

Appendix A: The Three-Point Diagnostic Protocol: A Complete Implementation Guide

Overview

This appendix gives you everything you need to run the three-point diagnostic protocol described in Chapter 1. The goal is formative assessment that actually teaches you something about your students, not just who is behind, but where and why. The protocol works in bar prep, upper-level review courses, or any doctrinal setting where you want real data about whether your interventions are working.

A quick reminder of how this works. Students take a 25-question diagnostic on a subject area. Civil Procedure is used throughout this appendix as the working example, but the protocol transfers to any bar-tested subject. They upload their answers to a GenAI platform, which analyzes their performance and generates personalized study materials targeted at their specific gaps. After studying those materials overnight, they take a parallel diagnostic weighted toward their individually identified weak areas. Three days later, they take a third parallel exam to measure what stuck.

Each section below walks through one phase of the protocol, with the specific materials you need: sample questions, word-for-word AI prompts, data templates, and guidance on what to do with the results.

Before You Begin: The IRB Question

Whether you need IRB review depends on what you plan to do with the data. There are two distinct use cases here, and they carry different obligations.

Using data solely to improve your teaching: Tracking aggregate class performance to identify where students are struggling and adjust your instruction generally does not require IRB approval. This is routine educational assessment, and most institutions treat it the same way they treat grading or formative quizzing. No submission is needed.

Publishing or presenting findings: Even at a conference, or even in a practitioner journal changes the calculus. Once you're generating knowledge intended for dissemination beyond your institution, the research/education line blurs, and your IRB will want to know about it.

The most likely pathway is an exemption under Category 1 of the Common Rule (45 C.F.R. § 46.104(d)(1)), which covers research conducted in established educational settings using normal educational practices. To qualify, you need to be testing an instructional strategy (this protocol qualifies), every student must receive the intervention (no control group that gets nothing), and you need to be able to demonstrate that participation does not affect grading. If all three conditions are met, the review is typically straightforward and fast.

My recommendation: send a brief email to your IRB coordinator before you start, describe the protocol, and ask whether it qualifies for an exemption. Most coordinators are happy to give informal guidance. It takes fifteen minutes and avoids a much longer conversation later.

Practical note: The exemption process is usually quick, often a two-page form and a response within a week. Even if you're only planning to use the data internally, the exemption documentation protects you if you later decide to share findings. Get it first.

Phase 1: The Baseline Diagnostic

Administration Instructions

Administer the baseline diagnostic during a class session or as a brief out-of-class assignment completed under timed conditions. Students should not use notes, outlines, or GenAI during the diagnostic. The point is to capture what they actually know, not what they can look up.

Suggested parameters:

- 25 questions, multiple choice, MBE format
- 50 minutes (2 minutes per question, standard bar exam pace)
- No materials permitted
- Students record their answers on a simple answer sheet (sample below) and retain a copy of their responses. They will upload this to GenAI in Phase 2

Sample Diagnostic Questions: Civil Procedure

The following fifteen questions illustrate the type and difficulty level appropriate for a bar prep diagnostic. A full 25-question set should cover all major Civil Procedure topics in proportion to their MBE weight. The current MBE allocates approximately 25% to jurisdiction and venue, 25% to pleading and joinder, 25% to discovery and disclosure, and 25% to pretrial, trial, and appellate procedures.

Questions 1–5: Jurisdiction and Venue

1. Paula, a citizen of Texas, sues Dan, a citizen of Oklahoma, in federal district court in Texas. Paula seeks \$60,000 in compensatory damages for breach of contract. Dan moves to dismiss for lack of subject matter jurisdiction. The court should:

- (A) Grant the motion, because the amount in controversy does not exceed \$75,000.
- (B) Grant the motion, because federal courts only have jurisdiction over federal questions.
- (C) Deny the motion, because the parties are citizens of different states and the amount meets the threshold.
- (D) Deny the motion, because the case arises from a contractual dispute.

Answer: (A). The diversity jurisdiction statute requires the amount in controversy to exceed \$75,000, not merely equal it. \$60,000 falls short. Note that complete diversity does exist (*Texas v. Oklahoma*) but it is insufficient without meeting the amount requirement.

