Lucidity, Made Measurable: An Outcome-Blind, Criterion-Free Cross-Domain Scale of Judgment

Construct and Validation

One Ruler for Newton, Napoleon, and Socrates

Institute of Lucidity · Working Paper · 2026

Disciplinary placement: Management / Leadership · Upper-echelons theory · Leader hubris · Executive succession & derailment

Disclaimer · Experimental project. Every lucidity score and conclusion in this paper is generated automatically by a large language model —not assessed by human experts, and with no ground-truth answer key. Scores may contain factual errors, are not authoritative judgments of any real individual's character, ability, or worth, may diverge from reality, and are offered for research and exploration only.

Abstract

Leaders' judgment —their lucidity about the situations they act in —is a core variable in upper-echelons theory, yet has resisted direct measurement: within-domain hard metrics (citations, victories, market capitalization) are precise but not portable; cross-domain proxies such as fame measure the echo of fate rather than the mind; and any assessment made *after the outcome is known* is entangled with hindsight. More fundamentally, judgment has *no ground truth*, which voids the conventional "score accuracy against a criterion"route to validation. This paper's contribution is accordingly a ruler, not a prediction. **(1) Construct:** lucidity is defined as Pattern $\otimes$ Mystery, decomposing into six operational axes, discriminable from adjacent constructs such as wisdom and metacognition. **(2) Instrument:** a large language model scores each figure from public biography using only then-available information, under a non-moral firewall and an anti-halo clause that strips identity before scoring, making lucidity *outcome-blind, identity-blind, and cross-domain computable* —so Newton and Napoleon land on one commensurable ruler. **(3) Criterion-free validation**, the methodological core: how does one defend a ruler with no ground truth? We demonstrate a suite that *never scores against an answer key* —coverage, structure (effective dimensionality ≈ 2), reliability (ICC 0.73–0.93 single-rater, up to 0.98 averaged), invariance (named $-$ blind $\Delta \leq 0.15$ per axis), discriminant validity (surface features explain only 9% of the score), and information (the two sub-dimensions each retain 27–37% independent variance) —plus three pre-specified experiments. The corpus comprises 459 cross-domain leaders. For upper-echelons and leader-cognition research, the paper supplies a missing *measurement instrument*, enabling longitudinal, cross-domain, hindsight-controlled studies that biography alone could not support.

Keywords: leader judgment · cognitive accuracy · scale development · criterion-free validation · outcome-blind measurement · upper-echelons theory · LLM-as-judge · discriminant validity

AI-use disclosure. The construct, scoring rubric, protocol, validation design, interpretation, and argument are the authors' work; the authors reviewed and verified the outputs and are responsible for all claims. A large language model acts as the rater, scoring each figure from public biography under that protocol, and assisted in drafting prose, tables, figures, and analysis code. Two limitations follow and are treated as first-class. First, because a single model family currently produces the scores, the validation evidence establishes *internal consistency* rather than convergence with human raters or across model families; closing that gap (cross-family re-scoring and human-rater ICC) is the highest-priority item on the research agenda (§7, R4). Second, the quantitative results inherit whatever biases the underlying model carries.

Highlights

The argument in brief. The first three points are the paper's contributions: construct, instrument, and criterion-free validation; the fourth is corroboration that the ruler behaves sensibly, and is deliberately not a claim.

1. **The construct: judgment can be operationalized.** Lucidity is not one blurry "merit."It decomposes into six scored axes (Pattern: Consequential Foresight, Ego Undistortion, Signal Discrimination; Mystery: Confidence Calibration, Boundedness Awareness, Irreducibility Reverence), giving a cross-domain definition of judgment discriminable from adjacent constructs such as wisdom and metacognition. Effective dimensionality is ≈ 2, and Pattern and Mystery each retain $27-37\%$ independent variance, so the two are real, semi-independent sub-dimensions rather than two names for one.
2. **The instrument: the ruler is outcome-blind, identity-blind, and cross-domain computable.** A language model scores from public biography using only then-available information, under a non-moral firewall and an anti-halo clause that strips identity before scoring. Newton and Napoleon thus land on the same ruler at last. Surface features explain only 9% of the score and named-minus-blind drift is ≤ 0.15 points per axis, so explicit identity and halo cues essentially do not enter, which is the precondition for an outcome-blind ruler (re-identification from conduct facts is a separate, untested channel; agenda R-ID).
3. **Criterion-free validation: with no answer key, a ruler can still be defended.** Judgment has no ground truth, so every "score accuracy against a criterion"route fails. We use instead a suite that *never scores against an answer key*: coverage, structure (effective dimensionality ≈ 2), reliability (ICC 0.73–0.93 single-rater, up to 0.98 averaged), invariance (named $-$ blind $\Delta \leq 0.15$), discriminant validity (surface features only 9%), and information (the sub-dimensions each retain $27-37\%$ independent variance), plus three pre-specified experiments. This paradigm generalizes to the whole LLM-as-judge family and is the paper's most transferable contribution.
4. **(Corroboration, not a claim.)** As a nomological net for construct validity, the structural spine is institution-building, strongest at both personal survival and organizational survival, and the one weight-bearing dissociation is that skill is enterprise-tilted across named, blind, *and* objective coding alike. A further tilt of lucidity toward personal fate appears only under the *outcome-blind rescore* (R2, App. E) and does *not* survive *purely objective outcome proxies* (R3a, App. F, directional only), so it is a lead for future work rather than a result relied on. These relations corroborate that the ruler measures something real; they *do not* amount to a claim that lucidity predicts outcomes (full regressions in §5, Tables 1 and 2, and Apps. D, E, F).

In a sentence. We build and validate a ruler for judgment, taking a thing everyone calls central, yet historically discussable only through the halo of hindsight, and making it *measurable, comparable, and open to falsification even without an answer key*. Its relationship to outcomes is a sanity check, not the thesis.

1 The Problem: Everyone Calls Judgment Decisive, and No One Can Measure It

A leader's success or failure is universally attributed to "judgment,"and yet judgment still has *no clean ruler*. Within-domain hard metrics are not portable, cross-domain fame measures the echo of fate rather than the mind, any assessment made after the outcome is known is entangled with hindsight, and judgment itself has *no ground truth*. Judgment, in short, can be discussed after the fact but not measured before it. The two portraits below are not evidence of anything; they simply put the gap on the table. Talent and judgment are different things, and to speak of judgment one first needs a ruler that measures judgment alone, without looking ahead to how the story ends.

Napoleon. In raw ability almost no one compares. He rebuilt France's administration, law, and finance, and the *Code civil* still runs across much of Europe two centuries later; on this paper's scale his skill is 89 (peak 90). Yet the same man marched on Russia in 1812 against all counsel, lost six hundred thousand men, repeated the error at Waterloo, abdicated twice, and died in exile on St. Helena with a great deal of time to reconsider (personal survival = 28). The atlas annotation puts it exactly: this was "a failure of lucidity, not a failure of ability."His skill was intact throughout, still 72–82 in the Russia and Waterloo phases; what collapsed was lucidity (peak $\approx$ 79, down to 16 in the Russia phase). The institutions he built survived (organizational survival = 70); he did not.

Washington. He was far less gifted: his pooled-entry skill is only 58, and the record describes a man who "lost battle after battle yet kept his army intact, wearing the British down through endurance rather than tactical genius."But he handed back power where he could have kept it, stopping at two terms (lucidity $\approx$ 84). His personal ending is at the ceiling (personal survival = 96), and his enterprise endures too (organizational survival = 96). As the annotation notes, the highest personal-survival reading in the whole field belongs to the man whose statecraft was not top-tier but whose restraint was.

The gap. The two men's talent diverges sharply (89 vs 58), while their lucidity diverges the other way ($\approx$ 64 vs 84); the question this paper asks is whether a ruler can register that second gap *without* reasoning back from how each life ended. Talent, it turns out, guarantees nothing about how a life ends. The difficulty is that, without a ruler that does *not* reason backward from the outcome, "Washington was more lucid than Napoleon"is forever just hindsight: impossible to measure on its own terms, compare across domains, or falsify. This paper therefore does not try to predict who will succeed; it first builds a ruler for judgment and shows that the ruler measures something real. The portraits are motivation, not evidence; their relation to outcomes (skill enterprise-tilted, lucidity fate-tilted) is deferred to §5 as a sanity check on the ruler, not as the headline.

The two portraits are not a special case. Widen the lens to a dozen figures across four domains, and the pattern holds (Figure 1): eminence compresses skill toward the high end (the full corpus averages skill 84.5, SD 12.2, and runs 30–99, versus lucidity's wider spread at mean 62.3, SD 18.2), while lucidity fans out across the whole range. Newton at 43 and Darwin at 84 are both first-rank scientists; Alexander at 28 and Genghis at 73 are both undefeated conquerors. Skill is the threshold for entering the atlas at all; lucidity is what actually varies, and it varies *across* domains on one shared axis, which is the entire

point of a cross-domain ruler.

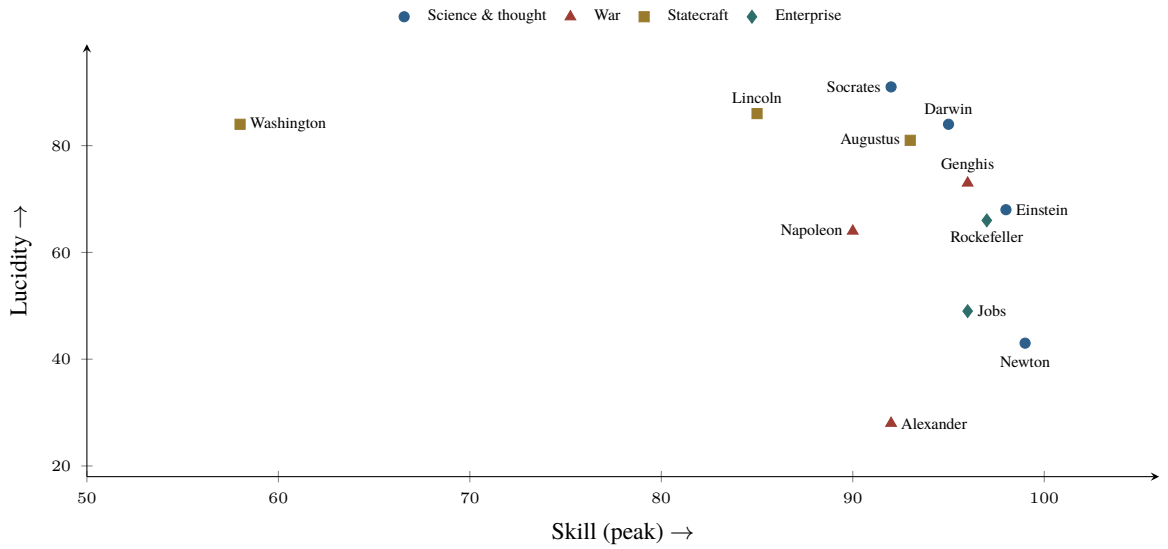

Figure 1 · Twelve exemplars across four domains on one ruler (open-dataset values; App. D). The horizontal axis (skill) is compressed high (corpus mean 84.5, SD 12.2): the eminent are, almost by definition, highly skilled, so the sample restricts skill's range —which attenuates its coefficients (a point returned to in §7) rather than proving it inert. The vertical axis (lucidity) is where they separate, and it separates them the same way across science, war, statecraft, and enterprise. This is what "one ruler for Newton and Napoleon" means operationally.

2 The Construct: What Lucidity Is

Definition. Lucidity, written 明 in the source rubric, is the product of two capacities: Pattern (理), how accurately one models the world, and Mystery (玄), how well one keeps company with the parts of the world that cannot be modeled. Formally, lucidity is the geometric mean of the two, $\sqrt{(\text{Pattern} \times \text{Mystery})}$ rescaled to 0–100, and each branch carries three axes scored 0–5:

- **Pattern:** Consequential Foresight (foreseeing the knock-on consequences of one's own decisions, reasoning only from contemporaneous information) · Ego Undistortion (accuracy of assessing one's own ability and limits, not modesty) · Signal Discrimination (telling real regularity from noise).
- **Mystery:** Confidence Calibration (confidence tracking evidence) · Boundedness Awareness (lucidity about the boundary of a framework's applicability) · Irreducibility Reverence (letting the irreducible core remain suspended rather than forcing a false closure).

The product rule means that if either factor goes to zero, lucidity goes to zero, because "one who believes he sees through everything is precisely the one who is not lucid." Lucidity is thus not a sum of six axes but the *product* of two capacities: seeing accurately, and knowing where one cannot see.

Structure: not one blob but two semi-independent sub-dimensions. Effective dimensionality is ≈ 2. Pattern and Mystery are highly correlated ($r \approx 0.80$) yet each retains 27–37% mutually independent variance, so they are genuinely two sub-dimensions rather than two names for one, a claim supported by the validation evidence in App. A and §4.

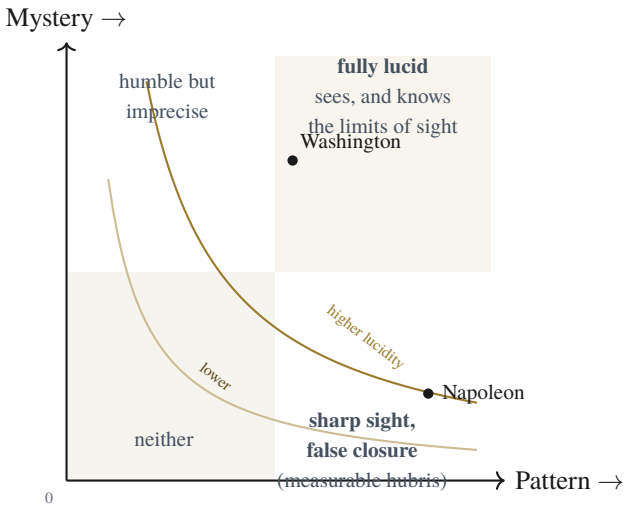

Figure 2 · Lucidity is a product*, not a sum. Because lucidity* $= \sqrt{(Pattern \times Mystery)}$*, the score collapses toward the shaded low corners whenever either factor is small: the high-Pattern / low-Mystery corner is exactly where sharp sight tips into hubris. Iso-lucidity curves are hyperbolas. Portraits use the pooled atlas values (App. D), rescaled to the 0–5 axis range.*

2.1 Relation to Adjacent Constructs (Discriminant Validity)

Lucidity is not wholly new; each of its six axes correlates strongly with an existing construct. The wedge is not "a new dimension"but the *mode of measurement*. Existing constructs are almost all measured by self-report or subordinate rating, and are thereby contaminated by impression management, social desirability, and the rater's knowledge of the outcome. Lucidity, by contrast, is inferred from public biographical behavior and is blind to both identity and outcome, which is the source of its hindsight-resistance, and of its ability to put Newton and Napoleon on the same axis at all.

Axis (branch)	Nearest construct	Existing measurement	Lucidity's distinguishing move
Consequential Foresight secondOrderSight · Pattern	Strategic foresight · second-order thinking (Tetlock 2005; Tetlock & Gardner 2015)	Forecasting-tournament · self-report	Scores foresight of one's own decisions' knock-on effects from only what was knowable at the time, not backward from the known outcome
Ego Undistortion egoUndistortion · Pattern	Leader humility (Owens & Hekman 2012); overconfidence (Malmendier & Tate 2005/2008)	Subordinate/self · option-holding proxy	Measures accuracy of self-assessment of ability and limits, not moral modesty; outcome-blind
Signal Discrimination signalDiscrimination · Pattern	Sensemaking (Weick 1995)	Qualitative · self-report	Operationalizes "telling regularity from noise" as a cross-domain scorable accuracy axis
Confidence Calibration calibratedUncertainty · Mystery	Calibration · overconfidence (Moore & Healy 2008)	Lab calibration tasks	Infers confidence–evidence matching from biographical behavior, not lab probability judgments
Boundedness Awareness boundednessAwareness · Mystery	Intellectual humility; practical wisdom (Grossmann 2010; Sternberg 1998)	Self-report scales	Measures lucidity about a framework's boundary; its negative pole is measurable hubris
Irreducibility Reverence irreducibilityReverence · Mystery	Tolerance of ambiguity · negative capability	Self-report scales	Lets the irreducible core remain suspended rather than forcing a false closure, the axis where existing constructs are thinnest

Table 3 · Per-axis discriminant validity (six lucidity axes × nearest adjacent construct × difference).

