



ARTICLE

Redesigning Empathy Maps with GenAI

Tracy L. M. Norton

Assistant Professor of Law

Louisiana State University, Paul M. Hebert Law Center

Published: September 2026

Introduction

Legal writing pedagogy runs on individualized feedback. We know it works. Students who receive it outperform those who don't, and the effect is strongest for the students who need the most support.¹ But here is the part we talk about less: differentiated feedback requires differentiated data. You cannot calibrate your comments to a student's needs if your intake data cannot distinguish one student from another.

I have been collecting first-day survey data from my LRW students for over twenty years. Names, hometowns, educational backgrounds, writing experience, goals. The information was useful for building rapport in the first week. It was not useful for sustained, individualized teaching.

In 2021, I read Jorge Valenzuela's *Edutopia* article on using empathy maps to connect with students, and I saw the potential immediately.² Empathy maps are structured profiles, borrowed from UX design, that organize what you know about a person

¹ Daniel Schwarcz & Dion Farganis, *The Impact of Individualized Feedback on Law Student Performance*, 67 J. LEGAL EDUC. 139 (2017).

² Jorge Valenzuela, *How a Simple Visual Tool Can Help Teachers Connect With Students*, EDUTOPIA (Apr. 12, 2021), <https://www.edutopia.org/article/how-simple-visual-tool-can-help-teachers-connect-students>.

around what they say, think, do, and feel. Dave Gray developed the empathy map at XPLANE, a design consultancy, and published it in his 2010 book *Gamestorming* as a tool for understanding electronic game users through four lenses: what they say, think, do, and feel.³ The framework spread quickly through UX design and product development, then migrated into education. Valenzuela had adapted them for K-12 classrooms as a tool for understanding students holistically, moving beyond test scores to capture motivations, strengths, struggles, and emotional states.⁴

I thought: what if I converted my survey data into empathy maps that I could build on across the semester? The survey data would no longer be read at the beginning of the semester and filed away; it could become a living teaching tool, helping me calibrate feedback and conference approach across the semester.

I started building my first empathy maps that fall of 2021. The concept worked, but the labor was punishing. Creating 40 profiles by hand, then updating each one after every conference round, was not sustainable alongside a full teaching load. The profiles decayed. By mid-semester, they no longer matched the students sitting across from me.

In 2023, generative AI offered the first breakthrough. I used it to automate the initial profile creation, exporting the Google Forms survey data first into a spreadsheet.⁵ From there, I set up the empathy map format in Google Docs and used a mail merge function to create a one-page empathy map for each student.⁶ The result was an empathy map for each student that not only contained the data from the first-day survey but also left room for me to make comments after each conference. This process took a few hours, compared to the weeks that it had previously taken me to create the empathy maps when I first started. That worked for two years. But the profiles stayed static. Updating them after each conference meant reopening each document by hand, and I rarely had time. By mid-semester they were again stale.

Having reached an impasse again with stale empathy maps—albeit empathy maps that were far easier to create—I returned to generative AI to troubleshoot the empathy maps

³ DAVE GRAY, SUNNI BROWN & JAMES MACANUFO, *GAMESTORMING: A PLAYBOOK FOR INNOVATORS, RULEBREAKERS, AND CHANGEMAKERS* (2010).

⁴ Valenzuela, *supra* note 2.

⁵ For the first attempt at using generative AI to help me create empathy maps, I used ChatGPT.

⁶ Generative AI coached me through this entire process. I started with a very basic prompt asking for the most efficient way to start with Google Forms responses and end up with a one-page empathy map for each student. Each time the GAI gave me a solution that was beyond my skills, I asked for a simpler alternative until we finally reached one that I could implement.

so they could become a living document throughout the semester, or even—I hoped—the full academic year that the course ran.

Because I have had considerable success using Claude Projects, I chose that for my generative AI platform. Claude is a generative AI assistant built by Anthropic, comparable in kind to ChatGPT or Gemini. You interact with it by typing in plain English the way you would brief a research assistant. It drafts and edits text, summarizes long documents, answers questions, works through analysis, and searches the web for current information. Anthropic positions Claude as particularly strong on writing, reasoning, and careful handling of professional material.⁷

A Claude Project is a shared workspace inside Claude where you keep everything related to a single line of work in one place. You upload the documents you want Claude to draw on, a syllabus, a rubric, a set of prior drafts, and you write standing instructions that tell Claude how to behave every time you work there. From then on, every conversation you start inside that project can see those files and follow those instructions, so you do not re-explain your context or re-attach your materials each session. Think of it as a folder that also remembers how you like things done: the work you do in it persists over time and stays available across your devices.

