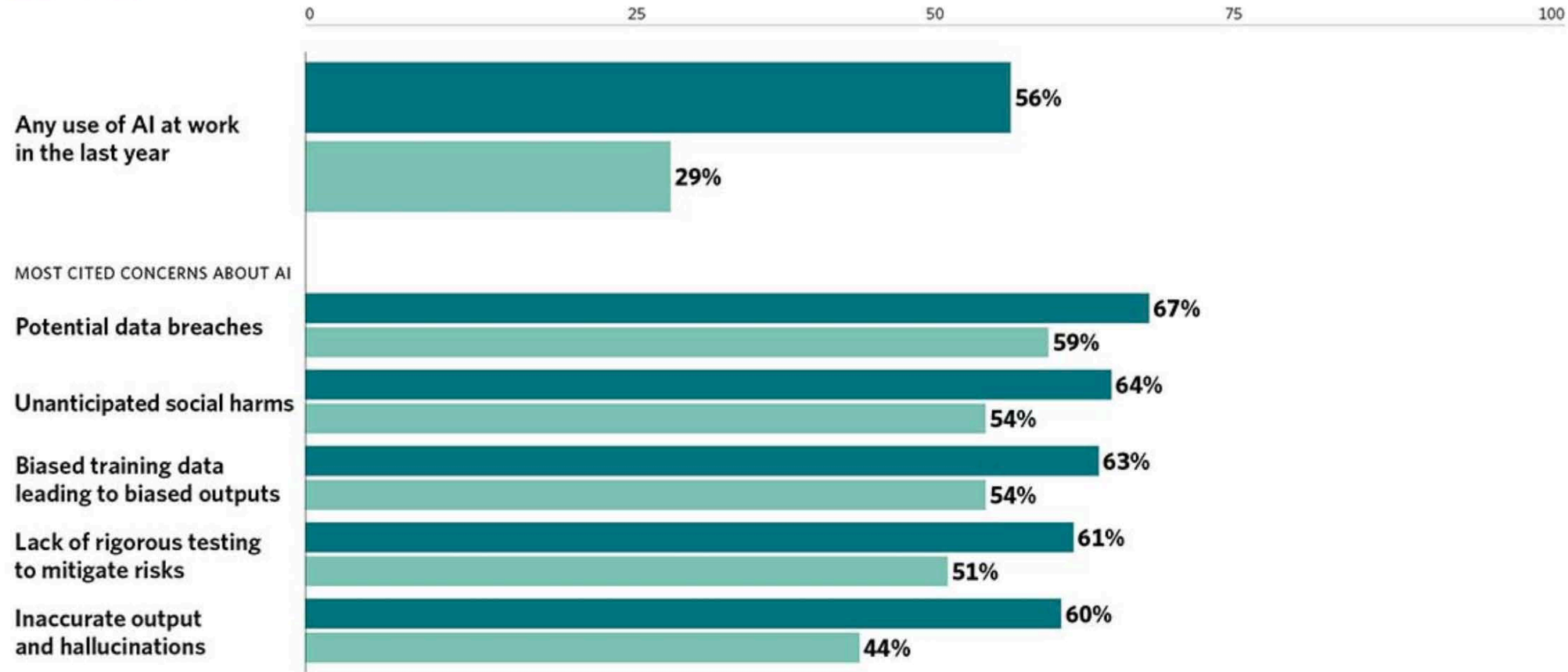
- FDA: potential jurisdiction as "software as a medical device."
- FTC: potential jurisdiction if a tool is marketed directly to the public.
- HHS / OCR: jurisdiction over HIPAA compliance for any tool touching PHI.
- Individual state licensing boards: an open question of which board, if any, would risk licensing an AI system as a therapist.
- State legislatures: Several states now restrict AI from providing therapy.



