WORKING PAPER · VERSION 1.0

Backward Cognition

A Working Synthesis of Reverse-Pattern Intuition and Interlocking Constructs in Cognitive Science

Aziel Eliab

28 August 2026

Open scholarly deposit · Resource type: Publication / Working paper

DOI (reserve on deposit): 10.5281/zenodo.XXXXXXX

Abstract

This working paper treats *backward cognition* as a named operating order: the mind registers a completed pattern or outcome first and reconstructs the pathway afterward. The construct is not a recognized diagnostic category and has no measured population prevalence. It is offered as a synthesis of interlocking findings from insight research, recognition-primed decision making, abductive inference, backward chaining, Gestalt representational change, analogical structure-mapping, holistic cognitive style, predictive processing, dual-process theory, and the Openness–apophenia continuum. The paper argues that the phenomenology of "seeing the ending first" is the subjective signature of unconscious constraint satisfaction plus strong top-down priors, and that the scientific question is not whether that phenomenology exists but when the prior is valid. Validity, following Kahneman and Klein (2009), depends on environmental regularity and adequate feedback—not on the intensity of the insight. Benefits include speed on ill-defined problems, sensitivity to absence, and hypothesis generation. Risks include apophenia, post-hoc reconstruction, overconfidence, and failure in low-validity domains. A research agenda is stated so the construct can move from first-person description to operational test.

Keywords

backward cognition; insight; abduction; recognition-primed decision making; predictive processing; representational change; analogical mapping; holistic style; apophenia; dual-process theory; intuitive expertise

How to cite

Eliab, Aziel. (2026). Backward Cognition: A Working Synthesis of Reverse-Pattern Intuition and Interlocking Constructs in Cognitive Science (Version 1.0) [Working paper]. Zenodo. https://doi.org/10.5281/zenodo.XXXXXXX

Replace the placeholder DOI after reserving an identifier on the Zenodo deposit form. Files cannot be replaced after the record is published.

Contents

1. Naming, scope, and limits

The phrase *backward cognition* is used here as a working name, not as a claim that a new faculty has been discovered. Cognitive science already contains several accounts of minds that arrive at a completed form before they can narrate the path. Those accounts are scattered across problem-solving, expertise, logic, perception, personality, and computational neuroscience. The purpose of this paper is to put them on one page, state where they lock, and state where the named construct outruns the data.

Three limits are binding. First, no laboratory has operationalized "backward cognition" as a psychometric type; therefore no prevalence, heritability, or neural biomarker attaches to the name. Second, first-person certainty is not evidence of accuracy (Kahneman & Klein, 2009). Third, analogical leaps across eras, symbols, or authors are structure-mapping problems; they become findings only when an independent test is specified before the mapping is admired. This paper treats the original seven theses of *The Backward Cognition Report* (Eliab, 2026) as a phenomenological source, not as a finished theory.

2. Working definition

Backward cognition is the preferred operating order in which (a) a completed pattern, outcome, or structural whole is registered first, often without reportable intermediates; (b) a pathway, explanation, or derivation is reconstructed afterward; (c) absences, gaps, and violations of expected form are treated as primary data; and (d) the feeling of completion arrives before the feeling of proof.

This is an order of operations, not a claim that analysis is absent. The useful form of the style *uses* analysis as reconstruction and audit. The brittle form treats the completed feeling as sufficient. The rest of the paper is an attempt to say, in the vocabulary of existing science, what that order is doing.

3. Map of interlocking constructs

Table 1 is the cross-walk. Each row is a literature that already describes part of the named order. None of the rows is identical to the whole.