I started a Claude Project with my existing questionnaire, the Google Doc template I'd been using, and a brief explanation of the problem I was trying to solve—the friction of updating the map throughout the semester. I then prompted Claude to integrate everything into a new system.

Claude didn't tell me how to solve my stale empathy map problem. It told me my questionnaire was wrong. I was asking the wrong questions to surface the information I wanted.

Somewhere in that exchange, a better empathy map stopped being the point. What I had found instead was a tireless collaborator, one willing to push back on assumptions I had stopped questioning years ago. The assumption I trusted most was that my intake

⁷ A paid Claude subscription is not necessary for this. Most modern generative AI platforms can do this work. What matters is not the tool but whether the platform, and your use of it, complies with FERPA. An institution standardized on Google Workspace might prefer Gemini or Gemini Notebook. A Microsoft campus like mine might prefer Copilot, which operates inside the same licensed environment as the Forms instrument feeding the pipeline. Confirm your institution's data-handling rules before entering real student information into any platform, or test with de-identified sample data first.

questionnaire was already telling me what I needed to know about my students. It was not.

1. The questionnaire problem

Claude asked to see not only the questionnaire but a representative empathy map. It even reminded me to redact identifying information like name, phone number, or email address to ensure I was complying with FERPA.⁸

My intake survey has changed over the past 25+ years, but it generally always asks for some basic rapport-building information like hometown and current favorite song along with educational background. I also ask some specific questions about their writing experience as well as their experience receiving feedback on writing.

I shared the questionnaire and the sample with Claude. The sample was for a student I will call Elena Castillo: a political science and art history double major from a Louisiana university, graduated in 2024, who had written several long papers including a senior thesis. She was motivated, articulate, and clearly engaged. On the writing skills portion of the survey, she reported being comfortable with every concept on the checklist and uncomfortable with none.

The empathy map I had built from her data was detailed. I knew her hometown, a song she felt represented her, her practice area interests, her educational background. I could open a conference with a personal connection. But the writing skills section told me nothing actionable. Elena looked identical to half the class.

The problem ran deeper than missing data. When I conferenced with students like Elena and read their early submissions, the self-assessments often turned out to be wrong. Students who reported comfort with organization submitted work structured by case rather than by issue. Students who reported comfort with clarity and conciseness produced long, tangled sentences. The instrument was not just undifferentiated. It was actively misleading.

Claude identified the structural flaw: checkbox self-assessment produces ceiling effects. Students entering law school lack the metacognitive framework to evaluate their own

⁸ I cannot say for sure that Claude would naturally warn me about FERPA. However, I've given it custom instructions that serve as a prompt add-on for all work that I do in Claude, and those instructions tell Claude that I am a law professor working at a state university, so Claude often raises issues that arise from that identity. Further, I use a paid version of Claude that affords some additional privacy protections, but it is not an enterprise version with negotiated privacy terms, which is what I would need to feel comfortable using identified student data.

writing skills in a domain they have not yet practiced.⁹ They are not being dishonest. They genuinely do not know what they do not know. A student who wrote a strong senior thesis in art history has real skills, but she cannot yet tell you which of those skills will transfer to legal analysis and which will not. The checkbox gives her no reason to try.

I had asked for better plumbing. Claude told me I needed electrical work first.

2. Redesigning the questionnaire: the human-AI collaboration

The questionnaire redesign emerged from the conversation with Claude that followed. I brought twenty years of conference experience and a clear sense of what information helps me differentiate instruction. Claude brought a structural diagnosis: the instrument was designed to collect self-assessment data, but self-assessment was exactly what incoming law students could not reliably provide. Together, we rebuilt it around three principles.

2.1 Forced ranking replaces checkbox self-assessment.

The redesigned questionnaire asks students to rank six writing skills from strongest to weakest: organizing ideas into a logical structure, expressing complex ideas clearly, conducting research, analyzing and applying rules to facts, editing and proofreading, and following citation requirements. The instructions are direct: "Be honest. Everyone has a bottom of the list." The format is a grid where each skill gets exactly one rank and each rank gets exactly one skill.

The difference is immediate. On the old instrument, a student could check "comfortable" for all six skills and move on. On the new one, everyone must place something at five and six. A student who ranks "analyzing and applying rules to facts" near the bottom needs different early feedback than one who ranks "editing and proofreading" there. The ranking does not ask students what they are bad at. It asks them to make relative judgments, which even confident students can do honestly. The bottom two items become the starting signal for differentiation.