2. Allison, domiciled in California, files suit in federal court in Nevada against two defendants: Baker Corp., incorporated in Delaware with its principal place of business in Nevada, and Carter Inc., incorporated in Nevada with its principal place of business in California. Allison seeks \$200,000 in damages. Does the court have diversity jurisdiction?

- (A) Yes, because Allison is a citizen of a different state than at least one defendant.
- (B) No, because Baker Corp. is incorporated in Delaware, not California.
- (C) No, because Carter Inc. is a citizen of California, the same state as Allison.
- (D) Yes, because the amount in controversy exceeds \$75,000.

Answer: (C). Complete diversity requires that no plaintiff share citizenship with any defendant. A corporation is a citizen of both its state of incorporation and its principal place of business. Carter Inc. is a citizen of Nevada (state of incorporation) and California (principal place of business). Since Allison is also a citizen of California, complete diversity is destroyed.

3. Doug operates a mail-order business out of his home state of Idaho. He has never advertised in, shipped goods to, or had any physical presence in Florida. A Florida consumer, relying on his website, placed an order that resulted in injuries. The Florida consumer sues Doug in Florida state court. Under the Due Process Clause, may Florida exercise specific personal jurisdiction over Doug?

- (A) Yes, because Doug's website is accessible in Florida.
- (B) Yes, because the plaintiff is a Florida resident who suffered injury there.
- (C) No, because Doug has not purposefully availed himself of the Florida forum.
- (D) No, because a website is insufficient to establish any form of jurisdiction.

Answer: (C). Specific jurisdiction requires that the defendant purposefully avail itself of the forum state, such that being hauled into court in that state is reasonably foreseeable. A passive website accessible everywhere, without deliberate targeting of Florida, does not satisfy purposeful availment under the standard derived from *World-Wide Volkswagen* and refined in *J. McIntyre Machinery*. The plaintiff's location is irrelevant to whether the defendant directed activity at the forum.

4. Plaintiff files a federal case in the Southern District of New York. Defendant moves to transfer to the Northern District of California under 28 U.S.C. § 1404(a), arguing that all witnesses and relevant documents are located there and that the case has no connection to New York. Plaintiff opposes transfer. What standard governs the transfer decision?

- (A) Transfer is available only if venue in the original district is improper.
- (B) Transfer is available when it would serve the convenience of parties and witnesses and the interest of justice, even if venue in the original district is proper.

(C) Transfer requires the defendant to show that the original forum is so inconvenient as to constitute a due process violation.

(D) Transfer is available only with the plaintiff's consent.

Answer: (B). Section 1404(a) allows transfer to any district where the action might have been brought, or to which all parties consent, for the convenience of parties and witnesses and in the interest of justice. Critically, the original venue need not be improper. § 1404(a) is a convenience transfer, distinct from § 1406(a), which applies when venue is improper.

5. A civil action was properly filed in state court. Defendant receives the summons and complaint on March 1. The forum state's long-arm statute permits removal to federal court. If defendant wants to remove this case to federal court, by what date must the notice of removal be filed?

(A) 30 days after service of the summons and complaint.

(B) 60 days after service of the summons and complaint.

(C) Any time before the case proceeds to trial in state court.

(D) Any time within one year of the commencement of the action.

Answer: (A). Under 28 U.S.C. § 1446(b), a defendant must file a notice of removal within 30 days of receiving the initial pleading or summons, whichever is first. If the case was not initially removable but becomes removable later (e.g., due to a change in the amount in controversy), the 30-day clock runs from receipt of the document making it removable, subject to the one-year limit for diversity cases.

Questions 6–10: Pleading and Joinder

6. Plaintiff files a complaint for negligence that states: "Defendant negligently operated his vehicle and struck Plaintiff, causing injuries." Defendant moves to dismiss under Rule 12(b)(6). Under the pleading standard established in *Twombly* and *Iqbal*, how should the court rule?

(A) Deny the motion, because notice pleading only requires a short and plain statement of the claim.

(B) Grant the motion, because the complaint states only a legal conclusion without factual support.

(C) Deny the motion, because the complaint gives defendant fair notice of the claim and the grounds upon which it rests.

(D) Grant the motion, because the plaintiff has not alleged all elements of negligence with specificity.