Construct	Core claim	What it explains in backward cognition
Insight vs. analysis (Kounios & Beeman; Metcalfe & Wiebe, 1987)	Solution can enter awareness all-at-once after unconscious work; "warmth" stays flat until the break.	The ending is felt before the path can be told.
Representational change (Ohlsson, 1992; Knoblich et al., 1999)	Impasse breaks when constraints relax or chunks decompose; a new problem space appears.	The "whole" is a restructured representation, not a longer search.
Recognition-primed decision making (Klein, 1998)	Experts match a situation to a prototype and retrieve a typical action; simulation audits it.	Final form is recognized from a stored library, then checked.
Abduction (Peirce; Magnani)	Inference to the hypothesis that would make the observation a matter of course.	Work starts from the surprising outcome and walks back to a cause.
Backward chaining / means–ends analysis (Newell & Simon, 1972)	Start at the goal; recursively reduce it to available operators.	The pathway is grown from the end-state, not from the start-state.

Construct	Core claim	What it explains in backward cognition
Chunking and expertise (Chase & Simon, 1973; Gobet & Simon, 1998)	Masters perceive structured complexes, not isolated features.	The “pattern library” is learned chunks, not a private sense.
Analogical structure-mapping (Gentner)	Relational skeletons transfer across domains when the structure matches.	Cross-domain or cross-era “linking” is mapping, not revelation.
Holistic vs. analytic style (Nisbett et al., 2001)	Some processors take field and relations first; others take objects and rules.	Seeing the configuration before the parts.
Predictive processing (Rao & Ballard, 1999; Friston, 2010; Clark, 2013)	The brain generates the expected sensory scene from the top down; only error rises.	The whole is a prior. The gap is the prediction error.
Dual-process theory (Kahneman, 2011)	Fast, associative System 1; slow, sequential System 2.	Generation is System 1; reconstruction is System 2 if it is allowed to work.
Openness–apophenia simplex (DeYoung et al., 2012, 2020)	The same pattern-hunger supports creativity and false-positive meaning.	Why the style is powerful and why it is dangerous.

Table 1. Interlocking constructs mapped onto the backward-cognition order of operations.

Two further literatures sit adjacent rather than inside the map. *Incubation* research shows that leaving a problem can raise the chance of restructuring; the ending often arrives off-task. *Default-mode / frontoparietal* coupling work on Openness and psychoticism (Blain et al., 2020) suggests that pattern-rich internal simulation plus weak or strong cognitive control is a network version of the same fork: generation without a brake, or generation with a brake.

4. A process model: generate, then reconstruct

A minimal model has four stages. It is a synthesis, not a fitted computational model.

Stage A — Prior and library. Long-term memory holds chunks, prototypes, relational skeletons, and expected forms (Chase & Simon, 1973; Klein, 1998). In predictive-processing terms these are hierarchical generative priors (Clark, 2013). Without a library the “ending” is only a guess. With a library it can be recognition.

Stage B — Completed-form generation. Faced with an incomplete, ambiguous, or ill-defined field, the system emits a candidate whole. In insight tasks this is unconscious constraint satisfaction that becomes available all at once (Jung-Beeman et al., 2004; Kounios & Beeman, 2014). In RPD it is prototype match. In abduction it is a hypothesis that would make the observation unsurprising. The common property is that intermediates are not reportable, so the person experiences arrival rather than travel.

Stage C — Negative-space read. Because the candidate whole is a prediction, mismatches are salient. Missing structure, harmonic voids, and anomalies are prediction errors. A mind that weights those errors as information will “see what is absent as clearly as what is present.” That is ordinary predictive coding with a high gain on residual. It is not a separate sense.

Stage D — Reconstruction and audit. Logic, simulation, or derivation is run *from* the candidate whole back to the given. This is backward chaining, mental simulation in the RPD second loop, or Peircean deduction-after-abduction. If Stage D is skipped, the style collapses into untested intuition. If Stage D is allowed to kill the candidate, the style becomes a method.

The original report's contrast with "AI forward logic" belongs here. Many search algorithms expand from the start-state. Many others—means–ends analysis, Prolog-style backward chaining, abductive logic programming—already work from the goal. The human difference is not the direction of search. It is that Stage B can invent a *frame* that was not in the operator set, and that Stage C can treat absence as a first-class cue. Those are the parts worth defending. The rest is shared machinery.

