Question-Driven Learning in the Age of AI

The Art, Science, and Business of Inquiry

Andrew Caplin

Professor of Economics, New York University

Chief Scientific Officer, Genrae

NBER Behavioral Macro

July 2026

We usually begin observing too late.

By the time the outcome appears, much of the inquiry history that produced it has disappeared.

What we observe	What has already disappeared
mortgage, lender, closing	borrower's inquiry history
student output, credential, first job	student's preparation history
factory, production, success or failure	organization's discovery history

Central claim

The observable endpoint misses critical determinants of decision quality.

From the Science of Mistakes to the Science of Inquiry

This talk introduces a forthcoming book and research program in cognitive economics.

It is directly downstream of:

An Introduction to Cognitive Economics: The Science of Mistakes (Caplin, 2025a)

Three earlier lines of work pointed toward the same missing object:

Housing Partnerships	Households preparing for commitment inside an imbalanced professional market (Caplin et al., 1997).
Propensity to plan	Organized preparation as an economically consequential capacity (Ameriks, Caplin, and Leahy, 2003).
Cognitive economics	Hidden mental objects and the records needed to diagnose economic mistakes.

The Thesis of the Book

The missing object is goal-directed inquiry under evolving commitment.

At the ideal limit, the record would preserve:

questions	sources or probes	answers	interpretations
revisions	stopping	commitments	future learning costs

Commitments of attention, time, money, identity, institutional credibility, and productive capacity accumulate while inquiry is still unfolding.

Our sciences and professions have studied endpoints far more successfully than the inquiry histories producing them.

Failure to study

The absence of these histories is not merely a data limitation. It is a fact about what our sciences and institutions have failed to make visible.

Three Record Environments

The three settings make inquiry observable in different ways.

Setting	Role	Record
Housing	build the adviser from assisted inquiry	structured question–answer trace
Education	develop and observe inquiry capability	student preparation trace
Frontier investment	reconstruct staged discovery	production record of probes and commitments

Different maturities

In housing, the institution and the record can be built together. In education and frontier investment, the records are more partial and exploratory.

From Model Objects to Built Records

Question-driven learning supplies a grammar for the missing history.

$$\text{model object} \to \text{ideal inquiry record} \to \text{feasible trace}$$
$$\to \text{measurement} \to \text{improvement.}$$

Modeling	names questions, probes, answers, diagnostic states, costs, interpretations, and stopping
Data engineering	asks what institutional record would make those objects observable (Caplin, 2025b)

Next

We now turn to the formal grammar: decision problem, questions, answers, and diagnostic environment.

Why AI Changes the Moment

AI changes the urgency, observability, and design of inquiry.

- Visible output can separate from owned capability.
- AI can preserve questions, revisions, source use, interpretation, and stopping while inquiry remains live.
- AI itself can be redesigned around sustained inquiry.

Future AI

A future AI should preserve what the user is trying to discover, where the inquiry stands, and why the next question matters.

From AI that answers questions $\Longrightarrow$ **AI that sustains coherent inquiry.**

The Underlying Decision Problem

Begin with a familiar economic problem.

$$\theta \in \Theta, \qquad y \in Y, \qquad u(y, \theta).$$

The agent wants to choose the action that fits the payoff-relevant state, but does not initially know θ.

Decision problem

Action must be chosen under uncertainty about the payoff-relevant state.

The full model and formal analysis are in Caplin (2026a).

Inquiry Before Action

Before choosing y, the agent can try to learn.

$$q \in \mathcal{Q}, \qquad a \in A_q.$$

Inquiry may unfold as

$$q_1 \to a_1 \to q_2 \to a_2 \to \cdots \to y.$$

Value of inquiry

Answers may change the best final action and which later questions are worth asking.

The Familiar Benchmark: An Experiment

In standard information economics, an experiment comes with a signal structure.

An experiment q is associated with a stochastic kernel:

$$\kappa_q(a \mid \theta).$$

Once q is chosen, its informational content is defined.