Overall, lucidity is closest to practical wisdom (Sternberg 1998; Grossmann 2010), from which it is distinguished on three counts: behavioral inference, outcome-blindness, and cross-domain commensurability. The empirical evidence for discriminant validity (the sub-dimensions each retain 27 – 37% independent variance; surface features explain only 9%) is in §4. One caveat: Irreducibility Reverence has no mature self-report scale to benchmark against, which is both where the measurement novelty is most concentrated and where human-rater convergence (R4) is most needed.

2.2 Positioning Against Upper-Echelons Theory

Upper-echelons theory (UET) is the natural home for this instrument, and also the source of the gap it fills. Hambrick & Mason (1984) argued that organizational outcomes are "reflections of"the cognitions, values, and perceptions of top executives, but conceded in the same paper that these psychological constructs are difficult to observe directly, and therefore proposed *observable demographic characteristics* (age, tenure, functional background, education, socioeconomic roots) as proxies for the underlying cognition. That single methodological move defined the empirical program for four decades (reviewed in Carpenter, Geletkanycz & Sanders 2004), and it drew an equally durable critique: demographics are a *black box*, correlated with outcomes for reasons the theory cannot see inside (Lawrence 1997), a limitation Hambrick (2007) himself acknowledged in calling for measures "closer to the psychological constructs they are meant to represent."The subsequent managerial-cognition turn (Helfat & Peteraf 2015 on managerial cognitive capabilities; Eggers & Kaplan 2013) reopened the box conceptually but still lacks a direct, cross-domain, hindsight-free *measure* of the cognition itself.

This is precisely the slot the ruler is built for. Where UET substitutes demographics for cognition, lucidity scores the cognitive construct directly; where its proxies are population-bound (a CEO tenure

variable does not travel to generals or scientists), lucidity puts Newton and Napoleon on one axis; and where demographic and reputational proxies absorb the outcome they are meant to predict, lucidity is scored outcome- and identity-blind. Table 4 lays the four measurement traditions side by side.

Measurement tradition	What it actually measures	Cross-domain?	Hindsight-immune?	Direct?
UET demographic proxies (Hambrick & Mason 1984; Carpenter et al. 2004)	Age, tenure, functional/educational background, standing in for cognition	No	Partly	No (proxy)
Hubris / overconfidence proxies (Roll 1986; Malmendier & Tate 2005/08; Hayward & Hambrick 1997)	One negative slice (acquisition premia, option-holding, media tone)	No	No	Indirect
Self-report wisdom / humility scales (Sternberg 1998; Grossmann 2010; Owens & Hekman 2012)	Cognition directly, via self/subordinate rating	No	No (impression management + hindsight)	Yes
Lucidity (this paper)	**Judgment directly, six behavioral axes**	**Yes (one ruler)**	**Yes (outcome- & identity-blind)**	**Yes**

Table 4 · Lucidity versus the three dominant traditions for getting at leader cognition. The wedge is not a new construct but a measurement that is simultaneously direct, cross-domain, and hindsight-immune, the combination none of the prior three achieve.

Read together with the adjacent literatures, the same gap recurs: the hubris literature (Roll 1986; Hayward & Hambrick 1997; Hiller & Hambrick 2005) lacks a blind longitudinal operationalization; practical wisdom lacks an automatically computable measure; and succession and derailment (McCall & Lombardo 1983; Shen & Cannella 2002; Finkelstein 2003) have never been distinguished from organizational survival (Stinchcombe 1965; Marquis & Tilcsik 2013) inside one framework. The undercurrent throughout is outcome bias (Fischhoff 1975; Baron & Hershey 1988): assessment usually happens after the outcome is known, so lucidity is hard to measure cleanly, and the outcome-blind design here is meant to answer exactly that. The paper's position is therefore modest and specific: it does not propose a new theory of leadership but supplies the measurement instrument UET has wanted since 1984, so that cognitive variables the literature "wants to control but cannot measure"finally admit measurement, comparison, and falsification.

3 The Instrument: How It Is Computed

Scoring protocol. Scoring is done by a language model from public biography alone. Each life is coded as a timeline of two or three consequential phases; per-phase lucidity, $\sqrt{(\text{Pattern} \times \text{Mystery})}$ rescaled to 0–100, is cost-weighted, and skill takes the peak across phases. Two hard rules hold throughout:

- **The amoral firewall.** Scoring credits accuracy of sight, not character. A pathological breakdown is still scored by the stakes at the time, and is never docked merely because the result was monstrous.

- **The anti-halo clause.** Before finalizing, the rater asks "if I did not know who this was, would I give the same score,"and strips out identity information. This is the step that turns "outcome-blind"from a slogan into a procedure. One honest boundary: the clause removes *explicit* identity cues (name, aliases, titles, and evaluative/legacy sections); it cannot guarantee the model does not *re-identify* a famous figure from distinctive conduct facts ("marched on Moscow in 1812 with six hundred thousand men"). "Identity-blind"should therefore be read as *identity-cue-blind*; whether re-identification survives redaction is an open verification (agenda R-ID), and until it is run the invariance evidence bounds, but does not eliminate, latent-identity leakage.

The properties on which credibility rests. Lucidity has no external criterion, so it is not validated by computing error against an answer key, which would only measure agreement. It is an opinion instrument, not a truth machine, and its credibility comes from the five criterion-free validations and three pre-specified experiments of §4. Only three numbers are needed here: surface features explain just 9% of the score; named-minus-blind drift is ≤ 0.15 points per axis; and inter-rater reliability is high (ICC(2,1) 0.73–0.93 single-rater, up to 0.98 averaged across raters; full ranges in §4), all between model raters, not convergence with humans (see R4). That "lucidity is outcome-blind"is a design claim (the Consequential Foresight axis reasons only from contemporaneous information), and its empirical verification is in §4 and App. E.

Cross-cohort scale (code-audited). The pooled regression reads each cohort's hand-authored raw axis values, with no normalization. Lucidity is consistent across cohorts on a shared fixed rubric, the axis the paper leans on most, and the only fully clean one. Skill and institution-building are directionally consistent but domain-relative in unit. Orientation is the problematic axis. It is pooled across incompatible pole conventions —rulers' "inward self-preservation $\leftrightarrow$ outward projection"versus presidents' and CEOs' "self-dealing $\leftrightarrow$ stewardship."The conventions disagree so sharply that Washington's single act of stepping down scores 43 under one and 95 under the other, so its organizational-survival coefficient cannot be read as a single construct. It is ≈ 0 on personal survival and non-core, with handling deferred to R0. The net judgment is that the skill/organizational-survival relation on which the core corroboration depends is unaffected by orientation.

Sample. The corpus is a hand-coded and model-scored panel of leaders: 459 carry lucidity, of whom 404 carry both outcomes, spanning twenty cohorts. Whether non-organizational leaders such as scientists and artists enter the main analysis depends on journal placement (agenda S2); the conservative option restricts the main analysis to political, military, and business leaders and moves the rest to a generalization appendix. Two fictional cohorts are excluded from the regressions. Variable coding and the full regression tables are in App. D.

3.1 Construction and Transparency: How the Model Is Used

The instrument is a human-authored rubric applied by a large language model; being explicit about which is which is a precondition for trusting the numbers. The construct, the six axes, their anchors, the two hard rules, and the aggregation formula are all authored by hand and fixed in advance. The model's job is narrow: read a biography and return an axis score, following the rubric. Three passes run over the same public source text (Figure 3). The *atlas* pass scores each phase on the twelve rubric components with the figure's identity visible; the *R2 blind* pass first redacts identity and outcome, then an independent judge re-scores the six axes; the *R3a objective* pass ignores the rubric entirely and extracts records-based facts, quote-anchored and blind to the predictor scores.

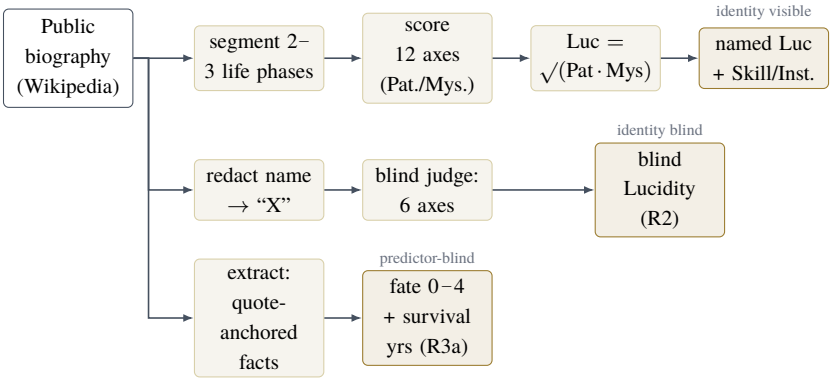

Figure 3 · The construction pipeline. One public source text, three passes: the human-authored rubric (grey) is applied by the model (Claude family) to produce the named atlas scores; the R2 pass re-scores after redaction; the R3a pass extracts objective records blind to the predictors. Gold boxes are the released quantities.

Stage	What the model does	Fixed by hand	Key limitation
Rubric design	nothing (authored by researchers)	construct, 6 axes, anchors, poles, formula	the construct choices are the authors'
Atlas scoring (named)	reads biography, scores 12 axes per phase	rubric + anchor exemplars	identity & outcome visible → halo risk (motivates R2)
Blind rescore (R2)	redacts text, re-scores 6 axes blind	redaction & judging protocol	single rater, one model family, imperfect blinding
Objective extraction (R3a)	pulls death-manner & survival years, quote-anchored	0–4 fate scale (all); survival years (builders only)	manner-of-death is coarse (n=449 / 136)
Aggregation	nothing (deterministic)	geometric mean, stakes-weighting	researcher degrees of freedom in weights

Table 5 · Division of labor between the human-authored rubric and the model. Every value in the dataset is a model reading of a fixed instrument, not a free-form model opinion; the model family throughout is Claude (Anthropic), single-rater, which is exactly the limitation R4 is designed to close.

3.2 The Corpus in Numbers

The corpus at a glance. 459 figures across 20 cohorts, spanning 723 BCE to 1985 CE. Lucidity has mean 62.3 (SD 18.2, range 10–91), with a long lower tail (the mean sits below the mode near 72). The outcome-blind rescore returns a valid score for 397 of the 400 re-scored figures; objective fate is coded for 449 and objective institution-survival for 136 builders. 156 figures are institution-builders. The objective records are 88% Wikipedia-sourced and 85% high-confidence.

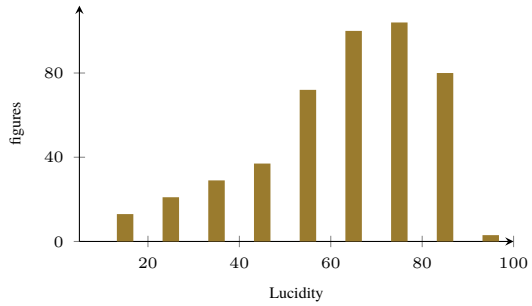

Figure 4 · Lucidity distribution (n=459).

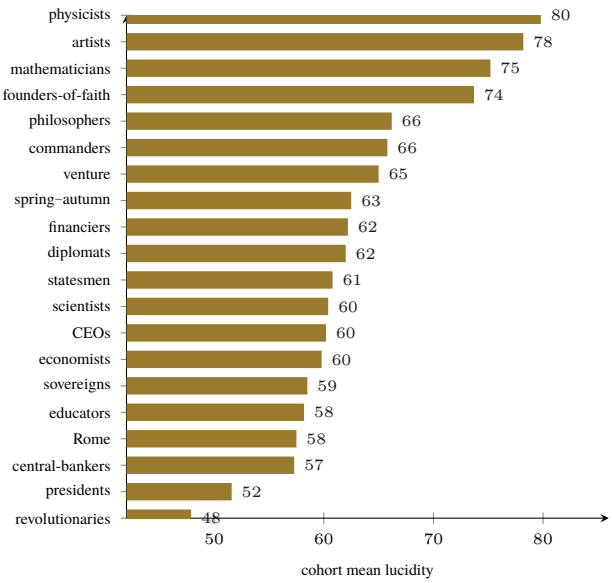

Figure 5 · Mean lucidity by cohort (20 cohorts).

Figures 4–5 · The score is not degenerate: it spreads across the full range with a sensible central mass, and cohort means order interpretably (contemplative and discovery domains sit high; power-seeking domains sit lower), which is descriptive context, not a claim.

4 Criterion-Free Validation: Defending a Ruler With No Answer Key

This is the paper's most transferable contribution. Judgment has no ground truth, so every "score accuracy against a criterion"route fails outright: one cannot validate lucidity against an answer key, because there isn't one. We therefore change the question. Instead of asking "are the scores correct,"we ask "does this set of scores have every internal property a credible measurement instrument should have."

Measurement philosophy: no answer key, so none is invented. Lucidity has no external criterion; the hand-authored reference set is merely "the $R{+}1$-th rater,"never an answer key; and computing MAE or ρ against "gold"is void, since it measures agreement rather than correctness. A distinction must be kept. The *measurement layer* (the criterion-free internal validation of this section) is one thing; *lucidity as an independent variable predicting real outcomes* is another, a criterion-bearing empirical claim to be defended separately via R2, R4, and out-of-sample prediction, and not under the criterion-free exemption. This is exactly why §5 keeps the outcome relations to corroboration and nothing more.

4.1 Five Criterion-Free Validations

The depth battery comprises 20 subjects × 3 raters × {named, blind}, for 120 scorings, spanning the knowledge, outcome, and faith domains (though not yet the regression's workhorse cohorts such as CEOs and central bankers; see R6), alongside 12 synthetic biographies. Each of the five validations answers one property a good ruler should have:

(1) **Coverage.** κ is a first-class output and confirms that the corpus covers the target domains adequately.

(2) **Structure.** Effective dimensionality is ≈ 2 (1.7–2.0 across methods): the reduction from 12 axes to 6, and then to the two-branch structure, is data-supported rather than imposed.

(3) **Reliability.** ICC(2,1) ranges 0.73–0.93 and ICC(2,3) ranges 0.89–0.98, a self-consistency between model raters, not convergence with humans, which is R4's job.

(4) **Invariance.** Named − blind $\Delta \leq 0.15$ per axis (on the 20-subject depth battery, $0-5$ axis scale): explicit identity cues essentially do not move the score, the most direct evidence of outcome- and identity-cue-blindness. (This per-axis invariance is not in tension with the ≈ 11-point mean named−blind gap on the full 397-figure pool at $0-100$ aggregate scale in App. B, B.8: the two are different samples and scales —a handful of high-consensus subjects scored per-axis at 0–5 versus the whole pool, including low-consensus figures, on the rescaled aggregate —and the aggregate arithmetically magnifies small per-axis shifts.)

(5) **Information.** The two sub-dimensions are highly correlated ($r \approx 0.80$) yet each retains $27-37\%$ independent variance, so they are not two names for one.

4.2 Three Pre-specified Adversarial Experiments

Beyond the criterion-free checks, three *pre-specified* experiments actively try to catch the ruler measuring something other than what it claims (*pre-specified* = the design, failure line, and analysis were fixed in writing before the scores were seen; they are not lodged in a public registry, so we do not call them pre-registered):

- **E3, confound regression.** Regressing the score on surface features (era, domain, fame, text length) yields $R^2 = 0.09$, far below the pre-set failure line of 0.50: the score cannot be reconstructed from surface features, and identity and halo did not enter.
- **E5, synthetic profiles.** Twelve biographies pre-designed by Pattern – Mystery quadrant are blind-scored back at $\rho = 0.82$ (Pattern) and 0.87 (Mystery), with a quadrant hit of 10/12: the ruler resolves engineered differences. Its weakness: item-writing and scoring share a model family.
- **E1, blinding ladder.** From L1 to L2 the bias does not grow; only under the L3 extreme rewrite do the two softest axes cross the line; unmeasured bias is ≤ 0.5 points out of 5.