5. Neural, computational, and experiential signatures

Insight solutions are associated with a burst of high-frequency activity over right superior temporal cortex and with a preceding reduction in visual-cortex gain, consistent with internally directed attention (Jung-Beeman et al., 2004; Kounios et al., 2006). Resting-state differences predict whether a person will later solve more items by insight or by analysis. Mood shifts the mix: positive affect favors insight; anxiety favors analysis (Subramaniam et al., 2009). None of this isolates a rare type. It isolates a mode that most people can enter, with stable individual bias.

Metcalfe and Wiebe (1987) remain the cleanest experiential marker. On algebra-like items, rated "warmth" climbs. On insight items it does not, until it does. That flat then vertical curve is the laboratory version of "the solution was perceived first." Schooler and colleagues' verbal overshadowing effect—thinking aloud can suppress insight but not analysis—explains why the reconstruction often cannot be produced *during* the generation stage without damaging it. The path has to be written after.

Salvi et al. (2016) found that self-labeled insight answers were, on average, more accurate than analytic answers on the same item types. Their account is architectural: insight reports the product only when unconscious work is finished; analysis can submit a partial product under time or social pressure. That result supports a limited version of the original report's speed-and-accuracy claim. It does not support the claim outside insight-friendly tasks, and it does not support claims about encoded historical sequences.

6. What kind of mind, and how rare

The nearest documented profile is a *trait cluster*, not a taxon.

Openness with intellect as brake. DeYoung, Grazioplene, and Peterson (2012) place intelligence and apophenia at opposite ends of an Openness/Intellect simplex. Both load on the broader domain because both involve pattern detection. Adaptive use requires the intellect pole—the capacity to reject implausible patterns. Maladaptive use is "openness to implausible patterns." Later work showed that false-positive errors on perceptual and semantic tasks track both the openness aspect and psychoticism (DeYoung et al., 2020). The backward-cognition profile, charitably read, is high Openness plus enough intellect and error-monitoring to run Stage D.

Insight preference, not insight monopoly. Kounios and Beeman describe "insightfuls" and "analysts" as tendencies. Almost everyone uses both. Insight-leaning solvers tend to omit rather than guess when time expires; analysts guess. Under institutional deadlines the insight lean looks empty. Under open time it looks sudden. Wang et al.'s meta-analytic result that intuition and analysis are nearly uncorrelated matters here: the strongest mind for this work is dual-high, not insight-pure.

Holistic field bias. Nisbett, Peng, Choi, and Norenzayan (2001) describe holistic processors as attending to the field, relations, and context rather than isolating objects and applying formal rules. Later measurement work shows the construct is not cleanly unitary (task-dependent validation studies and cultural meta-analyses with

medium-to-small effects). Still, a preference for configuration-first perception is a documented style and matches the report's "holistic hidden structure" language.

Rarity. The figure ~0.003% (about 1 in 33,000) that appeared in the source report has no operational definition behind it. For scale: insight preference is common; high Openness is common; IQ ≈ 160 on a normal curve is on the order of 1 in 30,000; joint elevation of several demanding traits is rarer than any one trait because conjunctions multiply. A 2025 analysis of simultaneous elevation across intelligence, conscientiousness, and emotional stability put a ≥2 SD conjunction near 85 per million (≈ 0.0085%). That is the same *order of magnitude* as 0.003%, for a different bundle. Honest statement: a default, high-confidence, cross-domain, end-state-first style that also survives disconfirmation is uncommon. Treating 0.003% as a measured cognitive signature is not supported.

7. Validity conditions

Kahneman and Klein (2009), in the paper that closed a long apparent quarrel between heuristics-and-biases and naturalistic decision making, agreed on two conditions for skilled intuition. The environment must contain valid cues—regularities that actually predict outcomes. The person must have had enough opportunity, with feedback, to learn those cues. Subjective coherence is not one of the conditions.