The familiar sequence is:

$$\text{choose experiment} \rightarrow \text{observe signal} \rightarrow \text{update} \rightarrow \text{act}.$$

Benchmark

The experiment itself carries the mapping from payoff states into signal distributions.

Why a Question Is Not Yet an Experiment

Real questions do not generally carry a fixed informational structure on their own.

Consider:

- "Can I afford this mortgage?"
- "Which course should I take next?"
- "Should we scale the pilot line?"

The same apparent question can be:

informative premature ambiguous redundant misleading

Missing object

What a question can reveal depends on the hidden inquiry environment in which it is asked.

The Diagnostic State

Call the hidden inquiry environment the diagnostic state:

$$D \in \mathcal{D}.$$

A diagnostic state is not merely a description of where inquiry currently stands.
It is the complete contingent answer structure of the inquiry process.

For a given D, the model specifies answer distributions:

- for every payoff-relevant state θ;
- for every feasible question or probe;
- at every contingent branch of the inquiry process.

Scale of the object

A single D contains the full contingent answer structure of the inquiry environment.

A Question Is an Experiment Only Conditional on D

A question does not carry a fixed signal structure on its own.
Conditional on the diagnostic state,

$$P(a \mid \theta, D, q).$$

For fixed D, the mapping from payoff states into answer distributions is defined.
Across diagnostic states, the same question may be

informative weak premature ambiguous misleading

Key distinction

The experiment is not q alone. A question becomes experiment-like only conditional on D.

Why Question Order Matters

Question order matters because answers change beliefs about the inquiry environment.

- D already contains the contingent inquiry structure.
- Observed answers update beliefs about which D is operative.
- Updated beliefs change the value of the remaining questions.

Order

Earlier answers matter because they change what the actor believes can still be learned, and which question is worth asking next.

Two Forms of Uncertainty and the Value of Stopping

The actor is uncertain about what is true and how inquiry works.

$$\mu \in \Delta(\Theta \times \mathcal{D}).$$

This joint belief covers the payoff-relevant state θ and diagnostic state D.
At any point, the actor can stop and choose

$$\bar{V}(\mu) = \max_{y \in Y} \sum_{\theta, D} \mu(\theta, D) u(y, \theta).$$

Stop or continue

Continue only when another inquiry stage is worth its cognitive, physical, financial, or organizational cost.

Two Cost Margins of Inquiry

Inquiry has two distinct cost margins.

Producing an answer	**Diagnostic orientation or refinement**
inspection	determine which source has authority
course or project	identify what the activity should reveal
prototype or pilot line	acquire expertise to interpret the result
production run	use prior evidence to identify the next probe

Distinction

Cheaper answers do not solve the problem of knowing which answer is needed.

Diagnostic Refinement Is the Second Margin

Before paying to generate an answer, the actor may need to refine beliefs about the diagnostic state.

$$\text{costly diagnostic refinement} \longrightarrow \text{better-informed question choice.}$$

Γ represents refinement of beliefs about D before question choice.

It is not the physical answer-producing probe itself.

Two margins

The physical probe produces an answer. Diagnostic refinement helps identify which question or probe is worth using.

Costly Diagnostic Refinement

Let the marginal belief over diagnostic states be

$$\nu(D) = \sum_{\theta} \mu(\theta, D).$$

A Bayes-plausible refinement satisfies

$$\Gamma \in \mathcal{B}(\nu), \qquad \sum_{\hat{\nu}} \Gamma(\hat{\nu})\hat{\nu} = \nu.$$

Following the RI tradition, refinement costs $\mathcal{C}(\nu, \Gamma)$.

Interpretation

The agent pays to sharpen understanding of which diagnostic environment may be operative.

Sims (2003); Matějka and McKay (2015); Caplin and Dean (2015).

Refined Beliefs

Suppose the refinement produces diagnostic belief $\hat{\nu}$.

The resulting joint belief is

$$\gamma_{\mu,\hat{\nu}}(\theta, D) = \mu(\theta \mid D)\hat{\nu}(D).$$

At this stage:

- beliefs about D have been refined;
- beliefs about θ conditional on D are held fixed;
- the refined joint belief $\gamma_{\mu,\hat{\nu}}$ is used to choose the next question.