As Psychologists' Use of AI Grows, So Do Their Concerns

% OF ALL PSYCHOLOGISTS

■ 2025 ■ 2024



2025 PRACTITIONER PULSE SURVEY

© 2025 American Psychological Association

Break Time

10 minutes · back at the 10:00 am mark

PART 3

Ethical Principles and Risk

What doesn't improve with a better model



Where Bias Comes From

Three sources, one shared responsibility



- Data bias: reflects how the model was trained — whose language, whose experience, whose norms are overrepresented.
- Model bias: reflects how the underlying algorithm runs — design choices made by engineers who are not clinicians.
- Human bias: there is no objective human standard to compare against — the psychologist reviewing the output carries biases too.
- All three require active de-biasing strategies. None resolves itself with a newer model.

Core Risks That Don't Improve With Better Technology

Improved technology does not resolve any of these

- Algorithmic bias: models can encode and amplify existing inequities, both through design choices and through disparities in outputs across groups.
- Automation bias: even physicians who completed dedicated AI-literacy training remained susceptible to subtly erroneous AI recommendations (Qazi et al., 2026).
- Deskilling: outsourced cognitive effort erodes competence gradually, one outsourced formulation or note at a time, rather than announcing itself.
- Each of these risks accumulates quietly. A skewed output, an accepted error, or a faded skill is ordinarily discovered after the fact, not while it's happening.

How Much Should the Tool Do Without Asking?

The risks that scale with autonomy



- Misaligned goals: what a system optimizes for may differ from what good clinical care requires.
- Unintended consequences: some harms surface only after a tool is deployed at scale.
- Lost accountability: each task handed to a system removes a point where a psychologist can catch an error.
- Compounding bias: an unreviewed output becomes the input to the next step.

Rules That Already Apply

- HIPAA and HITECH: any tool that accesses, processes, or stores protected health information requires a signed business associate agreement. Many general-purpose platforms do not offer one.
- FERPA: applies when a tool handles student records in training and education settings.
- APA Ethics Code (2017) and APA's 2025 ethical guidance on AI: the psychologist remains responsible for the work, however sophisticated the tool.
- Most of these rules predate the technology they now govern. That gap calls for caution now, before clearer rules arrive.



Diagnosis & Treatment Concerns

What a transcript can't capture

- Interpreting data in the context of the patient's needs — a model has no access to that context beyond what's typed into it.
- The need for human connection as part of what makes treatment work at all.
- Healing as a component of the relationship itself, not just the content exchanged within it.
- Creating a safe space involves neurobiological events between two people that a model cannot participate in.

Ethical Principles and Process

Principle-based ethics and intentional sequencing



Principle-Based Ethics

Psychologists incorporate these principles into practice directly — no separate model required



Autonomy

Respecting the patient's or supervisee's right to self-determination

Beneficence

Acting in ways that promote the welfare of others

Nonmaleficence

Avoiding harm, including harm that arrives gradually or indirectly

Fidelity to Relationship

Honoring the trust placed in the professional relationship

Justice

Ensuring fair and equitable treatment, including in how AI tools are applied across groups

Reason First

Why order matters

- Reason independently before consulting a model. Output that arrives first becomes an anchor — your own analysis is reduced to a revision of it.
- Reasoning first also gives you something to compare the model's output against, which converts its contribution from a conclusion into a second opinion.
- Treat AI-generated material as an object of critical analysis, not a draft awaiting light editing. Verify any claim about statute, regulation, or the literature against a primary source.
- This is the structural safeguard that awareness alone cannot provide.



There Are Wrong Answers — But Not the Kind You Expect

A note on how this material is usually taught

- Many traditional ethics workshops highlight codes and clear-cut violations.
- Sometimes that approach becomes "gotcha" or unnecessarily esoteric.
- These are usually fear-based experiences, and fear is not a durable foundation for good clinical judgment.
- Not every ethical or clinical dilemma is a gut-wrenching, heart-pounding decision — most of the work is quieter than that.

A DIFFERENT APPROACH

In Most Ethical Dilemmas, We Construct a Solution

Instead of looking it up in a sacred psychology text.

Research suggests that most mental health professionals do not use a structured ethical decision-making process. Instead, they tend to rely on intuition, often because of time constraints and the pressure to resolve the situation quickly.

PART 4

Case Practice



Apply the principles before you apply the model

Clinical complexity alone does not constitute an ethical dilemma; an ethical dilemma requires competing values or principles.