Answer: (C). *Twombly* and *Iqbal* require factual allegations sufficient to state a plausible claim: not just possible, but plausible. "Negligently operated his vehicle and struck Plaintiff"

is a thin pleading but likely survives because it alleges a specific act by a specific person with a specific result. It is meaningfully different from the stripped conclusory allegations Iqbal rejected. Courts have generally held that basic negligence complaints survive 12(b)(6) when they allege a specific act, a specific actor, and a specific harm.

7. After filing her Answer, Defendant realizes she has a claim against her co-defendant for contribution. She wants to bring that claim in the current lawsuit. Which procedural device should she use?

- (A) A counterclaim under Rule 13(a).
- (B) A cross-claim under Rule 13(g).
- (C) A third-party complaint under Rule 14.
- (D) An amended pleading under Rule 15.

Answer: (B). Rule 13(g) governs cross-claims against a co-party (a party on the same side of the litigation). The contribution claim here is against a co-defendant, making it a cross-claim. A counterclaim under Rule 13 would run against the opposing party. A third-party complaint under Rule 14 brings in a new party who is not yet in the lawsuit. Note that cross-claims are permissive, not compulsory.

8. Plaintiff sued Defendant in federal court for breach of contract. After judgment for Defendant, Plaintiff sues again in state court on a fraud theory arising from the same transaction. What doctrine governs Defendant's defense?

- (A) Res judicata (claim preclusion), which bars re-litigation of all claims that were or could have been brought in the prior action.
- (B) Collateral estoppel (issue preclusion), which bars re-litigation of issues actually litigated and necessarily decided in the prior action.
- (C) The full faith and credit clause, which only applies to state court judgments.
- (D) Stare decisis, which binds the state court to the federal court's legal conclusions.

Answer: (A). Res judicata bars a second action on claims arising from the same transaction or occurrence as a prior action, even if the plaintiff advances a different theory in the second suit. The fraud theory arose from the same contractual relationship and could have been brought in the original action. Issue preclusion (B) applies only to issues actually litigated. Since fraud was not litigated in the first action, collateral estoppel would not reach it.

9. Twenty plaintiffs, all customers of the same bank, want to bring a class action against the bank for a uniform fee that the bank charged without disclosure. What requirement is most likely to be contested in determining whether to certify the class?

- (A) Numerosity: whether the class is large enough to make joinder impracticable.

- (B) Commonality: whether the class members share common questions of law or fact.
- (C) Typicality: whether the named plaintiff's claims are typical of the class.
- (D) Predominance: whether common questions predominate over individual ones.

Answer: (D). For a damages class under Rule 23(b)(3), plaintiffs must satisfy not only the four Rule 23(a) prerequisites (numerosity, commonality, typicality, adequacy) but also predominance and superiority. The bank will argue that individual questions (whether each customer actually incurred the fee, whether they had notice, whether damages differ), predominate over the common question of whether the fee was unlawfully undisclosed. With only twenty plaintiffs, numerosity may also be questioned, but predominance is typically the hardest certification hurdle in consumer cases.

10. Plaintiff timely served Defendant with a complaint but forgot to serve one of three defendants. Forty days after filing, Plaintiff moves for an extension of time to complete service on the omitted defendant. The Federal Rules of Civil Procedure provide a presumptive service deadline of 90 days from filing. How should the court analyze the motion?

- (A) Deny the motion because the 90-day deadline is jurisdictional and cannot be extended.
- (B) Deny the motion because fewer than 90 days have elapsed and no extension is yet needed.
- (C) Grant the motion only if Plaintiff shows good cause for the failure.
- (D) Grant or deny based on whether good cause exists, but the court also retains discretion to extend time absent good cause.

Answer: (D). Rule 4(m) requires dismissal without prejudice if service is not made within 90 days, unless the plaintiff shows good cause. But the Supreme Court's decision in *Henderson v. United States* and subsequent circuit decisions confirm that courts retain discretion to extend the time even absent good cause. In any event, the motion here is premature. Only 40 days have passed, and the 90-day window is still open. However, (D) correctly states the legal standard.

Questions 11–15: Discovery, Summary Judgment, and Appellate Procedure

11. Defense counsel receives an interrogatory requesting "all documents relating to the incident of March 15." The request seeks documents that include a report prepared by an accident reconstruction expert whom Defendant hired to advise litigation counsel on trial strategy. Must Defendant produce the report?