Separation

Diagnostic refinement occurs before question choice.

Questions, Answers, and Updating

Given refined belief γ, the agent chooses the next question q.

The predictive answer distribution is

$$P_\gamma(a \mid q) = \sum_{\theta, D} \gamma(\theta, D) P(a \mid \theta, D, q).$$

After observing answer a, the agent updates to

$$B(\gamma, q, a).$$

Two distinct learning steps

Refinement sharpens beliefs about the diagnostic environment before the question.
The answer then produces Bayesian updating after the question.

Timing Within a Period

With inquiry opportunities remaining, the within-period sequence is

$$\mu \longrightarrow \Gamma \longrightarrow \hat{\nu} \longrightarrow \gamma \longrightarrow q \longrightarrow a \longrightarrow B(\gamma, q, a).$$

1. Begin with joint belief μ.
2. Decide whether to stop.
3. If continuing, choose costly refinement Γ.
4. Receive refined diagnostic belief $\hat{\nu}$.
5. Form joint belief γ.
6. Choose question q.
7. Observe answer a.
8. Update and enter the next period.

Finite-Horizon Inquiry

Assume t inquiry periods remain.

$$V_t(\mu) = \max\{\text{stop now}, \text{continue inquiry}\}.$$

Continuing means:

- refine diagnostic beliefs;
- choose and ask a question;
- observe an answer and update;
- continue with one fewer period.

Formal result

Under suitable conditions, the finite-horizon problem is well defined and admits an optimal policy.

Full recursion and existence conditions: Caplin (2026a).

Same Current Decision Quality, Different Inquiry Quality

The model now separates two performances that are often observationally confounded.

Two actors may be equally good if forced to act now:

$$\max_{y} \mathbb{E}[u(y, \theta)] \quad \text{is the same for both.}$$

Yet they may differ sharply in how well they continue inquiry:

- which question is worth asking next;
- which source has authority;
- whether diagnostic refinement is worth its cost;
- when the inquiry should stop.

Implication

Current decision quality can look identical even when inquiry quality is very different.

Same Current Recommendation, Different Best Next Question

Information about the inquiry environment can matter even when it does not change the current recommendation.

Diagnostic state	Valid evidence route
D_P	direct elicitation is informative
D_E	external evidence is informative

Learning whether D_P or D_E applies may reveal nothing directly about θ and may leave the current best recommendation unchanged.

But it can completely change the best next question.

Point
Information can matter entirely through the inquiry it organizes.

Two Organizations of Inquiry

Diagnostic structure can organize inquiry in different ways.

Identification	**Sequence preservation**
find the decisive question	preserve the effective path
failure removes a candidate	failure destroys continuation value
continuation survives	continuation collapses

Order value

The same nominal questions can have different value in different orders because their diagnostic effectiveness depends on where they occur.

Economic Data Engineering

Economics often separates theory from the data institutions already provide.
This project takes another route:

$$\text{theory} \rightarrow \text{ideal record} \rightarrow \text{institutional construction}$$
$$\rightarrow \text{new data} \rightarrow \text{revised theory and practice.}$$

Method

This is not theory without data or data without theory. It is theory-led data creation.

Three Research Programs Now Emerging

theory $\longrightarrow$ missing record $\longrightarrow$ constructed research object.

Environment	**Missing object**	**Research construction**
Mortgage choice	borrower inquiry and mistakes	diagnostic map, CCA, governed trace
Working with AI	correct human–AI boundary	guided/discretionary assignment and separate training
Frontier entry	discovery history before verdict	public ledger of probes, answers, commitments

Construction

The program is beginning to construct the experiments, institutions, and records that make inquiry measurable.

Housing Finance in the Seventh Inning

Mortgage inquiry often becomes most urgent after the path has already narrowed.

By the time financing is seriously reconsidered, the borrower may already be attached to:

- a property;
- an agent;
- a lender;
- an imagined future in the home;
- an offer and closing timetable.

Inquiry under evolving commitment

A question that would have opened the decision early can feel like an obstacle once the transaction is moving toward closing.