Vignette 1 — Name the Principles

Solo Exercise 10 minutes

- What are the competing ethical principles in this scenario? In other words, which ethical principles conflict or create tension?
- Does this vignette raise strong moral or personal values for you?
- Does the vignette evoke any type of emotional response?
- Notice them without letting them be the basis for professional judgment.

Vignette 1 — Reaction and Resolution

General Questions after this learning experience

- What is your emotional response to the scenario, and how might that response shape the decision you're inclined to make?
- What is a potential path forward? Work this through yourself first — only then would it make sense to check your reasoning against an LLM's.

Vignette 1: A New Referral?

A psychologist had an intake appointment with a new client. As soon as he looked at the intake form, he realized she is the ex-wife of a former client. Therapy with the ex-husband was brief and focused on his wish to leave the marriage. The prospective client was using her maiden name and gave no indication she knew of her ex-husband's treatment with this psychologist.

When asked about the referral, she said her physician referred her for anxiety. During the session, she mentioned she is still in communication with her ex-husband because they share three children.

Vignette 2: The AI-Assisted Note

A psychologist is behind on documentation after a full day of sessions. She uses an LLM to listen to the session and turn her shorthand into full progress notes for three patients. Reviewing the drafts before signing, she notices that for one patient — a Black man who spent the session discussing financial strain within his extended family — the AI-drafted note repeatedly frames his account using language like "entitlement" and "avoidance."

She doesn't recall using that framing in her shorthand, and it doesn't match her sense of the session. The draft is otherwise well organized, and she is already running late for her next appointment.

Vignette 2 — The AI assisted Note

10-minute exercise

- Apply the same questions from Vignette 1 — competing principles, personal reactions, emotional response, and a path forward — before opening any AI tool.
- Then, using the prompts to help evaluate the dilemma.
- Compare: did reasoning first change what you noticed once you saw the model's response?

Constructing Prompts for Multiculturalism and Diversity

Extending RISEN toward cultural awareness

- Ask the model directly where its answer might reflect a majority-culture default. An answer that carries no warning has not been checked for bias.
- Request at least two alternative, culturally situated interpretations of the same clinical material. A single alternative just swaps one default for another.
- Name the population or context, within de-identification requirements, so the model is not reasoning from an unstated default.
- Treat the model's answer to “What am I missing?” as a starting point for your own reflection, not a completed cultural competency check.

Sample Prompt for Ethical Analysis with an Emphasis on Culture

“You are an expert in ethics, ethical decision-making, and clinical psychology. I am a psychologist licensed in [your state], working with [population or context, de-identified]. Review the following situation and my initial analysis. Identify the competing ethical principles, offer at least two culturally situated interpretations, and suggest a path forward consistent with the APA Ethics Code and [your state]’s licensing regulations. Then tell me where your answer may reflect majority-culture assumptions, what I am missing, and what additional information you would need.”

Sample Prompt for Ethical Analysis with Diversity

You are assisting a licensed psychologist who works with [population or context, de-identified]. Review the following [note / treatment plan / case formulation] and identify any language, assumptions, or interpretations that may reflect a majority-culture default rather than this patient's specific cultural, linguistic, or socioeconomic context. Offer at least two alternative, culturally situated interpretations of the same material. Then tell me where your own suggestions may carry the same default, and what you don't know about this patient that would change your answer."

What Are the Main Points?

Recap before we close

- We have to train ourselves to think about larger ethical principles first, before reaching for any tool.
- We need the ability to slow ourselves down prior to making good ethical decisions — speed is where most of today's risks live.
- We need to think through the ethical dilemma first, including our values, therapeutic approach, conceptualization of the patient in context of the therapeutic relationship, countertransference reactions, values, beliefs, emotional state, and risk tolerance.
- After our own contemplation, then move toward LLM model or human consultation.
- We are integrating ethics with clinical practice, and using AI for specific tasks.

