Research-Quals: A Faculty Guide

Our PhD students are going to use AI for research. Here's one attempt at a structured response — and an invitation to build something better together.

The Situation We're All In

AI tools can now generate polished research outputs: literature reviews, data analysis, theoretical framings, even full manuscript drafts. Our students know this. Some are already using these tools, with or without guidance.

The choices we face aren't as simple as "ban it or allow it." There are at least two independent dimensions that matter:

How We Think About AI Access and Skill Development

	Skill verification implicit (apprenticeship, osmosis)	Skill verification explicit (structured assessment)
AI access open (use what you want)	Students learn by doing, with all the risks and upside that implies. Works well for self-directed learners with strong mentorship. Fails silently when mentorship is thin.	Students demonstrate competency through structured challenges, then use AI freely. Adds overhead but creates a legible floor.
AI access graduated (earn tool access)	Traditional apprenticeship plus tool restrictions. "You can use X after your second year." Familiar but	Competency gates tied to specific capabilities. Most structured, most overhead, clearest accountability.

	Skill verification implicit (apprenticeship, osmosis)	Skill verification explicit (structured assessment)
	hard to calibrate — what counts as "ready"?	

Most of us are operating in the top-left quadrant by default — implicit skill development, open access — because we haven't built the infrastructure for anything else.

Research-quals is an experiment in moving rightward: making skill verification explicit. It doesn't prescribe where you land on the access dimension. That's a judgment call that depends on your students, your program, and your tolerance for risk.

We don't claim this is the right answer. It's a starting point — one researcher's attempt to operationalize the tacit knowledge that expert researchers bring to this work. We're actively looking for collaborators who see the problem differently.

What Research-Quals Is

A competency-based training system for mixed-methods management research. Seven provisional domains. Three levels. Real data.

Domain	The Question It Answers
1. Pattern Recognition	Can you tell signal from noise?
2. Theoretical Positioning	Can you frame a contribution?
3. Research Design	Can you match design to question?
4. Qualitative Rigor	Can you code defensibly?
5. Literature Synthesis	Can you position in a conversation?
6. Adversarial Evidence	Can you find disconfirming data?
7. Claim Verification	Do your claims match your evidence?

Where these domains came from: These seven emerged from one researcher's (Matt Beane, UCSB) experience operationalizing what seemed to matter when building AI-assisted research tools — specifically, the places where AI went wrong and human judgment was essential. They draw on existing methods literature and three months of building tools and finding every way they fail. They are not exhaustive. They may not be the right decomposition. They're a working hypothesis about what matters, offered as a starting point for collective refinement.

Three Levels Per Domain

Foundation: Can you identify good practice when you see it? Conceptual exercises. Unlimited retakes — this is learning.

Practice: Given real data, can you do it yourself? Work on actual research problems. Detailed feedback against expert baselines.

Mastery: Can you do it independently and defend your choices? Single attempt. This is where it counts.

The Capstone

Two options (students can do one or both):

AI Output Supervision: Review an AI-generated paper and catch all the issues. The AI draft is deliberately produced without safeguards — genre violations, overclaimed contributions, hallucinated evidence, buried disconfirmation. Can the student find what's wrong?

Induce the Problem Space: Given an unfamiliar dataset and the task "build an AI-assisted analytical pipeline for this," the student must first *induce* the categories of skill and judgment required to do this work well — before seeing the seven domains, before seeing any existing tool's command list. What are the areas where human judgment is essential? What can be safely automated? What requires supervision? Then they build a pipeline informed by their own decomposition and defend it.

This is harder than copying someone else's pipeline. The point — inspired by Paul Leonardi's insight — isn't "can you build a machine?" (anyone can copy one). It's "can you derive what the machine *should* do from first principles, given a research problem you haven't seen before?" That's the skill that makes a researcher irreplaceable.

Why This Might Matter for Our Programs

The Apprenticeship Model Has Gaps

Most PhD training in research craft happens through osmosis: watch your advisor, absorb tacit knowledge, hope you pick up the pattern recognition. Some students get great mentorship. Many don't. Luck plays too large a role in determining who develops these skills.

Research-quals tries to make some of the implicit explicit. Whether these seven domains are the right decomposition is genuinely open — but the attempt to name what expert researchers do intuitively seems worth pursuing.