2.2 Behavioral and scenario questions replace comfort ratings.

Two examples. First, the writing process question: "When you wrote a major paper in undergrad, what did your process typically look like?" The four options range from outlining and revising in stages to writing close to the deadline under pressure. All of them sound neutral. None sounds like the wrong answer. But a student who writes in

⁹ Justin Kruger & David Dunning, *Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments*, 77 J. PERSONALITY & SOC. PSYCHOL. 1121 (1999).

one session the night before a deadline needs fundamentally different scaffolding than a planner who revises in stages. This is a behavioral question. It captures what students actually do, not what they think they should say.

Second, the setback question: "Imagine you receive a grade significantly lower than you expected on your first legal writing assignment. What would you most likely do first?" The five options range from "re-reading the feedback carefully" to "worrying about whether you belong in law school." This question is scenario-based and low-stakes, but it is remarkably predictive. Students who select the impostor-syndrome response or the fairness-focused response are not necessarily in trouble, but they are the students whose first round of graded feedback needs the most careful calibration. Students who say they want to talk to the professor are your conference-seekers. You know before the semester starts how to prepare for those first conversations.

2.3 Indirect elicitation replaces self-diagnosis.

The strongest question on the redesigned form is an open-ended prompt: "Think about the most critical or challenging feedback you've received on your writing. What did you do with it?" Students who describe reading the feedback, making a plan, and improving are signaling resilience and a growth orientation. Students who describe confusion, frustration, or avoidance are telling you how to calibrate your own delivery. Students who say they have never received critical feedback on their writing are telling you something important, too.

This is the design philosophy running through the entire instrument: differentiation data is a byproduct of honest answers, not a product of self-diagnosis. Students do not need to know what they are bad at. The questions are designed so that the useful information falls out of truthful responses to prompts that feel low-stakes and personal.

I want to be specific about credit. Claude identified the ceiling-effect problem and proposed the forced-ranking mechanism. I would not have framed the checkbox issue as a structural design flaw on my own, even though I had been living with its consequences for years. But I chose which skills to include in the ranking, which scenarios to pose, and how to frame every question so that it felt approachable rather than clinical. The redesign was a genuine collaboration: the AI made a diagnosis I had missed, and I made the pedagogical decisions that turned that diagnosis into a working instrument.

This process not only helped me improve my empathy map; it diagnosed and helped me resolve a problem that was not even on my radar. I lacked the survey design expertise to know how to solve the problem, but I did have the expertise to recognize

the solution once the problem and solution were identified. This is exactly the kind of partnership in which generative AI shines.

3. From intake data to living profile: the empathy map pipeline

With the redesigned questionnaire feeding better data, we turned to the problem of how to keep the empathy map growing with me and the student throughout the year. The new pipeline works like this. Students complete the survey in Microsoft Forms at the start of the semester.¹⁰ Their responses flow into an Excel spreadsheet and from there, through a mail merge process, into a one-page (front and back) empathy map for each student.¹¹

I could have stopped here. The improved questionnaire already makes even static empathy maps more useful than they were. However, Claude did help me improve the map and the process I use to update it. To do this, I return to my Claude project after I have the empathy maps completed, and I input the data to Claude, one student at a time with identifying information redacted.¹² Claude then synthesizes the raw data into three outputs for each student.¹³

The first is a narrative written in second person as a teaching tool: "You're working with a student who..." It covers what the student likely brings to the table, where the gaps or risk areas probably are, how to calibrate feedback delivery, and one or two rapport hooks from their personal information. The second is a set of differentiation tags drawn

¹⁰ I use Microsoft Forms rather than Google Forms throughout this new pipeline. Claude suggested the switch. LSU is a Microsoft campus, and Forms, along with the rest of the Microsoft 365 suite, operates inside the institution's licensed environment, which is covered by LSU's data-handling agreements for student educational records under FERPA. Keeping intake data within the university-sanctioned toolset avoids routing student responses through a consumer platform the institution has not vetted.

¹¹ To make them easier to navigate, I've chosen to put all empathy maps into a single Word doc, each map tagged with the student number (which I assign) and name so that I have a clickable menu in the Word navigation. Other options I considered were an Excel workbook with each empathy map a different spreadsheet tab; a folder of Word docs, one document for each student; and a Notion database. The specific platform used is less important than the utility to your workflow. Here, as with many things, the best tool is the one you'll use.

¹² Again, here there are a number of ways to do this. You can simply delete the information from the map before you input it for synthesis. I will redact it using Adobe Acrobat because that will work faster within my workflow. One caution about that method: a black box drawn with Acrobat's comment or drawing tools covers the text visually but leaves it in the file, where a generative AI platform will read it. Only the Redact tool, followed by applying the redactions and saving the file, removes the underlying text. The synthesis prompt I use begins by checking the file for identifying information for this reason.