Exceptionally experienced mortgage professionals describe this as the “seventh-inning” problem.

Information-Rich, Diagnosis-Poor

The modern mortgage market is not short of answers.

Borrowers can consult:

disclosures	calculators	lender materials	agency guides
comparison sites	housing counselors	professionals	AI

But the market does not reliably determine:

- which question matters for this borrower;
- which source has authority over that question;
- which answer is incomplete or interested;
- which uncertainty should be resolved next.

Diagnosis

The problem is not information scarcity. It is the absence of borrower-centered diagnostic organization.

The Interested Advisory Ecosystem

Mortgage beliefs are formed inside a network of partial and interested sources.

agent	lender	broker	builder
affiliated lender	family and peers	internet	AI

These sources differ in:

- product access;
- incentives;
- practical knowledge;
- legal authority;
- organizational objectives;
- ability to see the full borrower problem.

QDL reading

The informational meaning of an answer depends on the source, the question, and the diagnostic environment.

What the Existing Record Does Not Yet Reveal

The existing record makes mortgage outcomes much more visible than the production of mortgage advice.

Important mechanisms remain poorly measured or not yet made visible:

- when serious lender engagement begins;
- how agent and builder referrals shape lender choice;
- the practical product menu considered for one borrower;
- how professionals acquire product knowledge;
- the role of lender account executives;
- workflow burden across products;
- pipeline pressure and satisficing;
- product branches omitted before the borrower ever sees them.

Record gap

The selected mortgage is visible. The inquiry process that produced it is largely missing from the current record.

Exceptional Practice as Scientific Input

The best mortgage professionals have spent decades observing mechanisms that the research record has rarely been designed to capture.

Their accumulated knowledge helps identify:

- when transaction momentum begins to dominate diagnosis;
- which questions distinguish apparently similar borrowers;
- how practical product sets become restricted;
- how workflow and institutional position shape recommendations;
- how borrower answers should be interpreted;
- when inquiry should continue, stop, or escalate.

Epistemic principle

Exceptional practice identifies what must now be modeled, measured, and tested. It does not by itself establish prevalence or causal magnitude.

The Borrower Is Also an Uncertain Source

The adviser's account of the future cannot treat borrower forecasts as settled facts.

Recommendations may depend on uncertain beliefs about

tenure income employment mobility family refinancing

The advisory task is to interpret, verify where possible, preserve uncertainty, and test whether advice survives alternative futures.

QDL reading

The answer itself becomes a diagnostic object.

Product Choice Is Diagnostic Logic

Mortgage choice is not a lookup problem.

The correct product can depend on one fact that changes the ranking.

Comparison	Questions that can reverse the ranking
FHA, conventional, and mortgage insurance	credit profile, down payment, insurance duration, expected tenure, likely refinancing
Down-payment assistance	grant or debt, forgiveness schedule, refinance trigger, sale trigger, expected mobility

Exceptional advice

The valuable capability is not memorizing every product. It is recognizing which question opens the consequential branch.

Closing Is Not the End

The mortgage creates a future decision environment.

Its value depends on:

- future rates and refinancing;
- mobility and expected tenure;
- insurance cancellation or continuation;
- sale and repayment triggers;
- forgiveness schedules;
- contingent obligations;
- the borrower's ability to recognize and act on later opportunities.

A product with the lower payment today may create a more restrictive path tomorrow.

Planning capital

Good advice should preserve the borrower's capacity to understand, revise, and act as future states unfold.

The Missing Institution

No current participant reliably owns the complete borrower inquiry.
The market has disclosures, originators, agents, counselors, comparison tools, underwriting systems, product guides, and AI.
What remains unowned is the function that:

- enters early enough to affect the path;
- identifies and sequences consequential questions;
- evaluates sources and interprets answers;
- compares lifecycle paths;
- knows when to stop or escalate.

Institutional failure

The system has many participants but no reliable owner of borrower-centered inquiry.

The Cooperative Advisory Problem

The advisory problem is already difficult with a cooperative consumer.