- (A) Yes, because the interrogatory encompasses all documents related to the incident.
- (B) No, because the report is protected by attorney-client privilege.

(C) No, because the report constitutes attorney work product.

(D) Yes, because expert reports are never protected from discovery.

Answer: (C). A report prepared by a consulting expert (one hired for litigation strategy, not expected to testify) is protected as opinion work product under Rule 26(b)(3) and (4). It is not shielded by attorney-client privilege, which covers communications between attorney and client, not third-party experts. Testifying experts have different disclosure obligations under Rule 26(a)(2), but a non-testifying consulting expert's report is generally protected absent exceptional circumstances.

12. Defendant moves for summary judgment, attaching an affidavit from a key witness and documentary evidence that supports Defendant's version of events. Plaintiff responds with a brief stating that the evidence "raises genuine issues" but submits no affidavits, declarations, or other evidence. How should the court rule?

(A) Deny the motion because summary judgment is disfavored and doubts should be resolved in Plaintiff's favor.

(B) Grant the motion because Plaintiff failed to offer evidence creating a genuine dispute of material fact.

(C) Deny the motion because the credibility of Defendant's witness is a question for the jury.

(D) Grant the motion only if Defendant's evidence establishes Defendant's version of events beyond a reasonable doubt.

Answer: (B). Under Rule 56, once the moving party has demonstrated the absence of a genuine dispute of material fact, the burden shifts to the nonmoving party to produce specific facts showing a genuine issue for trial. An unsubstantiated assertion in a brief does not satisfy that burden. The court views facts in the light most favorable to the nonmovant, but there must be actual evidence (not just argument) to avoid summary judgment.

13. The district court enters final judgment for Plaintiff on August 1. Defendant files a motion for a new trial on August 10. The motion is denied on October 5. When must Defendant file a notice of appeal?

(A) Within 30 days of the August 1 judgment.

(B) Within 30 days of the October 5 denial.

(C) Within 30 days of whichever comes first: the judgment, or the denial of post-trial motions.

(D) Within 14 days of the October 5 denial.

Answer: (B). Under Fed. R. App. P. 4(a)(4), a timely post-trial motion, including a Rule 59 motion for new trial, tolls the time for filing a notice of appeal. The 30-day appellate window runs from the entry of the order disposing of the last such motion, not from the original

judgment. Defendant filed the Rule 59 motion within 28 days of judgment, making it timely, so the appeal clock reset when the motion was denied on October 5.

14. A district court in the Ninth Circuit grants summary judgment to Defendant on two of three claims but certifies under Rule 54(b) that there is no just reason to delay entry of final judgment on those two claims. Plaintiff wants to appeal the summary judgment rulings immediately. May the Court of Appeals hear the appeal?

(A) No, because the litigation is not yet complete in the district court.

(B) Yes, because the district court certified the ruling under Rule 54(b).

(C) No, because Rule 54(b) certification is only available in cases involving multiple parties.

(D) Yes, but only if the appellate court independently finds that immediate review serves judicial efficiency.

Answer: (B). Rule 54(b) allows a district court to certify a partial final judgment for immediate appeal in cases involving multiple claims or multiple parties, when there is no just reason for delay. The Ninth Circuit is bound to hear the appeal once certification issues because it does not independently second-guess the district court's certification decision except in extreme circumstances. Without Rule 54(b) certification, the appellate court would lack jurisdiction because the final judgment rule requires a complete resolution of all claims.

15. In a diversity case governed by Texas law, the federal district court applies its own procedural rules for summary judgment but applies Texas substantive law on the elements of the underlying negligence claim. A Texas statute of limitations of two years bars the claim. The district court applies the federal four-year statute. Was this correct?

(A) Yes, because statutes of limitations are procedural and federal courts apply their own procedural rules.

(B) Yes, because the Federal Rules of Civil Procedure govern all procedures in federal court, including limitations periods.

(C) No, because under the Erie doctrine and the outcome-determinative test, statutes of limitations are treated as substantive for Erie purposes.

(D) No, because federal courts must apply state law on all matters in diversity cases.

Answer: (C). Guaranty Trust Co. v. York established that in a diversity case, statutes of limitations are outcome-determinative and must be borrowed from state law under the Erie doctrine. Applying the federal four-year period instead of Texas's two-year period would substantially affect the outcome in a way that would encourage forum shopping and produce inequitable administration of the law, exactly the concerns Erie was designed to address.