Why this matters beyond the present construct: any LLM-as-judge tool faces the same challenge: "you have no answer key, so why should your scores be believed?"The answer offered here is not "we score accurately"but a validation paradigm that defends a ruler rigorously without an answer key: coverage, structure, reliability, invariance, and information, together with pre-specified adversarial experiments. That paradigm generalizes to the entire LLM-judge family, beyond the present construct.

5 What the Ruler Lets You See: Descriptives and Nomological Corroboration

This section connects the ruler to the world and asks whether its relation to real outcomes matches theory. Everything below is corroboration of construct validity*, evidence that the ruler measures something real, and* not *a claim that lucidity predicts outcomes. Full regressions in Apps. D, E, F.*

Connect lucidity to the two outcomes (personal survival and organizational survival) of nearly four hundred leaders, and a sensibly behaving ruler should show structured, interpretable relations rather than noise. It does, and the pattern is the one theory would predict.

First, a within-cohort look. Before the pooled regressions, hold the domain fixed. Among US presidents, skill and lucidity come apart sharply within the single office (Figure 6): Lyndon Johnson (skill 95) and Andrew Jackson (skill 88) are among the most capable operators ever to hold the office, yet score low on lucidity (33 and 23); Eisenhower sits high on both; and the high-skill / low-lucidity corner is populated while its mirror is nearly empty. This is the two-portrait contrast reproduced inside one cohort,

where domain, era, and role are held roughly constant, so it cannot be an artifact of comparing generals with CEOs.

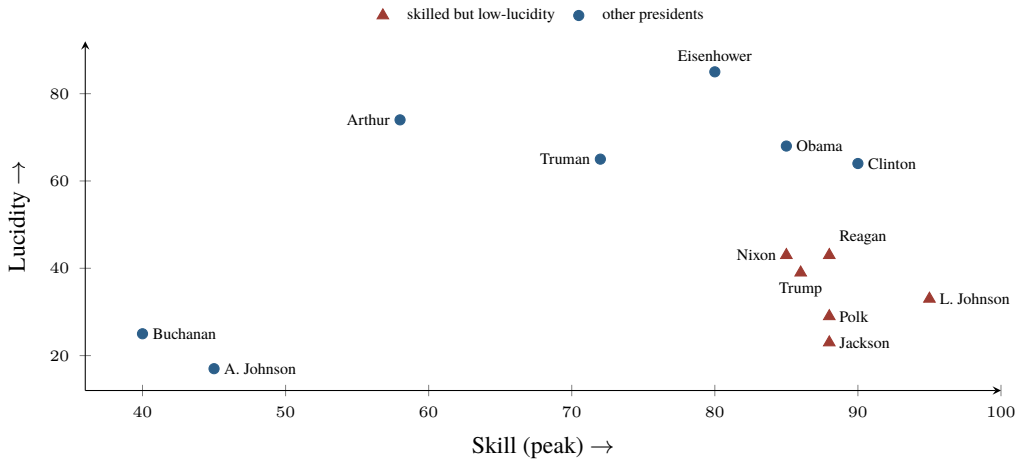

Figure 6 · US presidents (cohort n=40; 13 labeled). Within one office, skill (horizontal) and lucidity (vertical) dissociate: the lower-right region (high skill, low lucidity) is populated, the upper-right (high skill, high lucidity) is sparse. Holding domain fixed rules out cross-domain artifact.

Predictor	→ Personal survival β (t) [ΔR^2]	→ Org. survival β (t) [ΔR^2]
Institution-building	**+0.49** (11.1) [.167]	**+0.51** (16.1) [.183]
Skill	−0.03 (−0.7, ns) [.001]	**+0.22** (6.9) [.034]
Lucidity	+0.27 (5.6) [.043]	+0.21 (5.9) [.025]
Orientation	+0.02 (0.5, ns) [.000]	+0.14 (4.3) [.013]
Difficulty	−0.15 (−3.9) [.021]	−0.09 (−3.5) [.009]
Model R^2	0.457	0.719

Table 1 · Per-outcome OLS ($n = 404$; standardized β; t; unique ΔR^2). Table 1 uses named (atlas) scores; after the outcome-blind rescore, lucidity's coefficients move as in the lower rows of Table 2, while the other four predictors are not rescored (their coefficients shift only slightly).

Capacity	Pers. β	Org. β	Diff	z(subj)	z(coh)	p	Verdict
Skill	−0.03	+0.22	**+0.25**	**5.46**	3.37	**<.001**	**Enterprise-tilted (robust)**
Institution-building	+0.49	+0.51	+0.02	0.37	n/a	.71	Equal both ends (largest)
Lucidity · named	+0.27	+0.21	−0.07	−1.24	n/a	.21	Directional (halo-contaminated)
Lucidity · blind	**+0.34**	**+0.08**	**−0.26**	**−4.66**	**−3.77**	**<.001**	**Fate-tilted (emerges)**

Table 2 · Capability – outcome differential test (stacked SUR, diff = org − personal, computed from unrounded coefficients; lucidity shown named vs outcome-blind rescore). The named → blind shift is theory-consistent, not random: lucidity's inflated organizational-survival link collapses while its personal-fate link strengthens.

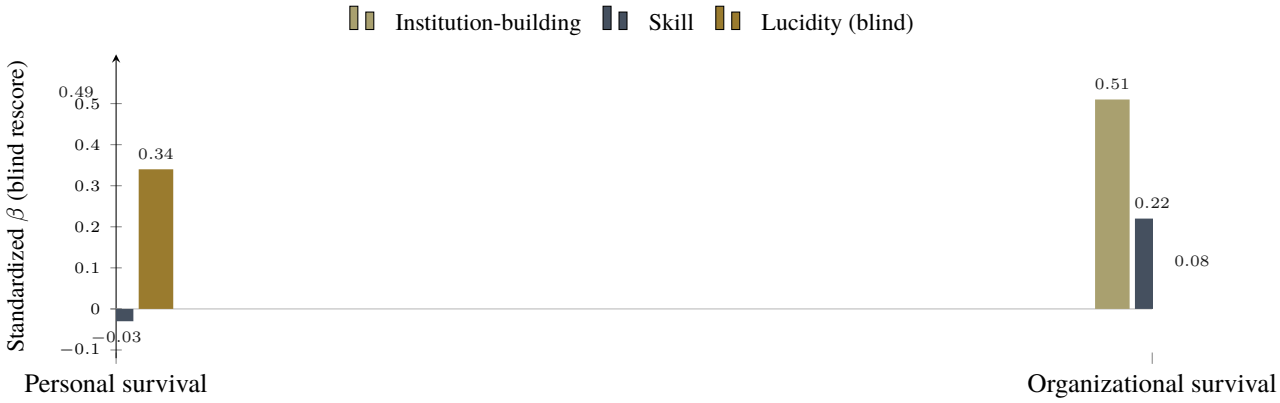

Figure 7 · The capability dissociation, plus one master beam. Skill loads on the enterprise and near-zero on the self; lucidity (blind) loads on the self and near-zero on the enterprise, the two crossing wings. Institution-building carries both ends at ≈ .5, the beam that holds the structure up.

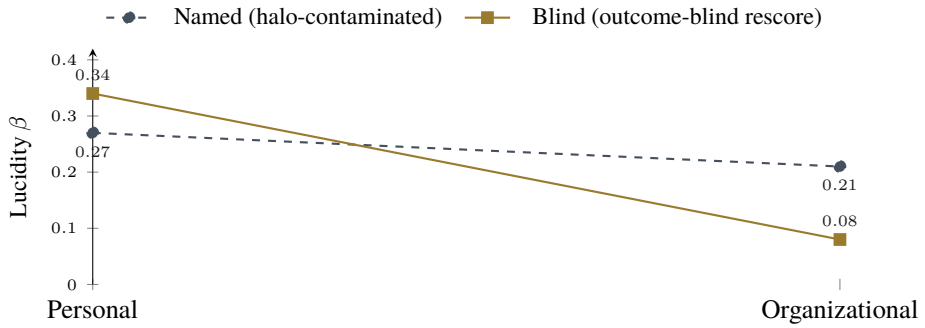

Figure 8 · Removing hindsight opens a scissors. Named scores read nearly flat across the two outcomes; the outcome-blind rescore lifts lucidity's tie to personal fate and collapses its tie to the enterprise. Had blinding merely added noise, both lines would fall toward zero together; instead they separate, the signature of removing a specific halo rather than adding error.

Corroboration one (robust across named and blind): skill is enterprise-tilted only. Skill's coefficient is $\beta = +0.22$ (t = 6.9) on organizational survival and $\beta = -0.03$ (n.s.) on personal survival; the difference, $\Delta = 0.25$ (subject-clustered z = 5.46, p < 0.001), is consistent across four SE specifications and, on the subject-clustered test, strengthens after the rescore (z = 5.5 → 6.0). Skill builds the organization but not the self.

Corroboration two (revealed by the blind rescore): lucidity is fate-tilted, within a coding boundary. With named scores lucidity is significant at both ends and the fate-tilt is not significant ($\Delta = -0.07$, p = 0.21), the halo-contaminated, conservative reading; and across 16 specifications the named fate-tilt sits at a median of ≈ 0 and keeps its (negative) sign in only about half of them (App. D), so *before* blinding it is not a stable effect at all. After the outcome-blind rescore (identity and outcome cues deleted, the whole pool independently re-scored), lucidity's personal β rises to $+0.34$ (t = 8.3), its organizational β falls to $+0.08$ (t = 2.4), and the fate-tilt becomes $\Delta = -0.26$ (subject-clustered z = -4.7, cohort-clustered z = -3.8, p < .001); drop any single cohort and the subject-clustered z stays between -4.0 and -5.0. This is not an artifact: had the rescore merely added measurement error, it would have pulled every coefficient uniformly toward zero, whereas here the opposite happened: lucidity's link to organizational survival collapsed and its link to personal survival strengthened, the signature of removing a specific confound (the "famous, large legacy, therefore surely lucid"halo). But the coding-dependence must be stated flatly: when outcomes are replaced with purely objective facts (a manner-of-death ordinal

plus institution-survival years, 459 figures blind-extracted), R3a finds that the fate-tilt does not replicate: lucidity on objective fate is $\beta = +0.08$ (t = 1.50; t = 0.99 dropping the still-living), and the tilt difference is z = −1.37, p = .17, directional only. Corroboration two is therefore confined to named and blind (expert-style) coding and does not extrapolate to the coarse objective proxy (method and full table in App. F; mechanism in §6).

The overall coordinate: institution-building dominates both ends. Its coefficients are ≈ .5 at each end, the difference is not significant (z = 0.37), and its dominance shares are 56% and 48%, first at both ends. Skill and lucidity are the two wings of a dissociation —robust for skill, coding-dependent for lucidity (§7) —while institution-building carries both, the master beam spanning the structure. R3a shows the master beam is not common-source (institution-building scores t = 3.7 on objective survival and t = 8.5 on objective fate); of the two wings, only the skill-tilt carries through to the objective operationalization, while the fate-tilt does not (see §7 and App. F).

Collinearity is not the culprit. All VIF fall in [1.03, 1.71], so skill's near-zero coefficient on personal survival and blind-lucidity's near-zero coefficient on legacy are real signals, not artifacts.

Taken together, these relations show that the ruler, once connected to the world, behaves sensibly, with an interpretable pattern that survives halo-removal. One distinction should be kept sharp, because the two wings are not equally secure: the skill→enterprise tilt is the *robust* anchor —identical in sign and significance across named, blind, *and* objective coding —whereas the lucidity→fate tilt is the *fragile* one, absent from the named specification curve, emerging only under the named→blind rescore, and converging to a null under objective proxies (R3a). The fate-tilt is therefore a directional lead for future work, not a finding the paper rests on; the weight-bearing empirical result is the skill dissociation. That is worth having, but even it is corroboration of construct validity, not a claim that lucidity predicts outcomes. The named-to-blind shift conflates halo-removal with a rater-and-method change, and a clean causal reading must wait for R4 (cross-model-family re-scoring plus human-rater ICC). The paper's claim stays at the measurement layer.

6 Discussion: What a Measurement Instrument Opens Up

The main contribution: turning "unmeasurable" into "measurable." Leader cognition (judgment, lucidity, hubris) is invoked constantly in the upper-echelons and hubris literatures, and has always lacked a direct, cross-domain, non-hindsight measure. This paper supplies one: a scale of judgment that is blind to outcome and identity, computable across domains, and defensible without a criterion. Its significance is not that one more predictor has been found, but that it opens a body of research that biography alone could never support.

Enabled questions. With computable lucidity, questions that could only be discussed qualitatively can now be tested longitudinally, cross-domain, and blind. Do boards betting on talent versus on lucidity get systematically different succession outcomes? Can leader hubris (the negative pole of Ego Undistortion and Boundedness Awareness) be read *before* a crisis rather than diagnosed after? Is lucidity trainable, and along what trajectory? Each of these shares one prerequisite: that judgment be measurable at all.

Methodological spillover: removing hindsight sharpens the picture. A finding here holds for the whole tradition of assessing leaders. Measuring judgment is trapped by outcome bias, because assessment usually happens after the outcome is known. The outcome-blind rescore does not weaken the signal but *sharpens* it: lucidity's link to organizational survival collapses while its link to personal

survival strengthens, the pattern expected if a halo, not the construct, drove the former. That yields a reusable test: if removing hindsight weakens the signal, the signal was hindsight; if it strengthens, the instrument is measuring the mind. The logic does not depend on this particular construct and transfers to any leader-cognition measure that must be cut loose from hindsight.

A reminder on dependent variables. The ruler also exposes a construction problem easy to miss in upper-echelons work: folding "did the person end well" and "did the organization survive" into a single "performance" erases, in one stroke, the distinct payoff profiles of talent (enterprise-tilted) and lucidity (fate-tilted). Napoleon is the limiting case: top talent, an enduring legal code, and death in island exile. But this is corroboration, not a claim; its strong causal form rests on R4.

7 Limitations and Robustness

1. **Residual confounding in the blind rescore (the key current constraint).** R2 has landed (App. E): deleting identity and outcome cues and re-scoring the whole pool made the fate-tilt significant. But the named-to-blind shift conflates two things: halo-removal and a rater-and-method change (single-rater model subagents, one model family, imperfectly blinded). The fate-tilt is thus currently a de-biased association; establishing it as hindsight-immune causation requires agenda R4, together with R3b.
2. **Common-source variance (R3a partly resolves it, and also bounds the fate-tilt).** Institution-building, skill, and outcomes were hand-coded, and so are not the same source as the model-scored lucidity. R3a (App. F) re-tests with objective outcome proxies: institution-building still holds on objective survival ($t = 3.7$) and objective fate ($t = 8.5$), so its dominance is not pure common-source and the master beam survives objectification. But the same test bounds the fate-tilt: on objective outcomes lucidity is insignificant on fate ($t = 1.5$; $t = 1.0$ dropping the still-living) and the tilt difference is $z = -1.37$, $p = .17$, directional only. So "lucidity saves the person" reads strictly as a named/blind-coding directional tilt, not a claim that carries through to objective facts. Two residuals remain: the objective manner-of-death proxy is coarse and partly mechanical for builders, and objective survival is builders-only ($n \approx 136$, low power); independent blind coding of the *predictors* is still outstanding, at agenda R3b.
3. **A single rater model-family, and no human convergence.** The remedy is agenda R4 (cross-model-family re-scoring plus human-rater ICC), the step that upgrades §4's self-consistency between model raters into convergence with humans.
4. **Validity extrapolation.** The 20-subject battery's validity must be extrapolated to the 404-subject full pool, at agenda R6.
5. **Selection bias as a collider.** Fame is produced partly by ability and partly by dramatic outcome, so sampling on fame can manufacture an association; this must be treated as an internal-validity threat, with the 459-to-404 attrition characterized.
6. **Orientation scale inconsistency.** See §3 and agenda R0 (unify the poles or drop the axis); it has no bearing on the core corroboration.
7. **Researcher degrees of freedom in aggregation.** Lucidity takes the weighted geometric mean and skill takes the peak; alternative aggregation rules should be tested for robustness (agenda R7).
8. **Source-level hindsight (a floor that scoring instructions cannot lift).** The scoring instruction confines the Consequential Foresight axis to contemporaneous information, but the *source text* is a

biography written by editors who knew the outcome, so which conduct facts are recorded, and how they are framed, is already hindsight-selected. Redaction removes evaluative adjectives and legacy sections but not this selection. Outcome-blindness is thus a property of the *scoring* layer, not of the source; the L3 extreme-rewrite arm of E1 is the only probe that reaches the source layer, and it is best read as an upper bound on residual source-hindsight, not as its elimination.