The adviser must:

- define executable protocols;
- construct the expert diagnostic map;
- choose literal questions, sources, and procedures;
- interpret a partial inquiry record;
- value the next question by its effect on the eventual protocol;
- decide when to stop;
- preserve an auditable trace.

This is not simply consumer search, expert advice, bargaining aversion, or conversational AI.

Benchmark

Reticence does not create question selection. It changes the technology through which question selection produces evidence.

The Consumer Has Two Economic Roles

In the stronger advisory problem, the consumer is not only the beneficiary of advice.

$$\theta = \text{what is good for the consumer}$$

$$\psi = \text{response-technology type}$$

It indexes observable or estimable response patterns, not diagnosed anxiety, embarrassment, distrust, or bargaining aversion.

The consumer is both:

- the person whose welfare the recommendation should serve;
- a holder, source, and producer of information needed for that recommendation.

Harder problem

Advice must learn about the consumer and about how inquiry with the consumer can proceed.

Responses Can Be Substantive or Procedural

Not every useful response is direct evidence about the welfare state.

Substantive response	evidence about θ
Procedural response	evidence about how inquiry can continue

"I do not know," hesitation, deferral, or refusal may reveal little directly about the best mortgage.

But it can reveal whether the next useful move is to ask again, change source, verify, repair, stop, or escalate.

A procedural response tells the system what may be feasible next; it need not reveal why the consumer responded that way.

Procedural evidence

A response can organize inquiry without directly settling the final recommendation.

Harmless and Harmful Refusal

Nonresponse is not one economic object.

Harmless refusal	the question fails, but later inquiry is unchanged
Harmful refusal	the attempt itself lowers the value of asking again

The stronger case begins when inquiry changes the future response technology.

Dynamic danger

A question can be costly not only because it fails, but because it damages future answerability.

Beliefs and the Continuation State

The advisory state must include more than beliefs about the consumer.

$$\text{state} = (\mu, b)$$

- μ records what the adviser believes about welfare-relevant circumstances and response technology.
- b is a reduced-form continuation state: how prior interaction has changed future answerability or inquiry feasibility.
- b is estimated from traces; it is not directly observed psychological burden.

Two histories can generate the same beliefs and require different actions because they leave different capacities for future inquiry.

Repair

Repair can be valuable even when it provides no direct information, because it restores the future inquiry technology.

The Concerned Consumer Agent

The CCA does not begin by recommending a mortgage.

Question-Driven Advice

↓

Endogenous Answerability

Layer 1	protocols, diagnostic map, questions and sources, partial record, sequencing, stopping
Layer 2	procedural responses, future answerability, repair, protection, delegation, escalation

It chooses among an issue–source–method probe, returning escalation, or a typed stopping protocol.

institution : governs,

controller : applies policy,

language model : parses, explains, renders.

The household retains decision authority; sources retain bounded authority; the CCA owns continuity of the inquiry.

Concern

The Feasible Inquiry Trace

$$\text{raw evidence} \to \text{qualified evidence} \to \text{state transition}$$
$$\to \text{admissible actions} \to \text{selected protocol} \to \text{execution}.$$

The trace preserves:

- why the issue became active and which source had authority;
- what remained unresolved and which actions were excluded;
- why inquiry continued, stopped, protected, abstained, or escalated;
- whether the structured protocol was executed.

Empirical object

A transaction record tells us what closed. An inquiry trace shows how advice was produced and where it failed.

$$\text{better advice} \quad + \quad \text{a new empirical object}.$$

Housing: Building the Missing Record

In housing, I do not have to wait for these data objects to appear.

Through Genrae, we can study assisted buyer inquiry before full commitment and build the CCA from what we learn.

$$\text{assisted inquiry} \longrightarrow \text{question–answer trace} \longrightarrow \text{diagnosis} \longrightarrow \text{better CCA}.$$

The record can reveal:

- which questions buyers ask and fail to ask;
- how answers are interpreted;
- where exceptional advice changes the path.

Construction

The CCA is a question–answer mechanism built from, and improved through, real housing inquiry.