¹³ The prompt I use to synthesize the initial intake data, along with the prompt I use to update each profile after a conference, can be found at <https://perma.cc/VP4N-ESAK>.

from a controlled list: Conference Priority, Feedback-Sensitive, High Autonomy, Process Support, Confidence Building, Citation Focus, and others. The third is a short set of resource recommendations matched to the student's profile, referencing the kinds of materials available in a lawyering skills course: research guides, writing mechanics refreshers, sample documents, citation checklists, flowcharts, office hours.

Claude produces a draft. I review every profile. I adjust where my professional judgment says the synthesis missed something or over-weighted a single data point. The AI does the labor-intensive first pass, the work I would do myself if I had unlimited time. Because I have this information at the beginning of the semester, I can review it to get a sense of the class as a whole even before we start meeting. I can also use this information to help me form the teams students will work in throughout the semester. Rather than form teams only with a variety of undergrad majors, hometowns, and other characteristics, I can form teams with a variety of working styles, feedback preferences, and independence.

A note on student data. This system processes survey responses through a generative AI platform, which raises reasonable questions about privacy. Two safeguards. First, use only platforms configured not to retain conversation data or use it for model training. Most major providers offer this setting; check before you start. Second, and more effective: anonymize the data before it reaches the AI. Replace student names with identifiers, Student A, Student B, or numeric codes. Claude does not need to know who the student is to synthesize a teaching profile. You reconnect the profile to the student in your own records after the synthesis is complete. The AI never sees identifiable information.

4. -The evolving profile: mid-semester refresh

The initial profile is useful, but the updated profile is where the system earns its place. I'll record conferences this year for all students who would like to have a recording, and this will help me update the empathy map after each conference.¹⁴ After each conference round, I will upload either a transcript of the recording or an AI summary of the recording and ask Claude to generate a post-conference narrative with this new information, noting important adjustments to how the student works and receives feedback. For students who opt out of recording, I'll create a voice note of my observations and upload a transcript of that instead of a conference recording. The post-

¹⁴ Of course, a student's name is likely to be spoken during the conference, the transcript must be read and the student's name removed before uploading the transcript or summary to Claude or any other platform.

conference prompt instructs Claude to flag divergences between the intake predictions and observed performance.¹⁵

This is what GenAI makes possible: The capacity to synthesize a semester's worth of observations into an updated teaching profile in minutes rather than hours. In 2021, I updated my Google Doc empathy maps when I could, which was rarely. By mid-semester they were stale. The new system is designed to keep up with what I already know about my students but cannot hold in working memory across forty profiles simultaneously.

5. What transfers, with or without AI

Not every legal writing instructor wants to create or update empathy maps or to run deidentified student data through a GenAI pipeline. The questionnaire redesign principles transfer without any of the technology.

Replace checkbox self-assessments with forced ranking. Ask behavioral questions instead of comfort ratings. Design prompts where the differentiation data falls out of honest answers rather than requiring students to diagnose their own weaknesses. Any instructor can make these changes to a Google Form in an afternoon and start the next semester with dramatically better intake data. The empathy map concept itself requires nothing more than a willingness to organize what you know about each student into a format you can reference and update.

For instructors who are ready to experiment with GenAI, though, the empathy map pipeline is a low-risk starting point. The AI is instructor-facing, not student-facing. It never interacts with students. It never grades. It never makes pedagogical judgments. It synthesizes data you already collect into a format you can act on, and you review every output before it informs your teaching. With deidentified data, the privacy concerns are manageable.

Conclusion

The lesson for me was broader than what I learned about designing the questionnaire and updating the empathy map. This is one model for what instructor-facing AI can look like in practice: not a replacement for professional judgment, but a way to do work that professional judgment demands and teaching schedules do not always permit.¹⁶ I

¹⁵ See *supra* note 13.

¹⁶ Ethan R. Mollick & Lilach Mollick, *Using AI to Implement Effective Teaching Strategies in Classrooms: Five Strategies, Including Prompts* (Wharton Sch. of the Univ. of Pa. Research Paper, 2023), <https://ssrn.com/abstract=4391243>.

plan to deploy this updated system with my Fall 2026 sections. The mid-semester refresh may work differently in practice than it does in development, and the AI-generated profiles may introduce problems I have not anticipated. But the system has already proved its value in ways I did not expect. I asked GenAI to help me automate a workflow. It helped me see a flaw in my instrument that I had lived with for years. The redesigned questionnaire produces better data regardless of what happens with the pipeline. The empathy map framework gives me a place to organize what I learn about each student across a semester. Even if every other piece of the technology changes tomorrow, those gains are real.

My hope when I started experimenting with GenAI was that it would make possible the teaching I never had time for. So far, it has. I am excited to discover what else becomes possible.