9. **Range restriction on skill.** Entry to the atlas requires eminence, so skill is compressed high (mean 84.5, SD 12.2, versus a presumably wider general-population range). Restriction of range attenuates skill's coefficients, so its near-zero loading on personal survival is a *conservative* reading, not evidence that skill is inert; a range-restriction-corrected estimate (agenda R7) would if anything widen, not close, the skill dissociation.

Conclusion

Judgment is universally treated as decisive to a leader's fate, and yet has long been discussable only through the halo of hindsight, because there was no ruler that refused to reason backward from the outcome. This paper builds and validates such a ruler: lucidity, decomposed as Pattern $\otimes$ Mystery into six operational axes, outcome-blind, identity-blind, and computable across domains. It then shows how to defend the ruler when there is *no ground truth*: through coverage, structure, reliability, invariance, and information, plus three pre-specified adversarial experiments. The ruler takes a thing everyone calls central, yet historically unmeasurable, and makes it computable, comparable, and falsifiable.

Connected to nearly four hundred leaders, the ruler behaves sensibly: the structural spine is institution-building, and a tilt of lucidity toward personal fate appears once hindsight is removed. But this is a sanity check, not a claim, and it carries one boundary worth stating plainly: replace the outcomes with purely objective facts and the fate-tilt yields to institution-building. The claim throughout is a measurement one: a ruler for judgment, not a model that predicts fate. Its value is that a whole class of "we wanted to measure it but couldn't"questions in upper-echelons and leader-cognition research now has an instrument to reach for.

Robustness Checks and Research Agenda

No.	Analysis	Status	Result / deliverable
R0	Refresh snapshot to n = 404; migrate methodology prose to the Pattern ⊗ Mystery formulation	done	Snapshot `--write` landed; prose migrated (orientation unification deferred)
R1	Stacked SUR + differential Wald; cohort FE + cluster-robust SE + bootstrap	done	Table 2: named skill-tilt z = 5.46; lucidity fate-tilt z = −1.24, p = .21
R2	Outcome-blind rescore (delete identity/outcome cues, re-score the full pool; 397 valid)	done	Table 2 and App. E: fate-tilt becomes z = −3.8, p < .001; dissociation emerges
R5	VIF / correlation matrix + dominance analysis + specification curve	done	VIF ≤ 1.71; dominance in App. D
R3a	Objective outcome proxies (manner-of-death ordinal + survival years, 459)	done	Institution-building t = 3.7 on objective survival (master beam not common-source); fate-tilt does not replicate (objective Δ z = −1.37, p = .17)
R3b	Independent blind coding of the *predictors* (skill/institution)	planned	Removes residual hand-coding common source
R4	Cross-model-family re-scoring + human-rater ICC	planned	Locks the causal reading (currently a de-biased association)
R6	Stratified re-run of validity, 20 → 404	planned	Validity extrapolation interval
R-ID	Re-identification probe: can a blind agent name the redacted subject, and at what confidence?	planned	Converts "identity-blind"from a design claim into a tested one; bounds latent-identity leakage
R7	Aggregation robustness (arithmetic vs geometric mean, peak vs mean skill) + range-restriction-corrected skill	planned	Confirms the skill dissociation is not an aggregation or range artifact

A The Full Measurement System

A.1 The six scored axes (canonical names as displayed in the LucidiAtlas; machine keys in italics). Pattern (理): Consequential Foresight (*secondOrderSight* —foreseeing the knock-on consequences of one's own decisions, reasoning only from contemporaneous information) · Ego Undistortion (*egoUndistortion* —accuracy of assessing one's own ability and limits, not modesty) · Signal Discrimination (*signalDiscrimination* —telling real regularity from noise). Mystery (玄): Confidence Calibration (*calibratedUncertainty* —confidence tracking evidence) · Boundedness Awareness (*boundednessAwareness* —lucidity about a framework's boundary) · Irreducibility Reverence (*irreducibilityReverence* —letting the irreducible core remain suspended rather than forcing a false closure). Six scaffold lenses (Doctrinal Non-Capture / Dissent Permeability / Affect Undistortion / Other-Mind Accuracy / Frontier Openness / Contingency Acknowledgment) only aid reasoning and are not scored; the reduction from 12 to 6 axes is justified by effective dimensionality ≈ 2.

A.2 The two hard rules. The amoral firewall (credit only cognitive accuracy; a pathological breakdown is still scored by the stakes at the time) · the anti-halo clause (a uniform exchange test strips out identity information).

A.3 Four auxiliary outputs. Coverage κ · record type · evidentialBasis · cost-weighting.

A.4 Measurement philosophy. Lucidity has no external criterion; the hand-authored reference set is only "the $R+1$-th rater"; computing MAE or ρ against "gold"is void. The distinction to keep: the measurement layer is criterion-free internal validation; lucidity as an independent variable predicting

real outcomes is a criterion-bearing empirical claim, defended separately via R2/R4 and out-of-sample prediction, not under the criterion-free exemption.

A.5 Worked exemplars from the top cohorts. A profile reads more plainly than a definition. Table 6 lists the six axis scores (0–5) and the aggregate lucidity (0–100) for eight figures drawn from the five highest-scoring cohorts; every value is an open-dataset score (`axes_long.csv`). Two regularities stand out. First, at near-equal skill lucidity still varies widely: Feynman and Heisenberg are both first-rank physicists (skill 97 and 95), yet score lucidity 88 and 57. Second, *which* axis gives way is diagnostic of how the lucidity was lost. Figure 9 sets two physicists of near-equal skill side by side: Galileo's Signal Discrimination is near the ceiling (he read the science correctly), yet his Consequential Foresight and Ego Undistortion sit low (the confrontation he neither foresaw nor stepped back from), whereas Feynman is high on all six. Lucidity is not a scalar a figure "has"; it is the shape of a profile, and the aggregate only summarizes it.

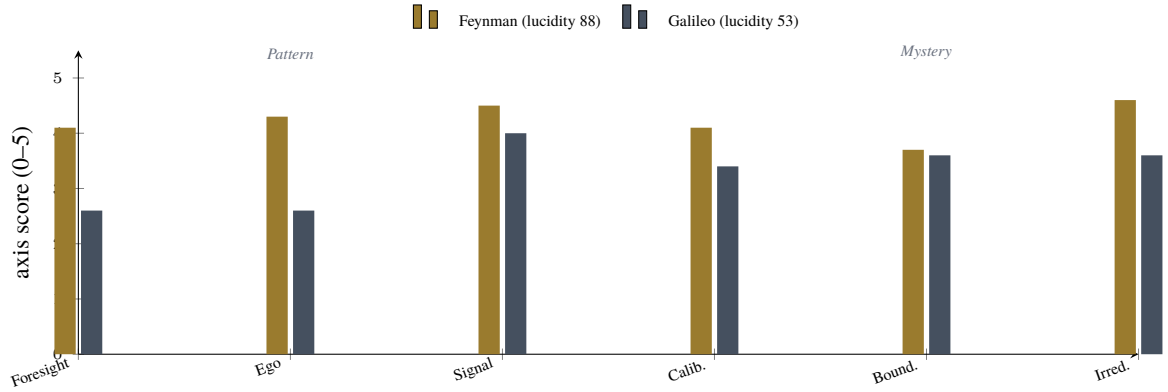

Figure 9 · Two physicists of near-equal skill, six axes each. Bar colour distinguishes the two figures (gold = Feynman, slate = Galileo); the labels above mark the two families (Pattern = first three axes, Mystery = last three). Galileo's shortfall is localised to foresight and self-appraisal, not to reading the pattern —information the lucidity scalar alone cannot carry.

				Pattern			Mystery		
Figure	Cohort	Skill	Luc.	Fore.	Ego	Sig.	Cal.	Bnd.	Irr.
Richard Feynman	physicists	97	88	4.1	4.3	4.5	4.1	3.7	4.6
Werner Heisenberg	physicists	95	57	2.8	3.3	3.2	3.3	2.9	3.7
Henri Poincaré	mathematicians	96	88	4.0	4.0	4.6	5.0	4.6	5.0
Carl F. Gauss	mathematicians	99	77	3.3	3.3	5.0	4.3	4.5	3.7
Paul Cézanne	artists	93	84	3.4	3.7	3.4	3.7	4.0	4.6
Richard Wagner	artists	97	55	2.4	2.0	2.7	2.0	2.6	1.7
Buddha	founders-of-faith	98	91	4.6	4.4	3.3	3.7	3.8	4.0
Socrates	philosophers	92	91	3.6	3.4	3.6	4.6	4.0	4.6

Table 6 · Worked exemplars from the five highest-scoring cohorts (mean lucidity: physicists 79.8, artists 78.2, mathematicians 75.2, founders-of-faith 73.7, philosophers 66.2). Axis scores 0–5 (`axes_long.csv`); lucidity is the aggregate 0–100. Keys —Fore.=Consequential Foresight, Ego=Ego Undistortion, Sig.=Signal Discrimination (Pattern); Cal.=Confidence Calibration, Bnd.=Boundedness Awareness, Irr.=Irreducibility Reverence (Mystery). The Mystery-heavy profiles of Socrates and Poincaré (calibration and reverence for the irreducible near ceiling) contrast with Wagner's uniform collapse at first-rank skill. The aggregate lucidity is the cost-weighted mean of

per-phase $\sqrt{(\text{Pattern} \times \text{Mystery})}$, so it need not equal $\sqrt{}$ of the tabulated cross-phase axis means, which are shown only to convey profile shape.

B The Atlas in Cross-Section: Additional Views

These four views render the open dataset the way the interactive atlas does, so a reader can see the distributions the regressions summarize. Colour is used the same way throughout: the gold iso-lucidity bands in Figure 10 are lines of constant lucidity; the red→amber→teal scale in Figures 11–12 is the atlas's *ending* language (red = collapse, teal = a durable end).

B.1 The two coordinates of lucidity: Pattern × Mystery. Lucidity is a pair, not a scalar. Figure 10 plots the two factors it is built from —Pattern (how accurately the world is read) against Mystery (how accurately the limits of the knowable are held) —for the 317 figures scored on the full construct. The faint gold contours are *iso-lucidity bands*: because lucidity $= \sqrt{(\text{Pattern} \cdot \text{Mystery})}$, a line of constant lucidity is the hyperbola Pattern·Mystery $= c^2$, and the shading deepens toward the top-right where both factors are high. The dashed diagonal is the balance line Pattern = Mystery; the region below-right of it (Pattern outruns Mystery) is the brilliant-but-closed reader —sharp sight, thin reverence.

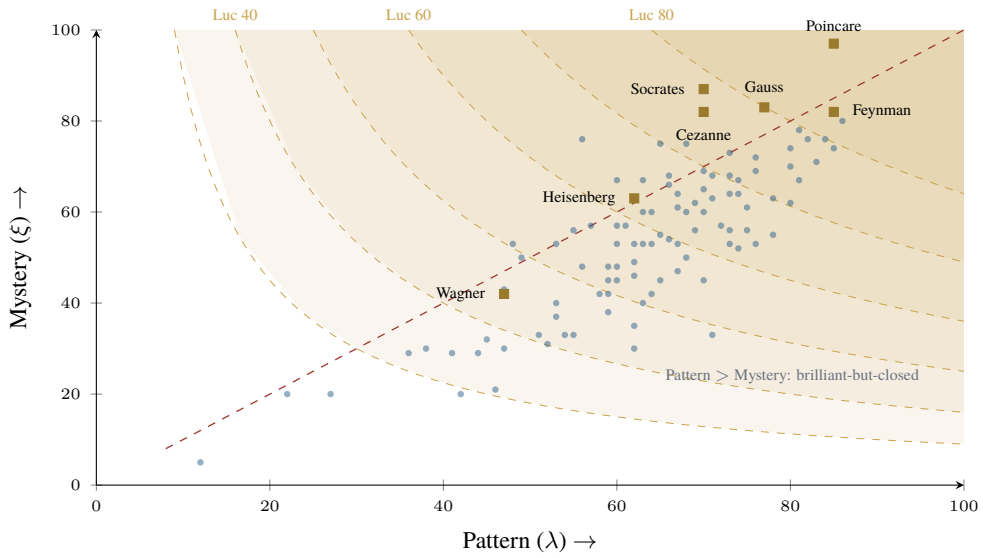

Figure 10 · Pattern × Mystery for the 317 fully-scored figures (`axes_long.csv`, means of each triad rescaled to 0–100). Gold contours are iso-lucidity bands (Pattern·Mystery $= c^2$); shading deepens toward high lucidity. The red dashed diagonal is Pattern = Mystery. Socrates and Poincaré sit high on Mystery (above the line); Wagner collapses on both.

B.2 The two capability axes: Skill × Institution-building, coloured by ending. Figure 11 places the two hand-coded capability variables against each other and colours every point by its *ending* (the atlas's red→teal scale). The upper-right —high skill *and* high institution-building —is where the durable (teal) endings concentrate; raw skill without an institution behind it (lower-right) collapses to red as often as not. This is the descriptive shadow of the regression result that institution-building, not skill, is the master beam.

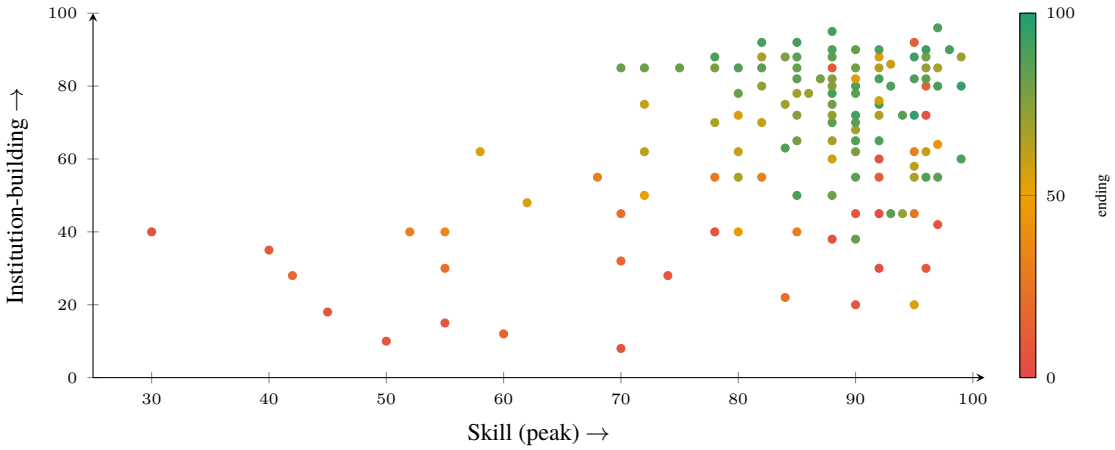

Figure 11 · Skill × Institution-building for 135 real leaders (`figures.csv`); point colour = ending on the atlas's red (collapse) → amber → teal (durable) scale. Durable endings cluster in the high-skill/high-institution corner; high skill with a weak institution (lower-right) is where the reds gather.

B.3 Lucidity × ending, with the outcome bands drawn in. Figure 12 is the single relation §5 leans on for nomological corroboration: lucidity against personal-and-organizational ending. The horizontal bands are the atlas's outcome zones (collapse / mixed / durable). The relation is real but loose ($r \approx 0.50$): high lucidity rarely ends in the red band, but the spread is wide —the ruler measures judgment, not fate, exactly as the measurement-first framing insists.