PART 5

Integration and Close

What this means for how you practice



Limitations of LLMs & the Integrated Psychologist

- Automation Bias: Fluency does not equal competency.
- De-skilling is an over-reliance on technology. Convenience is not competence, and the gap between them doesn't announce itself.
- Professional judgment and accountability remain with the psychologist regardless of what informed the clinical or ethical decision — the Ethics Code and Practice Guidelines assign that accountability unambiguously.
- Ethical framework limitations: garbage in, garbage out (GIGO) applies to ethical reasoning as much as to any other output — a model given a biased or incomplete framing will reason from it uncritically.

What This Means for Psychologists

- We are fortunate to have a human-centric profession — the relationship, not the output, is the mechanism through which the work succeeds. No LLM update will change this.
- We can become better collaborators with ethical, well-governed LLMs without treating them as co-clinicians.
- We need to understand which tasks to delegate and which tasks to keep.
- The goal is to integrate the parts of these tools that make us more efficient and effective, while keeping accountability, verification, and the therapeutic relationship squarely in human hands.

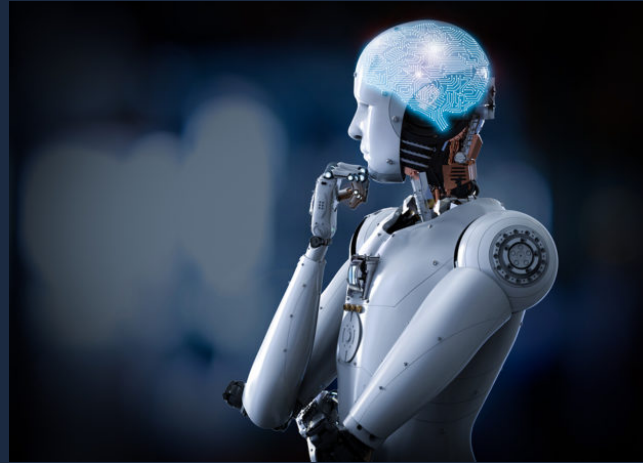
Revisiting this slide: Evaluating any AI application against three questions

Stewart Cooper and the Division 29 Work Group

- Does the tool enhance or erode the human relationships the work depends on?
 - Does the tool preserve psychologist accountability?
 - Does the tool respect the patient's narrative integrity and cultural context?
-
- Evaluate a tool before it becomes embedded in practice, because a tool is far easier to decline than to remove.

Wrapping Up

- Questions
- Critiques
- Concerns



Please complete the course evaluation.



References

- American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct* (2002, amended effective June 1, 2010, and January 1, 2017). <https://www.apa.org/ethics/code>
- American Psychological Association. (2025). *Ethical guidance for AI in the professional practice of health service psychology*. <https://www.apa.org/topics/artificial-intelligence-machine-learning/ethical-guidance-professional-practice.pdf>
- Bai, X., Wang, A., Sucholutsky, I., & Griffiths, T. L. (2025). Explicitly unbiased large language models still form biased associations. *Proceedings of the National Academy of Sciences*, 122(8), e2416228122. <https://doi.org/10.1073/pnas.2416228122>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610–623). *Association for Computing Machinery*. <https://doi.org/10.1145/3442188.3445922>
- Farmer, R. L., Lockwood, A. B., Goforth, A., & Thomas, C. (2024). Artificial intelligence in practice: Opportunities, challenges, and ethical considerations. *Professional Psychology: Research and Practice*, 56(1), 19–32. <https://doi.org/10.1037/pro0000595>
- Qazi, I. A., Ali, A., Khawaja, A. U., Akhtar, M. J., Sheikh, A. Z., & Alizai, M. H. (2026). Automation bias in large language model–assisted diagnostic reasoning among physicians trained in AI literacy — A randomized clinical trial. *NEJM AI*, 3(5). <https://doi.org/10.1056/aioa2501001>
- Society for the Advancement of Psychotherapy. (2026, September 7). Artificial intelligence and psychotherapy: Opportunities, challenges, and recommendations [Report of the Presidential Work Group on AI and Psychotherapy]. <https://societyforpsychotherapy.org/artificial-intelligence-and-psychotherapy-opportunities-challenges-and-recommendations/>