Note on expanding the question bank: The fifteen sample questions above cover the most frequently tested Civil Procedure topics. To build a full 25-question diagnostic, add questions

covering: (1) supplemental jurisdiction under § 1367; (2) the Erie doctrine applied to judge-made federal common law; (3) compulsory counterclaims and the risk of waiver; (4) mandatory versus permissive intervention; and (5) judgment on the pleadings under Rule 12(c). Use the same format — fact pattern, four options, answer with explanatory reasoning. GenAI can draft additional questions efficiently; see Phase 2 for the prompt structure.

Phase 2: AI Analysis and Personalized Study Materials

Step 1: The Analysis Prompt

Students upload their completed answer sheet and the scored diagnostic to the GenAI platform of their choice. They then use the following prompt. They should paste it exactly, then add their completed answer sheet below it.

PROMPT 1: DIAGNOSTIC ANALYSIS

I just completed a 25-question Civil Procedure diagnostic exam. My answers and the correct answers are listed below. Please analyze my performance and do the following:

1. Identify which questions I got wrong and which I got right.
2. Group my errors by topic area (jurisdiction, pleading, discovery, etc.) and calculate my accuracy rate for each topic.
3. Rank my topic areas from weakest to strongest.
4. For each topic where I got more than one question wrong, explain the likely conceptual gap my errors suggest. Do not just say what I got wrong, but what I probably misunderstand.
5. Identify any patterns in my wrong answers. For example, am I consistently choosing answers that reflect a common misconception, or mixing up two related rules?

Present your analysis as: (a) a brief overall summary, (b) a ranked topic list with accuracy rates, and (c) a "gap analysis" for each weak area explaining what I likely misunderstand and what I need to review.

[Student pastes their answer sheet here, formatted as: Q1: my answer / correct answer, Q2: my answer / correct answer, etc.]

Step 2: The Study Material Prompt

After reviewing the analysis output and confirming it looks accurate, the student uses a second prompt to generate personalized study materials. This prompt references the gap analysis from Step 1, so it works best in the same conversation thread.

PROMPT 2: PERSONALIZED STUDY MATERIALS

Based on the gap analysis you just provided, please create a personalized study guide targeting my three weakest areas. For each weak area, provide:

1. A clear, concise explanation of the rule or doctrine, written as if you're explaining it to someone who almost understands it but keeps getting one thing wrong.
2. A concrete example that illustrates the rule in action, using a realistic fact pattern.
3. A "common mistake" section that explains the specific error I appear to be making and why it's wrong.
4. A mnemonic, visual, or memory trick that will help me remember the rule under exam pressure.
5. Two practice questions targeting this specific gap, with complete explanations of the correct answer and why the wrong answers are wrong.

Keep the tone direct and practical. This is for bar prep, not a treatise. I need to understand this well enough to apply it correctly on multiple-choice questions under time pressure. If there are competing rules or circuit splits that regularly appear on the bar, flag them.

A Faculty Note on Prompt Quality

Students will vary in how well they execute these prompts. Some will paste them exactly; others will summarize or shorten them. Spend five minutes in class walking through the prompts before students use them. Show a live example of what a good AI analysis looks like versus what happens when the prompt is vague. That fifteen-minute investment pays off in data quality.

Also: some students will get an AI analysis that misidentifies their gaps, particularly if their answer sheet has transcription errors. Tell them explicitly to review the analysis before generating study materials. Garbage in, personalized garbage out.

Phase 3: The Day 2 Parallel Diagnostic**Generating the Parallel Exam**

The Day 2 diagnostic is not the same exam, but a parallel exam with different questions on the same topics, weighted toward each student's individual weak areas. Because each student has different gaps, their Day 2 exam should look different. This is where the protocol gets genuinely individualized.

There are two ways to handle this. The simpler approach is to have each student generate their own parallel diagnostic using GenAI. The more controlled approach is for the faculty member to create several parallel exam versions and assign each student the version that addresses their particular weaknesses. For most implementations, the first approach is sufficient.