A Template for Inquiry-Preserving AI

A Broader Template

The CCA is a narrow-domain prototype of inquiry-preserving AI: an agent that retains the goal, state, interpretation, and next question of a consequential inquiry.

The exceptional adviser can improve one consequential decision.

The mortgage setting shows how an institution can organize inquiry for a person. The next question is whether part of that capability can be measured and taught in the person.

Education faces the larger task of helping people acquire the capacity to organize consequential inquiry for themselves.

The Exceptional Adviser Cannot Be Everywhere

Housing builds an adviser for the person.

That can improve one consequential inquiry.

But no person can rely on a dedicated exceptional adviser for every:

- educational choice;
- career transition;
- new technology;
- unfamiliar institution;
- uncertain future.

Education's task

Some of the capacity to organize inquiry must be built into the person.

A natural place to begin is not with the whole of inquiry skill, but with one sharply defined higher-order capability created by AI.

AI Skill Includes Choosing What to Delegate

Productive AI use requires choosing the human–AI boundary.

$$A \subseteq \{1, \ldots, K\}, \qquad \Delta_{ik} = v_k^{AI} - v_{ik}^{H}, \qquad A_i^* = \{k : \Delta_{ik} > 0\}.$$

Under a capacity constraint, the person selects the best limited subset.

Recognition

The relevant skill is not merely using AI. It is recognizing where AI has comparative advantage over oneself.

Execution and Recognition Are Different Skills

AI-assisted performance = execution + AI operation + recognition.

Environment	**Participant controls**	**What it measures**
No AI	human execution	direct task capability
Guided AI	assignment supplied	prescribed human–AI workflow
Discretionary AI	delegated subset	recognition of the human–AI boundary

Identification problem

A single AI-assisted score cannot reveal whether a person executed well or delegated well.

Can the Two Skills Be Taught Separately?

execution training recognition training active control.

Training	**Guided AI**	**Discretionary AI**
Execution	improves	improves
Recognition	little/smaller effect	improves substantially

Recognition training elicits own-performance and AI-performance beliefs, constrained subset choice, observed outcomes, and feedback on selection regret.

Treatment signature

A treatment should improve performance where the capability being taught is activated.

From One Measurable Skill to a Science of Inquiry

$$\text{task execution} \rightarrow \text{comparative-advantage recognition}$$
$$\rightarrow \text{question/source selection} \rightarrow \text{verification/revision} \rightarrow \text{planning capital.}$$

The delegation experiment is a tractable first case:

- the higher-order capability is defined;
- environments suppress or activate it;
- training can target it separately;
- better recognition changes audited productivity.

Pedagogic route

Teaching inquiry may begin with capabilities that can be isolated, activated, and experimentally moved.

What We May Be Outsourcing Before We Understand It

AI can preserve visible output while replacing part of the human inquiry capability that produced it.

The risk is not merely that students use AI.

It is that AI may take over the capacity to identify:

- what should be asked;
- what should be delegated;
- what should be checked;
- what should be pursued next.

Warning

We may be outsourcing the skill of asking questions before we understand what that skill is.

A Broader Inquiry Observatory

The delegation experiment isolates one capability under controlled conditions.

The university measurement program asks how inquiry capability appears in a broader preparation environment:

- what students are trying to discover;
- how they use AI;
- which sources they trust;
- where visible output separates from owned capability;
- who, if anyone, owns the next step.

Measurement

Controlled experiments can identify specific capabilities. Broader traces can reveal where those capabilities matter institutionally.

Ongoing work with Ali Bagherzadeh and Luminita Stevens.

From Inquiry Capability to Productive Capability

Education asks how inquiry capability can be built into a person.

Frontier investment asks how productive capability can be built into an organization when the production system is not yet known.

At the frontier, questions are asked through:

partnerships prototypes pilot lines production runs

These commitments do more than add capacity.
They generate evidence and change what can be learned next.

The frontier turn
The inquiry record becomes a production record.