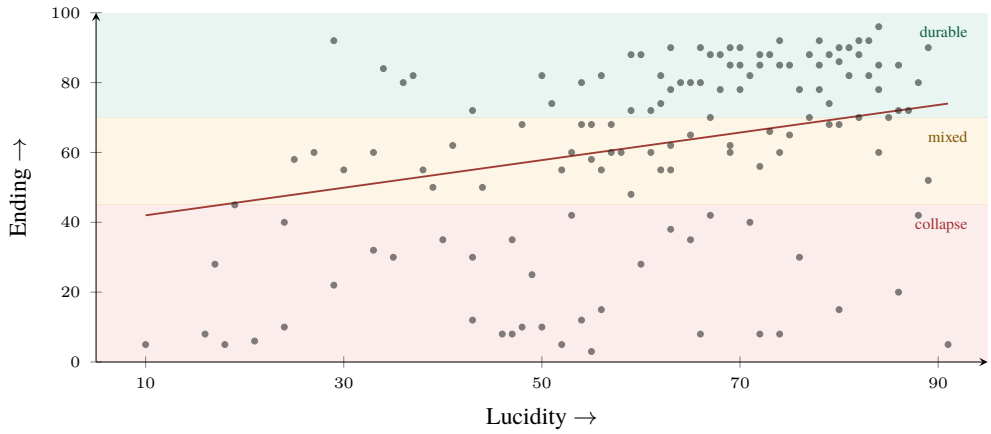

Figure 12 · Lucidity × ending for 404 figures (`figures.csv`); horizontal bands are the atlas outcome zones. The red guide line is the least-squares fit ($r \approx 0.50$): a real positive relation with wide scatter, the visual case for "a ruler of judgment, not a predictor of fate."

B.4 Mixed cohorts on one axis: Skill × Lucidity. Figure 13 overlays four very different cohorts on the shared skill–lucidity plane. They stack almost vertically at the high-skill edge, yet occupy different lucidity bands: physicists sit high (mean lucidity 80), commanders in the middle (66), revolutionaries low (48), with financiers spread across. Skill barely separates them; lucidity does —the cross-domain commensurability claim, shown rather than asserted.

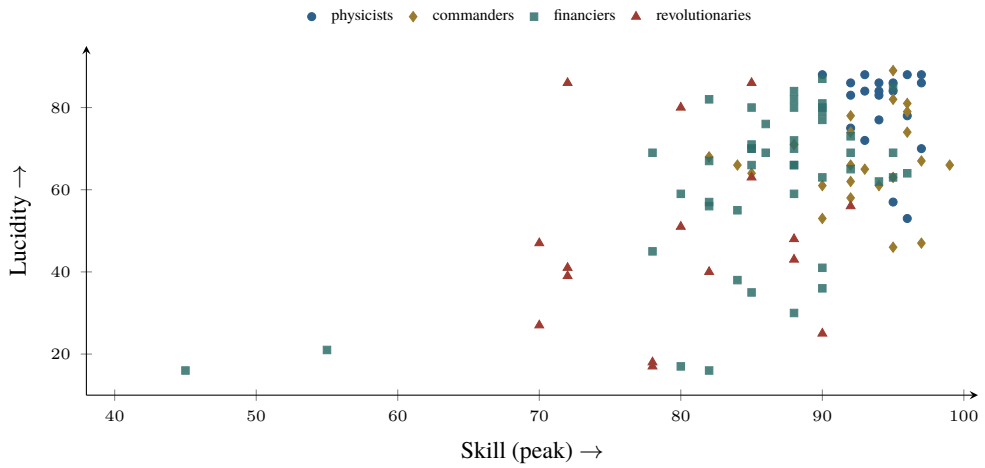

Figure 13 · Four cohorts on the skill–lucidity plane (`figures.csv`): physicists (blue, mean lucidity 80), commanders (gold, 66), financiers (teal, 62, wide spread), revolutionaries (brick, 48). At the high-skill edge the cohorts separate almost entirely along lucidity, not skill.

B.5 The lucid elite of sixteen fields: one summit, scattered fates. Figure 14 takes the most lucid figure from each of sixteen cohorts — seventeen in all, since the arts, uniquely, contribute two ceiling-level figures (Beethoven and Bach) — and places them on the same skill–lucidity plane, encoding two further variables at once: marker *shape* is the broad domain family (thought, arts, power & state, enterprise & finance), and marker *colour* is the figure's *ending* on the atlas's red→amber→teal scale. Two things read at a glance. First, the field-toppers pack into a narrow ceiling —skill 72–99, lucidity 81–91 —regardless of domain, which is the cross-domain commensurability claim in its strongest visual form: a founder of a religion, a mathematician, a central banker, and a liberator occupy one small corner of the plane. Second, and against any temptation to read lucidity as a fate-predictor, their *endings* fan across the entire range within that same corner: Darwin and Buddha end in the deep teal (95, 90), while Socrates —tied for the highest lucidity in the whole atlas —ends at 5, the hemlock. Shape shows judgment travels across domains; colour shows it does not buy a soft landing.

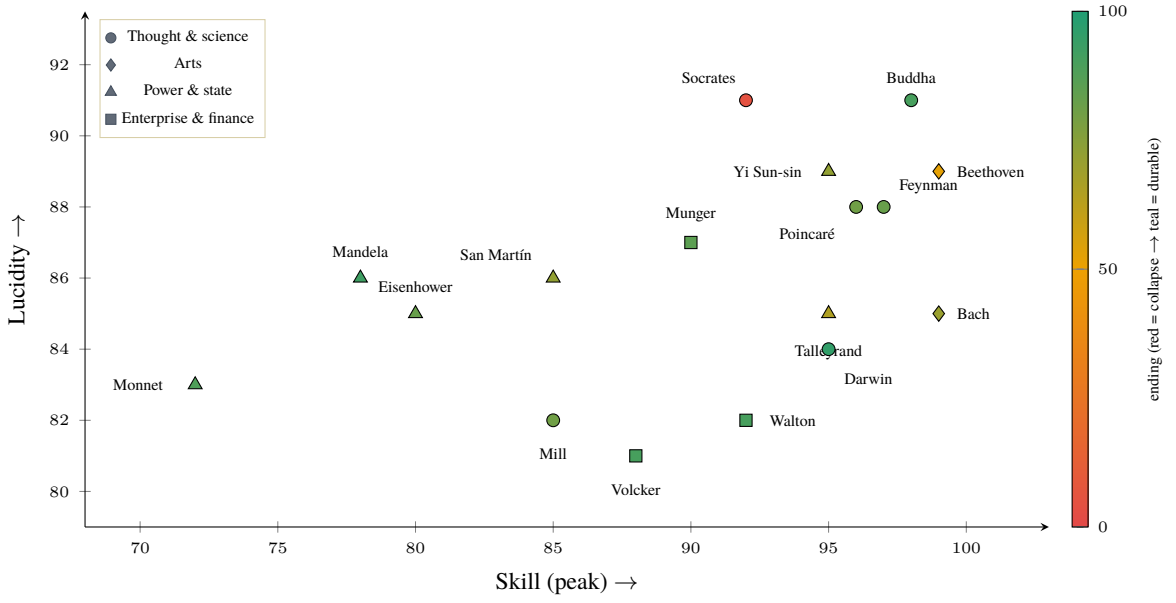

Figure 14 · The most lucid figure from each of sixteen cohorts — seventeen figures, the arts contributing two (`figures.csv`): physicists (Feynman), mathematicians (Poincaré), founders-of-faith (Buddha), philosophers

(Socrates), scientists (Darwin), economists (Mill), artists (Beethoven, Bach), commanders (Yi Sun-sin), revolutionaries (San Martín), sovereigns (Mandela), presidents (Eisenhower), statesmen (Monnet), diplomats (Talleyrand), financiers (Munger), CEOs (Walton), central-bankers (Volcker). Marker shape encodes domain family; marker colour encodes ending (red = collapse → teal = durable). The field-toppers occupy one narrow high-skill/high-lucidity corner (commensurability across domains), yet their endings span the full range within it (lucidity is a ruler of judgment, not a predictor of fate).

B.6 The two capacities that carry the structure: Institution-building × Lucidity. Figure 15 plots the master beam (institution-building) against the fate-tilted wing (lucidity) for all 404 leaders, coloured by *ending*. The two are positively but only loosely related ($r \approx 0.46$, about a fifth of the variance shared): high institution-building neither requires nor implies high lucidity. Both off-diagonal quadrants are populated — 79 leaders build lasting institutions without commensurate lucidity (lower-right, e.g. Wagner), and 70 are highly lucid yet leave little institution behind (upper-left, e.g. Socrates and Feynman). This is the discriminant-validity case for treating the two as separate capacities, and the visual root of the dissociation-and-master-beam structure of §5: precisely because they vary semi-independently, they can load differently on the two outcomes.

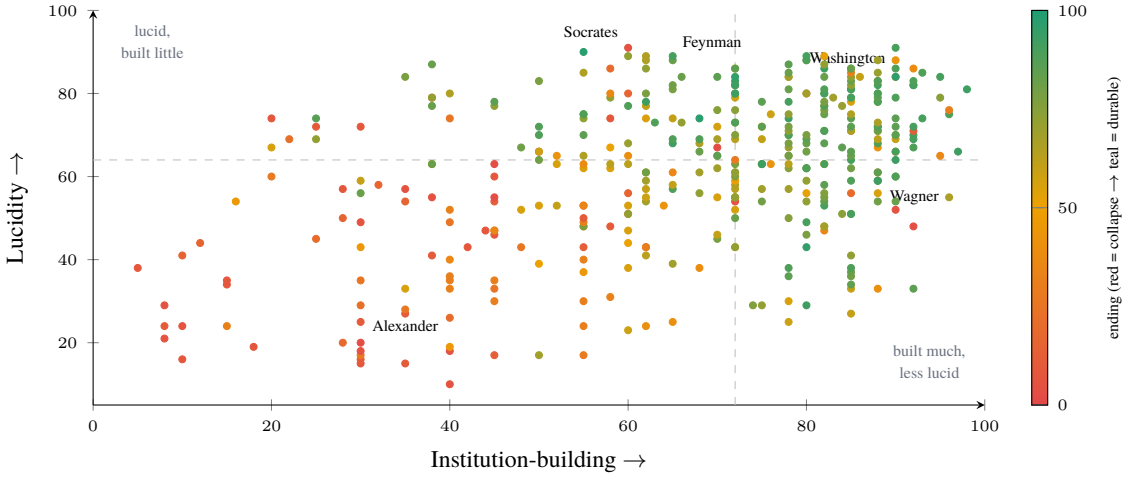

Figure 15 · Institution-building × Lucidity for 404 leaders (`figures.csv`); point colour = ending (red = collapse → teal = durable). The two capacities correlate only moderately ($r = 0.46$); dashed lines mark the medians (institution 72, lucidity 64). Both off-diagonal quadrants fill in: institution-builders low on lucidity (lower-right, e.g. Wagner) and highly lucid figures who built little (upper-left, e.g. Socrates — the atlas's joint-highest lucidity, yet an ending, in red, at the hemlock). That institution-building and lucidity are distinct capacities is why they can carry different outcomes in §5.

B.7 The twenty cohorts at a glance. Table 7 collapses the whole atlas to one row per cohort, sorted by mean lucidity. Two columns are heat-shaded on the atlas's own scales — *Luc* on the white→gold lucidity ramp, *End* on the red→amber→teal ending ramp — so the cross-cohort structure reads at a glance. Lucidity ranges from physicists (80) down to revolutionaries (48); skill is high and nearly flat across almost all cohorts, which is precisely why skill cannot be doing the discriminating work.

Cohort	n	Sk	In	Or	Luc	Df	End	Leg
Physicists	20	94	72	81	80	62	71	95
Artists	13	97	79	75	78	60	56	97
Mathematicians	21	95	76	74	75	63	49	96
Founders of faith	10	94	85	81	74	67	65	94
Philosophers	35	89	76	71	66	63	62	90
Commanders	27	91	62	65	66	74	59	64
Venture	6	85	80	64	65	65	81	84
Pre-Qin thinkers	41	84	45	62	62	72	47	60
Financiers	48	86	65	54	62	59	65	62
Diplomats	9	82	69	65	62	80	65	67
Statesmen	16	84	77	75	61	72	77	78
Scientists	14	88	66	65	60	60	71	81
CEOs	32	87	72	56	60	62	67	69
Economists	15	86	71	76	60	54	78	77
Sovereigns	59	83	74	65	59	73	63	70
Educators	9	75	76	76	58	59	76	78
Rome	19	78	47	62	58	68	41	53
Central bankers	9	74	71	67	57	73	56	58
Presidents	40	68	58	60	52	66	51	56
Revolutionaries	16	80	58	61	48	79	40	56
All figures	**459**	**85**	**67**	**66**	62	**67**	60	**72**

Table 7 · Per-cohort means (`figures.csv`, n=459; Or = orientation, Df = difficulty, End = ending, Leg = legacy). Luc and End cells are shaded on the atlas's lucidity and ending scales. Cohorts sorted by mean lucidity; the final row pools all figures.

B.8 The identity signal, measured then removed: named vs blind lucidity. Figure 16 plots each figure's *named* lucidity (scored with the name visible) against its *outcome-blind* lucidity (scored from a redacted biography) for the 397 fully re-scored figures. Two facts sit together. The two agree on average — both means are 61.1, so naming adds *no* aggregate inflation — yet they disagree per figure by about 11 points on average ($r = 0.68$), and the fit line (slope 0.49, flatter than the dashed 45° identity) shows the blind scores *shrink toward the mean*: naming stretches the tails. This is the halo made visible and quantified — the reason §5 and Table 2 run the fate analysis on the blind scores rather than the named ones.

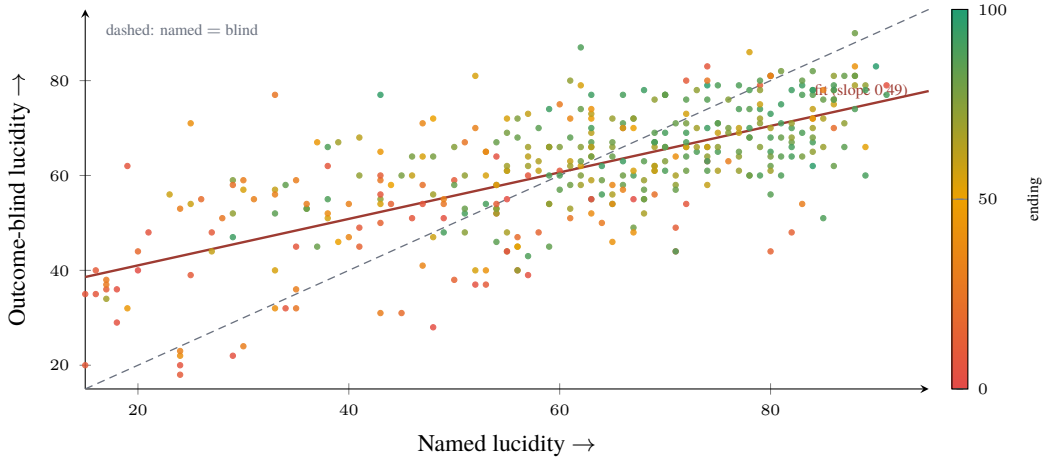

Figure 16 · Named vs outcome-blind lucidity for 397 figures (`figures.csv`); point colour = ending. Means coincide (61.1 vs 61.1); per-figure $r = 0.68$, mean $|\Delta| \approx 11$. The fit (brick) is flatter than the identity (dashed), so blind scoring regresses the extremes toward the middle — the halo the blind rescore removes.