PROMPT 3: PARALLEL DIAGNOSTIC GENERATION

Based on my gap analysis and the study materials you created, please generate a 25-question parallel diagnostic exam for me. The exam should:

1. Include 12–15 questions targeting my three identified weak areas (weighted proportionally to how weak each area was).
2. Include 10–13 questions covering my stronger areas, to confirm I've maintained that knowledge.
3. Use the same format as the original diagnostic: MBE-style multiple choice with four answer options.
4. Include questions of varying difficulty: some straightforward rule-recall, some requiring application to a slightly unusual fact pattern.
5. NOT repeat any question from the original diagnostic verbatim, though the same doctrinal issues may be tested.

After generating the exam, provide a separate answer key with brief explanations. I will take the exam first, then check my answers.

Administration of Day 2 Exam

Administer under the same conditions as the baseline: timed (50 minutes), no materials, no GenAI during the exam itself. Students record their answers before consulting the AI-generated answer key.

Students should then do the following, again in a new conversation thread to avoid context contamination from the previous session:

- Score their own Day 2 exam
- Compare their performance on each topic to their Day 1 baseline
- Note where improvement occurred and where it did not

PROMPT 4: DAY 2 COMPARISON ANALYSIS

I just completed my Day 2 parallel diagnostic. Below are my Day 1 results and my Day 2 results, organized by topic. Please:

1. Calculate the improvement (or regression) in each topic area from Day 1 to Day 2.
2. Identify which gaps the overnight study successfully closed.
3. Identify which gaps remain: where I still performed poorly despite targeted study.

4. For any area where I did not improve, explain what additional review might help and suggest a different approach to studying that material.

[Paste Day 1 and Day 2 results by topic here]

Phase 4: The Day 4–5 Retention Check

Three to four days after the Day 2 exam, students take a third parallel diagnostic. The purpose is to distinguish immediate learning from durable retention. Some students will show strong Day 1-to-Day 2 gains that largely disappear by Day 4. Others will show modest initial gains that hold or even improve. The difference matters because it tells you something about the depth of learning, not just the speed of it.

Generate the Day 4 exam using the same prompt structure as Day 2 (Prompt 3), adapted to focus on the areas that showed the most improvement. You want to test whether that improvement held.

Students should take the Day 4 exam cold: no review of study materials beforehand. The point is to measure what they actually retained, not what they can access with a fresh review.

Data Templates

Individual Student Tracking Sheet

Each student maintains this record across all three assessments. They submit it (along with their AI chat logs) at the conclusion of the protocol.

Topic Area	Day 1 Score	Day 2 Score	Day 4 Score	Net Change
Personal Jurisdiction	___/5	___/5	___/5	
Subject Matter Jurisdiction	___/5	___/5	___/5	
Venue & Transfer	___/3	___/3	___/3	
Pleading (Twombly/Iqbal)	___/4	___/4	___/4	
Joinder & Claims	___/3	___/3	___/3	
Discovery & Work Product	___/3	___/3	___/3	
Summary Judgment	___/2	___/2	___/2	

Appellate Procedure	___/3	___/3	___/3	
Erie / Preclusion	___/2	___/2	___/2	
TOTAL	___/25 (___%)	___/25 (___%)	___/25 (___%)	

Class Aggregate Data Template

The faculty member collects individual results and enters class-level data here. This is where the research value lives. The aggregate tells you where the curriculum is working and where it isn't.

Topic Area	Class Avg Day 1	Class Avg Day 2	Class Avg Day 4	Avg Gain D1→D2	Retention D2→D4
Personal Jurisdiction					
Subject Matter Jurisdiction					
Venue & Transfer					
Pleading (Twombly/Iqbal)					
Joinder & Claims					
Discovery & Work Product					
Summary Judgment					
Appellate Procedure					
Erie / Preclusion					
Overall					

How to calculate retention rate: Divide the average Day 4 score by the average Day 2 score for each topic. A retention rate above 90% means knowledge held; 75-90% indicates some decay but adequate retention; below 75% suggests surface-level learning that did not consolidate. If most students show strong Day 2 gains that fall below 75% retention by Day 4, the intervention is producing cramming, not learning.

Interpreting What You Find

What the Data Can Tell You

The three-timepoint design produces several distinct data patterns, each with a different interpretation.