A Sovereign Cannot Buy a Production Function

Frontier entry is inquiry under evolving commitment.
The entrant does not begin with a settled mapping:

$$Y = F(K, L).$$

It must discover:

- which process can work and where bottlenecks lie;
- which capabilities must be built internally;
- which knowledge can be transferred;
- what scale is required to learn;
- what evidence warrants the next commitment.

QDL reading

Investment is not what happens after inquiry. It is one of the principal ways questions are asked.

The Cost of Asking a Question

At the frontier, questions are asked through costly commitments.

Commitment	**Question made answerable**
technology partnership	can access become absorption?
pilot line	can process failures become interpretable?
facility construction	can capacity become a learning environment?
customer qualification	can the process meet external discipline?
next financing stage	does capital preserve discovery or extend motion?

Scale

A frontier stage can be a billion-dollar way of making the next question answerable.

A commitment counts as inquiry only if its trace can alter the next commitment.

Production Records Before Production Functions

The production record is historical.

It preserves the sequence:

$$\text{question} \rightarrow \text{probe} \rightarrow \text{trace} \rightarrow \text{interpretation} \rightarrow \text{next commitment.}$$

The production function is counterfactual.

It asks what output would be produced under alternative combinations of inputs once the relevant capabilities and production relation can be represented.

Ordering principle

The production record must be read before the production function can be interpreted.

History Lowers Future Inquiry Costs

History can lower the cost of learning what must be learned next.

$$\text{doing} \rightarrow \text{lower production cost}, \qquad \text{doing} \rightarrow \text{lower inquiry cost}.$$

Earlier stages can reveal:

- the relevant bottleneck and measurement system;
- which partner can transfer usable knowledge;
- which next question is feasible;
- which failure can now be interpreted.

Planning capital

Production history can create a lower-cost position for discovering what can be produced next.

Good Sequences and Bad Sequences

Equal expenditure can generate different discovery histories.

Better sequence	Worse sequence
fast learning	slow learning
well-ordered questions	badly ordered questions
capability-building partnership	dependence-building partnership
informative failure	motion without learning
preserved option value	destroyed option value

Useful failure sharpens the next question or lowers the cost of answering it.

Central danger

The worst outcome is motion without learning.

The Verdict Comes Too Late

A final verdict can tell us whether production began, whether the firm survived, and whether output met expectations.

But by the time we know whether a frontier project succeeded, much of the inquiry history needed to understand it has disappeared.

The verdict does not reveal:

- what was being learned;
- which questions were asked in which order;
- what capabilities were created;
- why commitment deepened, changed direction, or stopped.

Frontier problem

Success and failure do not explain the inquiry history that produced them.

What the Standard Record Preserves—and Loses

The public record is strongest on visible commitments and outcomes.

It preserves subsidies, investment, factories, patents, employment, output, partnerships, financing rounds, production targets, restructuring, and exit.

It rarely preserves the production history as inquiry:

- what uncertainty each stage was meant to resolve;
- what evidence changed the plan;
- which failure was informative;
- which commitment created capability;
- whether the project learned or merely moved.

Record gap

We observe the project after its inquiry history has been compressed into milestones and verdicts.

Reading the Trace

Each public stage should be read as a partial inquiry record.

$$\text{stage} \rightarrow \text{probe} \rightarrow \text{trace} \rightarrow \text{interpretation} \rightarrow \text{next commitment.}$$

For each stage, ask:

- What question did the commitment embody?
- What trace appeared, and what did it settle?
- What remained unresolved?
- Did the stage lower future learning costs?
- Did it build capability, dependence, or only motion?

Housing builds the trace; education activates the skill experimentally; frontier analysis reconstructs inquiry from public commitments.

build	experiment	reconstruct

Rapidus Before the Verdict

Rapidus is a live public trace record of sovereign re-entry under unresolved production knowledge.

The visible stages include:

- IBM and Albany: access to a research path;
- imec and other partnerships: external knowledge and ecosystem entry;
- Chitose and the pilot line: domestic production learning;
- prototypes and design infrastructure: technical trace formation;
- customer engagement: external discipline;
- successive public financing stages: continued commitment under uncertainty.

Before the verdict

Rapidus is not yet a success case or a failure case. It is an evolving record of staged discovery.