B.9 Reputation is not a soft landing: legacy vs ending. Figure 17 plots posthumous *legacy* (how the world came to rate the work) against personal-and-organizational *ending*. They are positively related ($r \approx 0.58$) but far from identical: the vertical scatter at every legacy level is wide, and the upper-left — towering legacy, hard ending — is well populated. Being remembered well and ending well are correlated, separable things, which is why the atlas scores them as two variables rather than one.

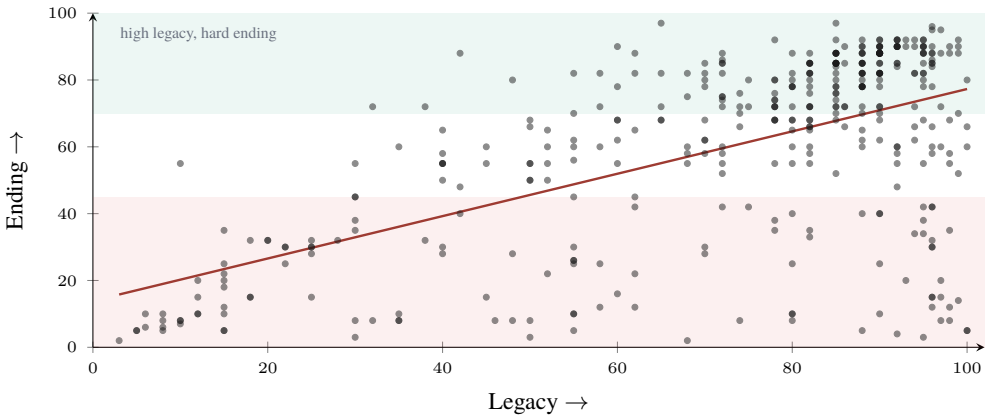

Figure 17 · Legacy × ending for 404 figures (`figures.csv`); the brick line is the least-squares fit ($r \approx 0.58$, ending $= 13.9 + 0.63$ legacy). The wide vertical spread — especially the well-filled high-legacy/low-ending region — is why reputation and personal fate are kept as distinct outcomes.

B.10 The shape of lucidity: a six-axis profile. Figure 18 draws the mean of each of the six axes on a radar — the three *Pattern* (理) axes on the upper-right, the three *Mystery* (玄) axes on the lower-left. The pooled profile is not round: Pattern (mean 3.2–3.3 of 5) consistently outreaches Mystery (2.5–3.0), and the single lowest axis across the whole corpus is Irreducibility Reverence (2.54) — these figures, on average, read the world better than they hold its limits. Two exemplars make the asymmetry concrete: Socrates (teal) balloons on the Mystery side; Napoleon (brick, dashed) is the mirror image — strong on Pattern, thin on Mystery, the brilliant-but-closed reader.

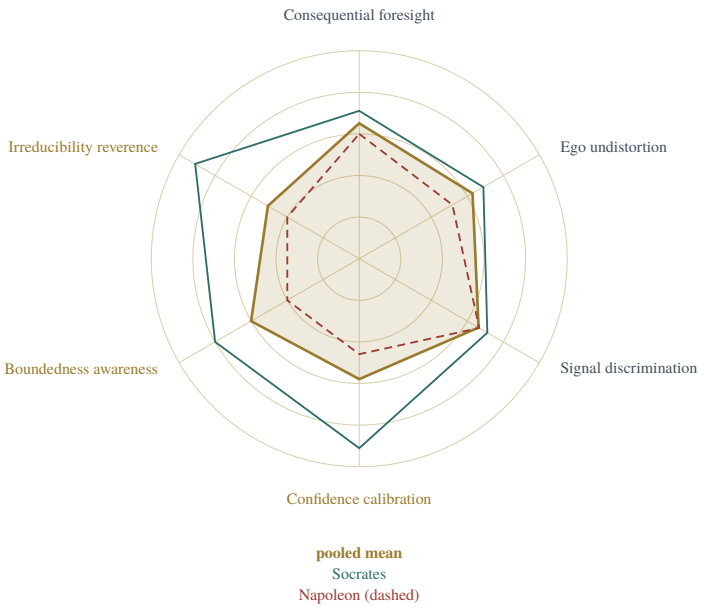

Figure 18 · Six-axis radar (`axes_long.csv`, axis means on 0–5). Pattern (理, slate labels): 3.25 / 3.14 / 3.31. Mystery (玄, gold labels): 2.89 / 3.00 / 2.54. The pooled profile leans to Pattern; Socrates leans to Mystery, Napoleon the reverse.

C The Atlas Index: All 459 Figures at a Glance

Table 8 is the complete atlas in one compact index — every one of the 459 figures with its full metric line, grouped by cohort (cohorts ordered by mean lucidity, figures within each cohort by lucidity). Two columns carry the atlas's colour language: **Luc** is shaded on the white→gold lucidity ramp (deeper gold = more lucid), and **End** on the red→amber→teal ending ramp (red = collapse, teal = a durable end). Reading down the two heat columns together is the paper's thesis in one view: the gold of lucidity and the teal of a good ending are correlated but visibly independent — Socrates and the other deep-gold names are scattered across every ending colour, red included. Column keys: **Coh** cohort, **C** century (negative = BCE), **Sk** skill, **In** institution-building, **Or** orientation, **Luc** lucidity, **Df** difficulty, **End** ending, **Leg** legacy.

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Physicists n=20 Luc 80 End 71									
Enrico Fermi	Phys	20	96	88	72	88	64	68	95
Lise Meitner	Phys	19	90	62	89	88	74	66	94
Richard Feynman	Phys	20	97	65	85	88	60	82	95
Subrahmanyan Chandrasekhar	Phys	20	93	80	86	88	68	88	95
James Clerk Maxwell	Phys	19	97	72	85	86	52	58	98
Ludwig Boltzmann	Phys	19	95	58	85	86	60	20	97
Michael Faraday	Phys	18	94	72	89	86	53	85	96
Niels Bohr	Phys	19	92	90	88	86	66	88	90
Chien-Shiung Wu	Phys	20	93	66	83	84	64	82	94
Lev Landau	Phys	20	95	90	81	84	76	42	96
Max Born	Phys	19	94	82	86	84	61	82	94
Hendrik Lorentz	Phys	19	92	72	83	83	58	92	94
Wolfgang Pauli	Phys	19	94	50	86	83	59	78	95
Ernest Rutherford	Phys	19	96	88	82	78	54	88	96
Erwin Schrödinger	Phys	19	94	45	80	77	66	72	94
Max Planck	Phys	19	92	76	80	75	71	58	96
J. J. Thomson	Phys	19	93	92	83	72	54	90	94
Paul Dirac	Phys	20	97	55	79	70	43	85	96
Werner Heisenberg	Phys	20	95	72	52	57	66	62	90
Galileo Galilei	Phys	16	96	55	70	53	67	34	95
Artists n=13 Luc 78 End 56									
Ludwig van Beethoven	Art	18	99	82	84	89	65	52	99
Johann Sebastian Bach	Art	17	99	88	83	85	54	70	99
Paul Cézanne	Art	19	93	86	86	84	63	60	97
Rembrandt van Rijn	Art	17	96	88	82	83	68	30	96
Wolfgang Amadeus Mozart	Art	18	99	72	78	82	56	14	99
Igor Stravinsky	Art	19	95	78	80	81	55	82	95
Caravaggio	Art	16	96	80	77	80	69	15	96
Francisco Goya	Art	18	96	62	72	77	54	60	96
Leonardo da Vinci	Art	15	99	60	77	77	52	88	99
Pablo Picasso	Art	19	97	96	60	75	58	88	98
Michelangelo	Art	15	99	78	75	74	61	92	99
Vincent van Gogh	Art	19	97	58	79	74	71	8	98
Richard Wagner	Art	19	97	96	41	55	54	66	96
Mathematicians n=21 Luc 75 End 49									
Emmy Noether	Math	19	96	90	89	88	62	42	96
Henri Poincaré	Math	19	96	80	83	88	70	80	95

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
John von Neumann	Math	20	97	92	71	86	66	38	95
Andrey Kolmogorov	Math	20	98	93	83	85	54	88	97
Bernhard Riemann	Math	19	97	85	82	85	66	32	96
Archimedes	Math	−3	97	60	73	80	53	15	97
Leonhard Euler	Math	18	97	88	86	80	56	86	96
Pierre de Fermat	Math	17	90	62	68	80	55	78	96
Joseph-Louis Lagrange	Math	18	95	85	78	79	57	90	96
Niels Henrik Abel	Math	19	95	38	80	79	62	12	96
Carl Friedrich Gauss	Math	18	99	82	51	77	43	88	96
David Hilbert	Math	19	95	96	81	76	76	30	96
Srinivasa Ramanujan	Math	19	95	40	79	74	84	25	90
Évariste Galois	Math	19	96	30	70	72	66	8	97
Euclid	Math	−4	88	90	78	70	42	68	98
Joseph Fourier	Math	18	90	72	65	69	64	62	94
Pierre-Simon Laplace	Math	18	96	82	56	68	59	85	95
Gottfried Wilhelm Leibniz	Math	17	96	88	65	67	55	42	95
Alexander Grothendieck	Math	20	99	95	80	65	72	35	98
Georg Cantor	Math	19	93	85	66	56	81	20	93
Kurt Gödel	Math	20	96	60	69	56	70	15	96
Founders of faith n=10 Luc 74 End 65									
Buddha (Siddhartha Gautama)	Faith	−6	98	90	94	91	55	90	98
Augustine of Hippo	Faith	4	96	85	85	84	70	72	97
Nagarjuna	Faith	2	96	78	83	84	59	60	92
Huineng	Faith	7	95	82	87	81	62	88	95
Xuanzang	Faith	7	95	82	90	80	70	90	86
Ashoka the Great	Faith	−4	82	88	68	73	60	66	86
Paul the Apostle	Faith	1	95	92	83	71	79	12	98
Constantine the Great	Faith	3	90	90	53	60	70	84	92
Martin Luther	Faith	15	92	88	69	59	68	84	96
Jesus of Nazareth	Faith	−1	96	72	96	54	80	5	100
Philosophers n=35 Luc 66 End 62									
Socrates	Phil	−5	92	60	92	91	83	5	100
David Hume	Phil	18	90	65	86	89	40	90	88
Zhuangzi	Phil	−4	88	62	79	89	46	65	88
Bertrand Russell	Phil	19	90	82	83	80	79	82	90
John Rawls	Phil	20	88	88	88	80	83	88	90
Laozi	Phil	−6	85	72	76	79	45	58	92
Nāgārjuna	Phil	2	92	85	77	78	53	55	95
Ludwig Wittgenstein	Phil	20	93	82	83	77	84	60	92
John Locke	Phil	17	88	82	76	75	53	82	90
Montesquieu	Phil	17	84	78	81	75	51	72	90
Hannah Arendt	Phil	20	86	66	79	73	76	72	86
Maimonides	Phil	12	88	78	66	72	67	78	90
Aristotle	Phil	−4	97	85	80	71	51	66	100
Epicurus	Phil	−4	82	78	72	71	50	72	82
Francis Bacon	Phil	16	88	88	79	71	77	40	90
Confucius	Phil	−6	88	90	69	69	66	60	100
Denis Diderot	Phil	18	88	68	87	68	73	66	82
Baruch Spinoza	Phil	17	92	60	83	65	72	40	90
Wang Yangming	Phil	15	86	78	71	65	69	70	88
Augustine of Hippo	Phil	4	90	82	52	63	66	55	98

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Immanuel Kant	Phil	18	96	82	68	63	38	90	95
Voltaire	Phil	17	90	72	77	63	66	78	88
Karl Popper	Phil	20	90	80	63	62	73	88	82
Averroes	Phil	12	88	65	68	61	70	35	82
René Descartes	Phil	16	90	72	71	59	51	48	92
Mencius	Phil	−4	84	75	70	57	62	68	90
Adi Shankara	Phil	8	92	88	55	56	54	55	92
Thomas Hobbes	Phil	17	84	68	44	56	70	72	88
Plato	Phil	−5	96	88	62	54	53	80	100
Friedrich Nietzsche	Phil	19	92	55	72	53	66	10	90
Zeno of Citium	Phil	−4	80	84	78	51	51	74	85
Seneca	Phil	−1	78	55	52	50	85	10	80
Jean-Paul Sartre	Phil	20	85	82	43	48	68	58	72
Jean-Jacques Rousseau	Phil	18	90	82	39	47	75	34	94
G. W. F. Hegel	Phil	18	92	85	55	37	49	82	88
Commanders n=27 Luc 66 End 59									
Yi Sun-sin	Cmdr	16	95	60	92	89	93	72	85
Li Jing	Cmdr	6	95	72	80	82	74	92	85
Subutai	Cmdr	12	96	65	58	81	56	78	60
Alexander Suvorov	Cmdr	18	96	83	77	79	81	68	80
Khalid ibn al-Walid	Cmdr	6	96	58	74	79	76	72	90
Xu Da	Cmdr	14	92	62	85	78	80	92	88
Duke of Wellington	Cmdr	18	96	68	85	74	75	97	85
Erich von Manstein	Cmdr	19	92	65	27	74	69	60	58
Robert E. Lee	Cmdr	19	92	55	78	74	63	72	32
Mikhail Tukhachevsky	Cmdr	19	88	85	60	71	84	10	80
Scipio Africanus	Cmdr	−3	92	60	74	68	77	70	90
Stanley McChrystal	Cmdr	20	82	68	64	68	52		
Huo Qubing	Cmdr	−2	97	48	81	67	88	80	75
Bernard Montgomery	Cmdr	19	84	68	72	66	60	82	72
Gustavus Adolphus	Cmdr	16	92	90	59	66	81	33	82
Yue Fei	Cmdr	12	99	50	83	66	80	8	30
George S. Patton	Cmdr	19	93	52	62	65	62	42	72
David Petraeus	Cmdr	20	85	72	58	64	49		
Duke of Marlborough	Cmdr	17	95	58	53	63	80	62	70
Belisarius	Cmdr	6	92	55	84	62	83	55	30
Heinz Guderian	Cmdr	19	90	68	29	61	55	68	60
Horatio Nelson	Cmdr	18	94	62	72	61	84	80	90
Georgy Zhukov	Cmdr	19	92	72	60	58	88	65	80
Erwin Rommel	Cmdr	19	90	50	54	53	81	55	10
Hannibal Barca	Cmdr	−3	97	45	58	47	94	35	15
Han Xin	Cmdr	−3	95	45	48	46	83	8	80
Publius Quinctilius Varus	Cmdr	−1	42	30	41	15	56	8	10
Venture n=6 Luc 65 End 81									
Georges Doriot	VC	19	82	85	79	72	79	85	88
Vinod Khosla	VC	20	84	72	59	67	65		
Don Valentine	VC	20	90	90	69	66	60	90	90
Eugene Kleiner	VC	20	85	80	64	66	63	80	80
Marc Andreessen	VC	20	85	80	57	64	64		
Tom Perkins	VC	20	85	72	57	55	57	68	78