Pattern	What It Suggests	What to Do
Strong D1→D2 gain, strong retention	The intervention worked; the gap was addressable through targeted review	No change needed. This is the target outcome
Strong D1→D2 gain, poor retention	Surface learning; students crammed but did not consolidate	Revisit spaced repetition; this topic needs multiple exposures across the course
Weak D1→D2 gain on a specific topic	The AI-generated materials did not address the actual gap, or the gap is deeper than targeted review can fix	Pull those student chat logs; review what study materials were generated for this topic
Consistent weakness across the class (Day 1)	The topic is undertaught in the curriculum — a structural issue, not a student preparation issue	Revisit how much class time and assessment you devote to this topic
Wide variance in Day 1 scores	Students entered with very different preparation levels	Consider offering the protocol earlier in the semester, before preparation gaps widen further

What the Data Cannot Tell You

The basic protocol has honest limitations. Students who performed poorly on Day 2 despite studying might have studied less thoroughly than students who improved, and you cannot control for effort. The AI-generated materials differ in quality across students, because students vary in how well they execute the analysis and study material prompts. And the parallel exams, while designed to test the same doctrinal territory, are not psychometrically equivalent in the formal sense.

None of these limitations make the data useless. They make it what it is: meaningful evidence that should inform teaching decisions, not peer-reviewed research that can support publication without further controls. If your results are striking enough that you want to publish them, that's when you go back to the IRB and the comparison group design.

Adding a Comparison Group

The basic protocol is a pre-post design with a single cohort. It can tell you that students improved from Day 1 to Day 2; it cannot tell you whether the AI-generated study materials caused that improvement. Students who know they'll be retested have reason to study harder regardless of what materials they use.

A comparison group addresses this. In the refined design, a subset of students take the same baseline diagnostic, but identify their own weak areas and study from traditional materials: their outlines, commercial bar prep books, or recorded lectures. They take the same Day 2 and Day 4 parallel diagnostics as the AI-intervention group. The only difference is the study material.

This design raises the ethical question the book flags: is it fair to give some students less effective study materials than others? The honest answer is that you do not yet know which materials are more effective, that is what the study is designed to find out. Both groups receive a structured intervention. Neither group is denied help entirely. If the AI materials prove substantially more effective, you can share them with the comparison group afterward.

Two practical notes on running this version:

1. Randomly assign students to groups if you want results that carry any methodological weight. Self-selection destroys comparability: students who opt into the AI group will likely be the more engaged students, and any difference in outcomes will reflect that selection bias rather than the intervention.
2. Track study time for both groups. Ask students to log how long they spent with their study materials. If the AI group spends 90 minutes and the comparison group spends 40 minutes, a performance difference tells you very little about the materials, only about time on task.

Adapting the Protocol for Other Subjects

Civil Procedure is used throughout this appendix because it is one of the more conceptually dense MBE subjects and one where students frequently have uneven preparation. The protocol transfers directly to any bar-tested subject.

A few subject-specific notes:

- **Contracts and Torts:** These subjects have large, well-established question banks. Commercial bar prep providers (Themis, Barbri, Kaplan) can serve as the source for diagnostic questions rather than requiring you to draft your own. Use their questions and have GenAI analyze performance against those questions.
- **Constitutional Law:** The conceptually heavy nature of ConLaw means the "common misconception" analysis in the AI study materials is especially valuable. Students often confuse the tests (strict scrutiny, intermediate scrutiny, rational basis) and their triggers. Ask GenAI to pay particular attention to distinguishing when each standard applies and why.
- **Evidence:** Evidence is rules-heavy, and mnemonics are genuinely useful here. The study material prompt's request for memory devices will produce particularly practical output for Evidence. Students who struggle with the hearsay exceptions benefit from structured rules-based review.
- **Professional Responsibility:** MPRE preparation is an excellent use case for this protocol. The questions are scenario-based and require applying Model Rules to ambiguous facts. The AI's analysis of which rules a student is misapplying is more useful here than raw topic scores.

A note on technology: The prompts in this appendix were drafted for use with any major GenAI platform: Claude, ChatGPT, or Gemini will all produce useful output. Claude tends to produce the most analytically precise gap analysis; ChatGPT produces reliably organized study materials. For the analysis prompts (Prompts 1 and 4), Claude's structured reasoning is generally preferable. For the study material generation (Prompt 2), any major platform will do. Whatever platform you use, run the prompts yourself before asking students to use them. You should know what good output looks like before you can help a student who gets mediocre output.