What Remains Unknown

The public record is informative, but radically incomplete.

Ledger field	What must remain distinct
Documented fact	funding, partnership, construction, milestone, output
QDL interpretation	apparent question, inferred probe, capability implication
Plausible alternative	another reading consistent with the same fact
Confidence	strength of factual or interpretive claim
Genuinely unknown	internal hypotheses, rejected paths, thresholds, learning costs

Scientific result

The missing internal inquiry record is itself evidence of a failure to preserve the history needed for understanding.

Genrae as a Recorded Frontier

Genrae is being designed to preserve two inquiry records at once.

mortgage inquiry record $\rightarrow$ identify mistakes and improve the CCA,

company discovery record $\rightarrow$ improve Genrae and the theory of staged discovery.

The first records how households prepare for consequential commitment.

The second records how a new advisory technology and institution are discovered.

The Reflexive Opportunity

The design objective is for QDL to help build the company while the company helps build the science.

Genrae's development will be recorded as staged discovery, not reconstructed only after the verdict.

One Program, Three Construction Strategies

Setting	Scientific problem	Construction strategy
Housing	advice is produced but its inquiry is unowned	build the adviser and the trace
Education	AI performance confounds execution and recognition	design environments and interventions that separate them
Frontier investment	the verdict hides the discovery path	reconstruct staged inquiry from public production records

Art: exceptional judgment reveals what the existing record has lost.

Science: model the capability and create discriminating measurements.

Business and institutions: build durable systems that improve inquiry and learn from its traces.

Common enterprise

The common enterprise is to make practical inquiry visible enough, structured enough, and measurable enough to improve.

From Data Engineering to Engineering

model the inquiry $\rightarrow$ specify the missing record $\rightarrow$ build or reconstruct it
$\rightarrow$ identify mistakes and capabilities $\rightarrow$ redesign theory and practice.

a mortgage adviser that can be tested,
an AI skill that can be separately trained,
a production history that can be read before the verdict.

The records do not yet exist in the form we need.
Build them, read them, and learn from them.

Let the engineering begin.

Backup: TSMC as a Production-Learning System

TSMC is a benchmark for accumulated production learning, not retrospective inevitability.
Its system joined demanding customers, repeated yield learning, suppliers, process histories, external discipline, and cumulative capability.
The foundry model helped turn recurring production problems into interpretable evidence.

QDL reading

TSMC matters because its production record lowered the cost of future learning, not because its path was obvious in advance.

Backup: Northvolt as a Breakdown Record

Northvolt is a breakdown record, not simply a failure story.

Public fact	large facilities, customer commitments, production difficulty, contract loss, financing pressure, and bankruptcy
QDL interpretation	scale, production difficulty, customer pressure, and financing pressure became decisive before the production system stabilized
Open question	useful learning may still have survived organizational failure in people, routines, suppliers, and institutional knowledge

Record geometry

Failure of the organization does not prove that no planning capital or useful production knowledge was created.

References I

Ameriks, John, Andrew Caplin, and John Leahy. 2003. "Wealth Accumulation and the Propensity to Plan." *Quarterly Journal of Economics* 118(3): 1007–1047.

Caplin, Andrew. 2025a. *An Introduction to Cognitive Economics: The Science of Mistakes*. Manuscript.

Caplin, Andrew. 2025b. "Economic Data Engineering." Manuscript.

Caplin, Andrew. 2026a. "The Organization of Inquiry: Question-Driven Learning under Diagnostic Uncertainty." Manuscript.

Caplin, Andrew. 2026b. "Investment as Staged Discovery." Manuscript.

Caplin, Andrew, and Mark Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition." *American Economic Review* 105(7): 2183–2203.

Caplin, Andrew, Sewin Chan, Charles Freeman, and Joseph Tracy. 1997. *Housing Partnerships: A New Approach to a Market at a Crossroads*. Cambridge, MA: MIT Press.

Matějka, Filip, and Alisdair McKay. 2015. "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model." *American Economic Review* 105(1): 272–298.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50(3): 665–690.