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Pre-Qin thinkers n=41 Luc 62 End 47									
Fan Li	PreQ	−6	95	55	65	90	60	97	65
Sun Tzu (Sun Wu)	PreQ	−6	97	80	61	89	39	90	99
Lin Xiangru	PreQ	−4	90	38	86	87	78	82	65
Yue Yi	PreQ	−3	96	35	68	84	70	78	70
Sun Bin	PreQ	−4	92	38	59	79	61	72	86
Zichan	PreQ	−6	85	88	77	79	67	85	88
Duke Wen of Jin	PreQ	−7	92	75	61	78	84	85	72
Sunshu Ao	PreQ	−7	84	88	82	78	69	78	88
Wang Jian	PreQ	−3	93	45	63	78	71	85	70
Yan Ying	PreQ	−6	83	38	81	77	62	82	70
Li Mu	PreQ	−3	90	20	77	74	81	8	32
King Zhuang of Chu	PreQ	−7	84	63	71	73	80	88	72
Bai Qi	PreQ	−4	97	25	64	72	82	10	55
King Zhao of Yan	PreQ	−4	85	50	87	72	80	88	62
Guan Zhong	PreQ	−8	95	82	63	71	59	88	85
Baili Xi	PreQ	−7	82	68	76	69	73	88	80
Lord Xinling	PreQ	−3	84	22	72	69	79	22	52
Tian Dan	PreQ	−3	90	25	69	69	78	72	58
King Wuling of Zhao	PreQ	−4	88	70	71	67	66	2	68
Li Kui	PreQ	−5	82	92	84	67	73	88	90
Zhang Yi	PreQ	−4	95	20	17	67	68	55	45
Goujian, King of Yue	PreQ	−6	88	50	54	64	67	80	48
Duke Huan of Qi	PreQ	−8	80	45	62	63	65	3	50
Duke Mu of Qin	PreQ	−7	80	38	65	63	64	82	62
Lian Po	PreQ	−4	88	20	64	60	69	28	48
Wu Zixu	PreQ	−6	90	45	44	60	89	8	48
Qu Yuan	PreQ	−4	74	28	88	57	79	8	74
Lord Pingyuan	PreQ	−3	70	30	59	56	52	82	58
Lord Mengchang	PreQ	−4	72	16	33	54	63	45	30
Lü Buwei	PreQ	−3	85	45	34	54	81	12	62
Fan Ju	PreQ	−3	88	52	45	53	76	60	68
Li Si	PreQ	−3	92	90	62	52	82	5	88
Shang Yang	PreQ	−4	95	92	70	48	81	3	95
Wen Zhong	PreQ	−5	90	58	77	48	82	10	55
Lord Chunshen	PreQ	−3	74	44	50	47	73	8	46
Wu Qi	PreQ	−5	97	42	61	43	79	8	50
Jing Ke	PreQ	−3	42	5	54	38	74	8	35
Su Qin	PreQ	−4	92	15	23	35	80	8	35
Pang Juan	PreQ	−4	70	8	46	29	69	5	8
Duke Xiang of Song	PreQ	−7	35	15	47	24	75	32	20
Zhao Kuo	PreQ	−3	50	10	69	24	62	6	8
Financiers n=48 Luc 62 End 65									
Charlie Munger	Fin	20	90	78	76	87	57	85	82
Warren Buffett	Fin	20	95	80	83	85	51		
John Templeton	Fin	20	88	70	78	84	62	85	85
Amadeo Giannini	Fin	19	88	92	81	82	77	88	90
John Bogle	Fin	20	82	92	85	82	66	90	92
Stanley Druckenmiller	Fin	20	90	55	59	81	59		
Adam Smith	Fin	18	90	72	75	80	32	92	92
Benjamin Graham	Fin	19	90	72	77	80	66	90	88

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Bernard Baruch	Fin	19	88	40	65	80	61	65	40
Howard Marks	Fin	20	85	70	62	80	52		
James de Rothschild	Fin	18	90	90	56	79	62	88	88
George Soros	Fin	20	90	55	53	77	62		
David Tepper	Fin	20	86	58	54	76	51		
John Maynard Keynes	Fin	19	92	72	68	73	62	88	85
Stephen Girard	Fin	18	88	75	70	72	75	85	72
Paul Warburg	Fin	19	85	88	75	71	68	82	90
Henry Kravis	Fin	20	85	78	57	70	55		
Jamie Dimon	Fin	20	85	78	60	70	51		
Ken Griffin	Fin	20	88	70	49	70	57		
Abigail Johnson	Fin	20	78	80	60	69	51		
David Ricardo	Fin	18	92	65	60	69	61	88	85
Larry Fink	Fin	20	86	82	57	69	54		
Nathan Mayer Rothschild	Fin	18	95	92	50	69	69	90	90
George Peabody	Fin	18	82	84	71	67	66	85	90
Carl Icahn	Fin	20	85	55	48	66	57		
Mayer Amschel Rothschild	Fin	18	88	97	54	66	81	92	93
Ray Dalio	Fin	20	85	68	60	66	56		
Steve Cohen	Fin	20	88	66	46	66	60		
Cornelius Vanderbilt	Fin	18	92	70	40	65	76	72	38
J. P. Morgan	Fin	19	96	90	62	64	57	92	92
Jay Gould	Fin	19	95	55	22	63	74	38	30
Milton Friedman	Fin	20	90	75	63	63	40	85	85
Jacob Fugger	Fin	15	94	80	45	62	57	82	55
Hetty Green	Fin	19	88	30	37	59	57	60	35
Marcus Goldman	Fin	19	80	85	54	59	63	88	90
Bill Ackman	Fin	20	82	62	50	57	64		
Masayoshi Son	Fin	20	82	60	53	56	64		
Stephen Schwarzman	Fin	20	84	75	47	55	53		
Daniel Drew	Fin	18	78	25	23	45	68	20	15
Jesse Livermore	Fin	19	90	10	25	41	79	15	18
Jay Cooke	Fin	19	84	60	54	38	69	55	40
Andrew Mellon	Fin	19	90	85	54	36	53	80	82
Nicholas Biddle	Fin	18	85	45	55	35	60	30	22
Irving Fisher	Fin	19	88	45	55	30	38	30	70
Bernie Madoff	Fin	20	55	8	7	21	37	6	6
John Law	Fin	17	80	45	36	17	58	12	15
Charles Ponzi	Fin	19	45	10	12	16	43	8	8
Ivar Kreuger	Fin	19	82	30	32	16	60	10	12
Diplomats n=9 Luc 62 End 65									
Charles Maurice de Talleyrand	Dipl	18	95	55	45	85	86	65	52
George C. Marshall	Dipl	19	85	88	84	83	87	88	90
James A. Baker III	Dipl	20	85	72	68	77	64		
Dag Hammarskjöld	Dipl	20	82	80	83	70	78	78	85
Condoleezza Rice	Dipl	20	78	62	62	61	83		
Henry Kissinger	Dipl	20	90	65	48	58	81	60	80
Kofi Annan	Dipl	20	80	78	79	50	73	75	72
Klemens von Metternich	Dipl	18	90	82	51	48	90	68	80
Neville Chamberlain	Dipl	19	55	40	63	26	80	20	12

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Statesmen n=16 Luc 61 End 77									
Jean Monnet	Stat	19	72	92	87	83	62	88	90
Angela Merkel	Stat	20	84	72	74	79	60		
Robert Peel	Stat	18	85	72	84	76	62	74	78
Clement Attlee	Stat	19	78	88	84	72	73	88	90
Camillo Cavour	Stat	19	90	78	76	70	75	82	82
William Gladstone	Stat	19	90	78	84	68	56	78	80
David Ben-Gurion	Stat	19	85	85	69	65	81	82	82
Benjamin Disraeli	Stat	19	88	80	57	62	53	76	78
Zhou Enlai	Stat	19	88	72	61	61	85	72	72
Benito Juárez	Stat	19	80	72	76	60	80	76	74
Charles de Gaulle	Stat	19	88	90	80	60	84	88	90
Georges Clemenceau	Stat	19	86	62	63	57	79	60	55
Mohandas Gandhi	Stat	19	85	72	88	50	82	82	82
Julius Nyerere	Stat	20	72	60	82	47	71	55	52
Jawaharlal Nehru	Stat	19	82	85	80	34	83	84	85
Margaret Thatcher	Stat	20	88	75	58	29	72	72	75
Scientists n=14 Luc 60 End 71									
Charles Darwin	Sci	19	95	72	80	84	51	95	96
Geoffrey Hinton	Sci	20	90	70	75	83	50		
Katalin Karikó	Sci	20	88	60	80	83	61		
Jennifer Doudna	Sci	20	90	75	76	81	48		
Venki Ramakrishnan	Sci	20	88	74	75	81	46		
Alan Turing	Sci	20	96	55	78	75	81	88	95
Albert Einstein	Sci	19	98	65	80	68	83	92	95
J. Robert Oppenheimer	Sci	20	92	78	60	63	69	62	70
Marie Curie	Sci	19	92	75	81	63	72	92	90
Louis Pasteur	Sci	19	92	82	71	53	58	90	92
Isaac Newton	Sci	17	99	80	48	43	38	95	97
Nikola Tesla	Sci	19	90	35	56	33	69	55	82
Fritz Haber	Sci	19	90	65	38	25	76	40	88
Trofim Lysenko	Sci	19	30	40	15	10	31	5	5
CEOs n=32 Luc 60 End 67									
Andrew Grove	CEO	20	90	82	62	87	69	72	78
Lisa Su	CEO	20	88	76	65	82	52		
Sam Walton	CEO	20	92	88	57	82	57	90	95
Alfred P. Sloan	CEO	19	88	95	56	79	58	74	72
Reed Hastings	CEO	20	86	80	61	79	55		
Sundar Pichai	CEO	20	82	78	65	77	63		
Demis Hassabis	CEO	20	90	74	66	75	53		
Tim Cook	CEO	20	82	80	65	75	61		
Andy Jassy	CEO	20	80	72	63	74	51		
Jeff Bezos	CEO	20	90	82	54	74	64		
Jensen Huang	CEO	20	92	78	61	74	77		
Konosuke Matsushita	CEO	19	88	90	78	74	74	85	82
Satya Nadella	CEO	20	84	80	65	74	60		
Bill Gates	CEO	20	92	85	64	73	53		
Marc Benioff	CEO	20	82	76	61	70	51		
Bernard Arnault	CEO	20	90	84	50	68	50		
Bob Iger	CEO	20	84	78	58	68	48		
John D. Rockefeller	CEO	19	97	85	52	66	67	85	88

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Sam Altman	CEO	20	82	55	51	60	69		
Akio Morita	CEO	20	90	72	62	55	69	72	74
An Wang	CEO	20	85	35	57	54	63	15	12
Mark Zuckerberg	CEO	20	84	70	44	52	73		
Ray Kroc	CEO	20	88	85	43	51	55	90	92
Elon Musk	CEO	20	88	55	46	50	78		
Steve Jobs	CEO	20	96	80	58	49	76	90	92
Andrew Carnegie	CEO	19	90	70	53	45	63	78	80
Walt Disney	CEO	20	95	78	64	38	75	88	88
Edwin Land	CEO	20	95	45	73	33	57	25	22
Thomas J. Watson Sr.	CEO	19	85	92	49	33	65	85	85
Howard Hughes	CEO	20	80	40	34	19	52	45	55
Kenneth Lay	CEO	20	55	30	29	18	50	5	5
Henry Ford	CEO	19	90	50	34	17	60	68	82
Economists n=15 Luc 60 End 78									
John Stuart Mill	Econ	19	85	65	85	82	48	80	78
Amartya Sen	Econ	20	90	75	88	79	42		
Daron Acemoglu	Econ	20	90	72	83	78	49		
Friedrich Hayek	Econ	19	88	80	72	69	60	85	82
Joseph Schumpeter	Econ	19	90	70	77	65	59	85	82
Thorstein Veblen	Econ	19	80	55	80	65	65	65	62
Alfred Marshall	Econ	19	88	85	85	64	44	90	90
Paul Samuelson	Econ	20	95	88	85	59	43	92	92
Carl Menger	Econ	19	85	82	82	58	52	85	85
Léon Walras	Econ	19	90	65	77	57	55	88	85
Friedrich List	Econ	18	75	62	55	54	77	68	70
Thomas Malthus	Econ	18	78	60	73	51	45	72	65
John Kenneth Galbraith	Econ	20	78	55	75	48	44	50	52
Joan Robinson	Econ	20	88	65	61	39	53	68	65
Karl Marx	Econ	19	90	80	68	29	80	92	90
Sovereigns n=59 Luc 59 End 63									
Abraham Lincoln	Sov	19	85	85	41	86	82	82	90
Nelson Mandela	Sov	20	78	82	49	86	85	92	78
George Washington	Sov	18	58	90	43	84	82	96	96
Tokugawa Ieyasu	Sov	16	88	95	44	84	77	90	92
Zhao Kuangyin	Sov	10	80	78	33	84	70	68	85
Sejong the Great	Sov	14	70	85	53	83	48	82	90
Umar ibn al-Khattab	Sov	6	85	85	73	82	74	72	82
Augustus	Sov	−1	93	98	68	81	77	92	95
Liu Bang	Sov	−3	88	85	54	81	83	85	92
Yongzheng	Sov	17	88	80	23	80	75	68	82
Deng Xiaoping	Sov	20	88	82	59	78	75	78	88
Peter the Great	Sov	17	68	84	78	77	79	70	85
Elizabeth I	Sov	16	85	70	55	76	77	75	68
Li Shimin	Sov	6	87	82	69	76	75	78	88
Genghis Khan	Sov	12	96	72	87	73	91	68	78
Akbar	Sov	16	84	88	68	72	72	80	78
Marcus Aurelius	Sov	2	72	70	58	72	63	56	55
Otto von Bismarck	Sov	19	92	75	65	72	66	52	72
Catherine the Great	Sov	18	85	72	78	70	70	75	72
Lee Kuan Yew	Sov	20	85	90	56	70	59	85	90

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Kangxi	Sov	17	72	75	64	69	73	60	82
Konrad Adenauer	Sov	19	78	88	56	68	78	88	90
Diocletian	Sov	3	80	78	40	67	84	70	55
Napoleon	Sov	18	90	72	91	64	79	28	70
Cyrus the Great	Sov	−6	88	85	78	62	81	80	85
Wu Zetian	Sov	7	92	70	39	62	72	68	50
Atatürk	Sov	19	86	88	54	61	83	85	85
Charles V	Sov	15	72	72	70	61	77	68	78
Darius I	Sov	−6	82	88	72	59	72	76	85
Frederick the Great	Sov	18	82	80	75	59	74	72	82
Charlemagne	Sov	8	85	78	84	57	78	68	60
Mehmed II	Sov	15	88	82	85	56	73	76	85
Ashoka	Sov	−4	78	72	49	55	65	66	50
Pericles	Sov	−5	80	70	65	55	66	55	50
Franklin D. Roosevelt	Sov	19	86	90	61	54	87	82	88
Queen Victoria	Sov	19	60	82	68	54	40	85	85
Winston Churchill	Sov	19	88	75	66	54	62	80	70
Yongle Emperor	Sov	14	85	80	71	54	73	72	78
Mao Zedong	Sov	19	97	64	46	53	73	42	75
Saladin	Sov	12	84	60	73	53	77	55	40
Thomas Jefferson	Sov	18	75	85	59	51	56	82	88
Julius Caesar	Sov	−1	90	40	75	49	78	25	55
Toyotomi Hideyoshi	Sov	16	90	55	73	49	84	32	20
Kublai Khan	Sov	13	82	70	77	46	76	60	45
Theodore Roosevelt	Sov	19	80	80	70	46	51	78	80
Frederick Barbarossa	Sov	12	78	60	81	44	74	50	40
Emperor Wu of Han	Sov	−2	80	72	85	43	63	55	80
Joseph Stalin	Sov	19	95	62	54	43	65	30	55
Suleiman the Magnificent	Sov	15	85	85	77	41	68	72	85
Zhu Yuanzhang	Sov	14	90	68	35	38	75	38	78
Simón Bolívar	Sov	18	85	40	87	36	86	30	40
Qin Shi Huang	Sov	−3	90	88	69	33	78	40	80
Timur	Sov	14	90	40	85	33	82	32	38
Louis XIV	Sov	17	80	78	58	30	61	55	72
Empress Dowager Cixi	Sov	19	55	30	24	29	71	22	15
Alexander the Great	Sov	−4	92	35	95	28	77	26	55
Justinian I	Sov	5	78	85	76	27	78	60	72
Aurangzeb	Sov	17	82	55	79	24	76	30	25
Adolf Hitler	Sov	19	92	30	91	20	83	2	3
Educators n=9 Luc 58 End 76									
Jimmy Wales	Educ	20	72	80	82	75	60		
Abraham Flexner	Educ	19	82	85	86	74	42	85	88
Salman Khan	Educ	20	75	72	83	73	56		
Wilhelm von Humboldt	Educ	18	85	92	82	70	70	90	93
Ezra Cornell	Educ	19	70	80	75	68	52	82	88
Booker T. Washington	Educ	19	82	82	79	55	90	78	82
Maria Montessori	Educ	19	85	85	78	38	82	88	88
Leland Stanford	Educ	19	72	78	50	36	31	85	92
Amos Bronson Alcott	Educ	18	55	30	72	35	49	25	15
Rome n=19 Luc 58 End 41									
Vespasian	Rome	1	88	78	71	78	64	90	60