AI Changes the Stakes

A student who can't recognize a hallucinated quote shouldn't be using tools that routinely hallucinate. A student who doesn't understand genre conventions shouldn't rely on tools that default to hypothesis-testing framing for every paper. A student who's never manually coded data can't evaluate whether AI coding is any good.

Some kind of floor seems necessary. Research-quals proposes one. We'd love to hear what you think the floor should look like.

What Our Students Are Telling Us

In a recent CSS Lab session at UCSB (February 2026), PhD students consistently expressed two things:

1. They want explicit instruction on research craft. One student said: "I've been begging for explicit instruction on how to do these things for years."

2. They want permission to use AI responsibly, not a blanket ban.

These aren't just our students — we suspect this is widespread.

How You Might Use It

For an Individual Advisor

1. **Have your student look at the domains.** Read through `competencies/` and see if the decomposition resonates.
2. **Start with Domain 1 (Foundation).** Takes 2-4 hours. Low stakes, unlimited retakes.
3. **Review their Practice and Mastery work.** The system provides expert baselines for comparison, but your domain expertise adds irreplaceable context.
4. **Use the capstone as a milestone.** "Before you use AI tools on your dissertation data, demonstrate you can catch what they get wrong."

For a PhD Program

1. **Evaluate as a methods supplement.** Research-quals complements existing methods courses — it doesn't replace them. It focuses on applied judgment that courses often can't teach.
2. **Adapt the domains.** If your program emphasizes different skills, modify the decomposition. The framework is open source and designed to be forked.

3. **Contribute expert baselines.** The system improves with more expert perspectives. Faculty can contribute anonymized data from their research — all processing runs locally, data never leaves your machine.

Timeline (Approximate)

Stage	Time	What Happens
Foundation (all 7)	2-4 weeks	Conceptual understanding
Practice (all 7)	4-8 weeks	Applied skill on real data
Mastery (all 7)	4-8 weeks	Independent demonstration
Capstone	1-3 weeks	Integration + AI supervision
Total	**3-6 months**	Full competency

Students with prior qualitative methods training can test out of Foundation immediately.

What It Connects To

Research-quals is part of a developing ecosystem:

- **Theory-forge**: An AI-assisted pipeline for producing theory-building papers. One way to use research-quals is as a prerequisite system for theory-forge capabilities, though the two can also be used independently.

- **Living Paper**: A verification system that creates auditable links between claims and evidence. Addresses the "how do I know your data is real?" concern that reviewers will increasingly raise.

All are open source (MIT license). All are works in progress. All benefit from faculty feedback, critique, and contribution.

Honest Objections and Honest Responses

"My students should learn by struggling, not by following a checklist." Agreed — and the Practice and Mastery levels *are* struggle. Real analysis on unfamiliar data, not following instructions. The domains provide structure; the work within them is genuinely hard.

"I don't trust AI-based assessment." Foundation assessments are auto-scored (multiple choice + short answer). Practice and Mastery work is designed for advisor review. The system provides rubrics and expert baselines, but you make the judgment call.

"This will slow down my students." Possibly. The bet is that structured competency building prevents costlier problems downstream — desk rejections, retractions, or students who can produce papers but can't evaluate whether they're any good. Whether that tradeoff makes sense depends on your context.

"Why seven domains? Who decided these?" One researcher, based on one experience building AI research tools. The domains are a working hypothesis, not a settled taxonomy. If your expertise suggests different or additional domains, that's exactly the conversation we want to have. The "Induce the Problem Space" capstone was designed partly for this reason — we genuinely believe students might derive a better decomposition than ours.

"This feels like it's solving your problem, not mine." It might be. We built what we needed and are offering it in case others face similar challenges. If it doesn't fit, that's fine. If it almost fits but needs modification, we'd rather collaborate than prescribe.

Get Started

```
git clone https://github.com/mattbeane/research-quals.git
cd research-quals
```

Start by reading the domain descriptions in `competencies/` and the assessment specs in `assessments/` to see if this resonates with your program's needs.

> ***Note***: The `research-quals` CLI (for taking and submitting assessments) is in development. The curriculum content, assessment rubrics, and expert baselines are complete and usable now. CLI automation is coming.

Questions? Disagreements? Better ideas? Open an issue or contact mattbeane@ucsb.edu

---

Research-quals: Because "I took a methods course" isn't the same as "I can do research." And "I can prompt an AI" isn't the same as "I can supervise one."