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Fabius Maximus	Rome	−3	90	55	79	75	63	85	86
Hadrian	Rome	1	92	85	70	75	58	65	82
Cincinnatus	Rome	−6	80	25	85	74	77	86	72
Constantine the Great	Rome	3	92	93	66	74	73	90	95
Trajan	Rome	1	95	82	74	71	62	82	85
Marcus Vipsanius Agrippa	Rome	−1	90	78	80	70	76	88	85
Sulla	Rome	−2	95	50	59	70	69	88	42
Gaius Marius	Rome	−2	92	88	75	61	69	25	80
Cato the Younger	Rome	−1	70	32	81	58	75	16	60
Tiberius Gracchus	Rome	−2	70	35	73	57	66	5	55
Cicero	Rome	−2	88	38	83	55	77	4	92
Pompey the Great	Rome	−2	92	45	62	55	70	3	30
Mark Antony	Rome	−1	85	28	44	50	78	15	25
Crassus	Rome	−2	70	30	45	49	63	5	15
Marcus Junius Brutus	Rome	−1	60	12	63	44	63	15	18
Nero	Rome	1	55	15	26	34	60	8	10
Commodus	Rome	2	42	8	15	24	64	10	6
Caligula	Rome	1	45	18	22	19	66	7	10
Central bankers n=9 Luc 57 End 56									
Paul Volcker	CBnk	20	88	85	88	81	86	90	90
Marriner Eccles	CBnk	19	78	85	79	79	85	78	88
Ben Bernanke	CBnk	20	84	78	72	78	49		
William McChesney Martin	CBnk	20	82	90	83	72	58	85	85
Christine Lagarde	CBnk	20	76	74	68	68	70		
Jerome Powell	CBnk	20	78	76	70	68	61		
Arthur Burns	CBnk	20	70	58	39	31	71	32	28
Montagu Norman	CBnk	19	72	62	57	24	86	40	42
Rudolf von Havenstein	CBnk	19	40	35	48	15	87	10	12
Presidents n=40 Luc 52 End 51									
Dwight D. Eisenhower	Pres	19	80	78	82	85	64	82	80
George H. W. Bush	Pres	20	76	72	74	82	66	70	74
Chester A. Arthur	Pres	19	58	62	61	74	55	55	78
James Madison	Pres	18	72	88	81	69	75	62	88
William McKinley	Pres	19	78	70	64	69	62	66	74
Barack Obama	Pres	20	85	66	63	68	79		
Rutherford B. Hayes	Pres	19	62	58	64	67	65	50	50
John Quincy Adams	Pres	18	60	50	86	66	67	58	40
Harry S. Truman	Pres	19	72	78	79	65	74	80	88
Bill Clinton	Pres	20	90	60	56	64	50		
Gerald Ford	Pres	20	58	52	71	63	68	55	50
James A. Garfield	Pres	19	68	38	65	63	62	48	42
John F. Kennedy	Pres	20	82	60	70	63	69	62	55
Martin Van Buren	Pres	18	85	76	48	63	71	42	62
James Monroe	Pres	18	68	78	77	62	58	74	78
William Howard Taft	Pres	19	50	58	70	62	51	55	60
Ulysses S. Grant	Pres	19	80	62	69	61	65	60	52
John Adams	Pres	18	58	62	76	59	65	68	65
Jimmy Carter	Pres	20	55	62	78	55	80	58	68
Joe Biden	Pres	20	62	60	60	55	74		
Calvin Coolidge	Pres	19	62	48	61	52	41	55	40
George W. Bush	Pres	20	72	60	61	52	88		

Name	Coh	C	Sk	In	Or	Luc	Df	End	Leg
Warren G. Harding	Pres	19	55	40	34	52	37	32	25
Woodrow Wilson	Pres	19	80	72	68	52	50	55	68
John Tyler	Pres	18	50	48	44	43	66	26	55
Richard Nixon	Pres	20	85	55	31	43	71	12	58
Ronald Reagan	Pres	20	88	72	59	43	59	72	78
Zachary Taylor	Pres	18	45	30	58	43	60	45	30
Grover Cleveland	Pres	19	72	62	68	41	72	62	60
William Henry Harrison	Pres	18	55	40	56	40	49	35	30
Donald Trump	Pres	20	86	28	24	39	76		
Benjamin Harrison	Pres	19	52	55	58	37	54	45	62
Millard Fillmore	Pres	18	52	40	35	35	62	30	25
Lyndon B. Johnson	Pres	20	95	85	60	33	71	60	88
Herbert Hoover	Pres	19	68	55	71	30	87	30	22
James K. Polk	Pres	18	88	74	59	29	71	76	82
James Buchanan	Pres	18	40	30	48	25	86	10	8
Andrew Jackson	Pres	18	88	60	48	23	54	58	70
Franklin Pierce	Pres	19	42	28	32	20	75	18	15
Andrew Johnson	Pres	19	45	30	48	17	77	28	25
Revolutionaries n=16 Luc 48 End 40									
José de San Martín	Rev	18	85	62	87	86	79	72	80
Tomáš Garrigue Masaryk	Rev	19	72	85	88	86	52	82	68
Michael Collins	Rev	19	80	58	58	80	83	35	78
Ho Chi Minh	Rev	19	85	75	60	63	83	66	82
Vladimir Lenin	Rev	19	92	80	49	56	88	52	85
Toussaint Louverture	Rev	18	80	60	68	51	92	25	58
Giuseppe Garibaldi	Rev	19	88	55	78	48	75	78	85
Emiliano Zapata	Rev	19	70	45	72	47	88	22	62
Leon Trotsky	Rev	19	88	62	51	43	87	15	45
Georges Danton	Rev	18	72	38	56	41	89	10	15
Oliver Cromwell	Rev	16	82	55	41	40	82	32	18
Sun Yat-sen	Rev	19	72	50	67	39	78	50	50
Che Guevara	Rev	20	70	35	71	27	76	10	35
Fidel Castro	Rev	20	90	78	41	25	72	58	82
Maximilien Robespierre	Rev	18	78	40	39	18	90	5	15
Kwame Nkrumah	Rev	20	78	55	45	17	54	28	40

Table 8 · The Atlas Index (`figures.csv`, n=459). All scores 0–100 except century (negative = BCE). Luc *and* End *shaded on the lucidity and ending scales. Generated directly from the open dataset by* `tools/gen_appendix.py`.

D Variables, Coding, and Regression Details (with full R1/R5 results)

- **Sample flow (why the n changes across analyses).** 459 figures carry a lucidity score → 404 carry both hand-coded outcomes and enter the main regressions (55 lack one outcome; two fictional cohorts excluded) → of the 404, 400 have a redactable biography and 397 return a valid blind lucidity (R2) → objective manner-of-death is coded for 449, and objective institution-survival years for the 136 institution-builders (R3a) → 317 carry the full six-axis profile → the depth validation battery (§4) is 20 subjects × 3 raters plus 12 synthetic biographies. Each analysis names its own n; the differences are these inclusion rules, not case selection.

- **Reproducibility.** All scores are produced by subagents of the Claude family (Anthropic), single-rater per pass, and every released quantity is regenerable from the open dataset and `tools/` scripts. Because the scores depend on the model, re-scoring under a future model is expected to drift —the cross-model-family re-run (R4) doubles as the reproducibility check.
- Lucidity derivation: 2–3 consequential phases per subject, per-phase lucidity $= \sqrt{(\text{Pattern} \times \text{Mystery})}$ rescaled to 0–100, cost-weighted; skill takes the peak.
- **Model fit:** personal survival $R^2 = 0.457$ (adj R^2 .450, rmse 20.7); organizational survival $R^2 = 0.719$ (adj R^2 .716, rmse 13.5). Lucidity univariate R^2: personal .247, organizational .387.
- **R1 differential test (stacked 808 rows = 404 $\times$ 2, standardized; subject-clustered SE):** skill $\Delta = +0.252$ (z = 5.46, p < .001; cohort-clustered z = 3.37); institution-building $\Delta = +0.022$ (z = 0.37); lucidity $\Delta = -0.065$ (z = -1.24, p = 0.21); orientation $\Delta = +0.114$ (z = 2.31 $\rightarrow$ clustered 1.64); difficulty $\Delta = +0.053$ (z = 1.37). Bootstrap (B = 2000) lucidity difference 95% CI [-0.165, +0.034], P(fate-tilt) = 0.886.
- **R5 dominance analysis (general dominance):** personal: institution-building 56.4% > lucidity 25.5% > difficulty 7.1% > skill 5.5% $\approx$ orientation 5.4%; organizational: institution-building 48.0% > lucidity 20.1% $\approx$ skill 19.6% > orientation 10.5% > difficulty 1.8%. **VIF:** skill 1.47 / institution 1.42 / orientation 1.40 / lucidity 1.71 / difficulty 1.03. **Specification curve (fate-tilt):** across 16 specs min -0.129 / max +0.071 / median -0.004, only 50% positive, univariate -0.125.
- Two-portrait data (pooled entries): Napoleon skill 89 / institution 72 / lucidity 64 / personal 28 / organizational 70; Washington skill 58 / institution 90 / lucidity 84 / personal 96 / organizational 96.

E R2 Outcome-Blind Rescore (full pool): Method and Results

Method. For every figure entering the regression, four steps are performed: (1) take the public Wikipedia biography; (2) an independent *redactor* agent deletes name, aliases, and titles (replaced uniformly by "Subject X"), the lead's summary judgements, and any Legacy, Reputation, Influence, Assessment, or pop-culture sections, and strips evaluative adjectives while preserving all conduct facts; (3) a *separate* independent *blind-judge* agent scores the six axes from the redacted text alone, knowing neither identity nor atlas score; (4) blind lucidity is computed by the same engine. All 400 pooled figures (of the 404 in the regression, four lacked a redactable biography) were re-scored, and 397 yielded a valid blind lucidity. The limitations (see §7, item 1) are that redaction and judging are both done by a model of the same family, single-rater, and imperfectly blinded.

Results (standardized; n = 397).

- Per-outcome OLS (blind): personal lucidity $\beta = +0.34$ (t 8.3), $\Delta R^2 = .089$; organizational $\beta = +0.08$ (t 2.4), $\Delta R^2 = .005$. Versus named (personal .27; organizational .21): the organizational coefficient mostly collapses while the personal one strengthens.
- Differential Wald (lucidity, org $-$ personal): $\Delta = -0.261$; subject z = -4.66, cohort z = -3.77 (named -0.065, z = -1.24). The fate-tilt goes from non-significant to p < .001.
- Skill enterprise-tilt (blind): $\Delta = +0.288$, subject z = 5.99 / cohort z = 3.28, stronger and in the same direction.
- Robustness (drop any single cohort; fate-tilt subject-clustered z): from -4.03 (drop spring-autumn) to -5.04 (drop mathematicians), so no single cohort drives it.

- Consistency: r(named, blind) ≈ 0.6, mean shift ≈ 0, a specific re-ordering (revered figures fall, harshly-docked ones rise) that is the signature of halo-removal rather than noise.

F R3a Objective-Outcome-Proxy Robustness (full pool): Method and Results

Purpose. This appendix cuts common-source from the *dependent-variable side*: it replaces the two hand-scored outcomes with external, records-based facts, holds the predictors at their atlas values, and re-runs the analysis. It is the "objective outcome proxy"half of agenda R3.

Coding (blind to atlas predictor scores; 459/459 extracted from cached Wikipedia, quote-anchored).

- **Personal fate = manner-of-death ordinal 0–4:** 0 executed/assassinated/killed/forced-suicide · 1 deposed & imprisoned/exiled-in-captivity · 2 fell-from-power/disgraced but died free · 3 natural death, out of power, in good standing · 4 died in power / at the height, honored.
- **Legacy survival = log1p(years the primary institution persisted after death)**, coded only for institution-builders; pure intellectual, artistic, scientific, or generalship legacies are N/A, which removes the institution-building-to-survival tautology from the DV side.
- Extraction: 58 model batches over bounded wiki excerpts, one quote-anchored record per figure (402 wiki / 24 mixed / 33 established; 392 high-confidence). Artifacts: `data/lucido-study/r3a-objective/`.

Results (standardized β; t).

- **Personal fate (objective, n = 449):** institution-building **+0.42 (t = 8.47, ΔR^2 .127)** > difficulty -0.17 > skill -0.10 > lucidity **+0.08 (t = 1.50, ns)** > orientation -0.06. **Deceased-only (n = 404):** institution-building t = 8.35 unchanged, lucidity **t = 0.99**.
- **Legacy survival (objective log years, builders n = 136):** institution-building **+0.36 (t = 3.72, ΔR^2 .091)** > orientation +0.15 > lucidity **-0.11 (ns)** > skill +0.09.
- **Stacked differential Wald (builders paired n = 133):** fate-tilt $\Delta = -0.22$, subject **z = -1.37, p = .17** (directional, far weaker than blind's z = -3.8); skill-tilt +0.11, z = 0.96.

Reading. Five observations follow. First, the master beam is not common-source: on objective survival institution-building still scores t = 3.7 and on objective fate t = 8.5, so the common-source worry is resolved. Second, the fate-tilt does not extend to objective manner-of-death: lucidity is insignificant on objective fate (t = 1.5, realized-only t = 1.0). Third, two causes operate together: halo-removal (the named legacy halo inflated lucidity) and construct compression (manner-of-death is not the same as a good end, since a composed Socrates drinking hemlock scores 0 objectively). Fourth, institution-building's link to objective fate is partly mechanical (rulers tend to die in power), so the cleaner test is on the survival side. Fifth, the five-level fate scale still detects institution-building at t = 8.5, so lucidity's null is not mere coarseness. The bottom line is that R3a *bounds* the fate-tilt: directionally preserved but not robust to a fully objective operationalization, and the one association surviving named, blind, and objective coding alike is institution-building's link to outcomes.

References (★ marks Institute-added sources)

- Baron & Hershey (1988). Outcome bias in decision evaluation. *JPSP*.
- ★Carpenter, Geletkanycz & Sanders (2004). Upper echelons revisited: composition, structure, process. *J. Management*.
- ★Eggers & Kaplan (2013). Cognition and capabilities: a multi-level perspective. *AoM Annals*.
- ★Finkelstein (2003). *Why Smart Executives Fail*. Portfolio.
- Fischhoff (1975). Hindsight ≠ foresight. *JEP:HPP*.
- Grossmann et al. (2010). *PNAS*.
- Hambrick & Mason (1984); Hambrick (2007). Upper echelons. *AMR*.
- Hayward & Hambrick (1997). CEO hubris. *ASQ*.
- ★Helfat & Peteraf (2015). Managerial cognitive capabilities. *SMJ*.
- Hiller & Hambrick (2005). Executive hubris. *SMJ*.
- ★Lawrence (1997). The black box of organizational demography. *Org. Science*.
- ★Malmendier & Tate (2005, *JF*; 2008, *JFE*). CEO overconfidence.
- ★Marquis & Tilcsik (2013). Imprinting. *AoM Annals*.
- McCall & Lombardo (1983). *Off the Track*. CCL.
- ★Meindl et al. (1985). Romance of leadership. *ASQ*.
- ★Moore & Healy (2008). The trouble with overconfidence. *Psych. Review*.
- ★Owens & Hekman (2012). *AMJ*. / ★Owens, Johnson & Mitchell (2013). *Org. Sci.*
- ★Roll (1986). The hubris hypothesis of corporate takeovers. *J. Business*.
- Shen & Cannella (2002). CEO succession. *AMJ*.
- Stinchcombe (1965). Social structure and organizations.
- Sternberg (1998). A balance theory of wisdom. *Rev. Gen. Psych.*
- ★Tetlock (2005). *Expert Political Judgment*. / ★Tetlock & Gardner (2015). *Superforecasting*.
- ★Weick (1995). *Sensemaking in Organizations*.