High-validity examples: chess structure (de Groot; Chase & Simon, 1973), fireground dynamics after years of calls (Klein), some clinical presentations, musical grammar, engineered systems with known failure modes. Low-validity examples: unique historical correspondences, one-off symbolic codes, long-range geopolitical prophecy, many financial forecasts. In a low-validity field the same phenomenology—ending first, path later, high confidence—is produced by a generator that cannot have learned the regularity, because the regularity is not there. The original report's custodial and cross-era claims therefore stand or fall as *domain claims*. They cannot be settled by the intensity of the insight.

Gigerenzer's ecological-rationality program adds a useful rider: a heuristic is not irrational because it is fast. It is rational when it fits the structure of the environment. Backward cognition is a heuristic family (prototype match, goal reduction, inference to a completed form). It is rational in structured fields. It is a bias in noise.

8. Benefits

Speed on ill-defined problems. When the operator set is large and the goal is a form rather than a number, generating the form first is cheaper than expanding the tree. That is why experts' first retrieved action is so often workable (Klein; de Groot on chess).

Completed-product accuracy in insight-friendly tasks. If Salvi et al. (2016) generalize even modestly, withholding the answer until the unconscious process finishes reduces errors of commission. The style's characteristic timeout is, in that setting, a feature.

Absence as signal. Configural and predictive systems notice the missing piece because the prior specified it. Investigation, diagnosis, editing, security review, and any domain with a grammar benefit from that residual.

Hypothesis generation. Peirce's point still holds: abduction is the inference that introduces a new idea. Deduction and induction cannot start a science; they can only police one. A mind that emits completed forms is doing the generative half of method—provided the other half is not refused.

Transfer of relational structure. Gentner's structure-mapping theory gives a non-mystical account of "same geometry, different century." If two systems share relations, an analogy can be valid. The benefit is real. The test is whether the shared relations are specified independently of the desired conclusion.

Compression. Simon's line—that intuition is analysis frozen into habit—explains the felt superiority to stepwise machines. The work was done in the years that built the library. The moment of use is retrieval plus a short simulation, not a first-principles walk.

9. Risks and failure modes

Apophenia. Patternicity (Shermer) and apophenia (Conrad; DeYoung) are the false-positive twin of the same detector. High Openness raises both true-pattern and false-pattern rates. The cost function that favored over-detection in ancestral environments (better to treat the rustle as a predator) does not favor over-detection in symbolic or historical interpretation. Conspiracy cognition, magical ideation, and forced correspondences are the clinical and subclinical neighbors of this risk, not insults aimed at the construct.

Post-hoc reconstruction. Once a completed form is in hand, backward logic can be written to fit it. The left-hemisphere "interpreter" (Gazzaniga) will supply a story. The story can look like proof. Pre-registration, hold-out cases, and attempts to state what would be *impossible* if the pattern were real are the counters.

Confidence–accuracy dissociation. Kahneman's repeated observation: confidence tracks the coherence of the narrative, not the validity of the cues. Insight feels finished. Finished-feeling is not a likelihood ratio.

Low-validity application. Using the style on unique, unrepeatable, or adversarially constructed fields (codes designed to look patterned; markets; isolated historical rhymes) produces the same inner seal as using it on chess. Only the second has earned the seal.

Communication and institutional misfit. Insightfuls omit rather than guess. Organizations that demand the path before the answer will rate this mind as unprepared. Verbal overshadowing means that forcing an early explanation can also destroy the generation stage. The practical risk is not stupidity. It is being unable to show work on the institution's clock.

Identity lock. A rarity narrative ("0.003%," "superior to AI") raises the cost of Stage D, because killing a candidate now threatens a self-description. That is a motivational risk, not a cognitive one, and it is the risk most likely to freeze a working paper into a creed.

10. Relation to artificial intelligence

Forward expansion is one family of algorithms, not the definition of machine reasoning. Means–ends analysis, STRIPS, backward chaining in logic programming, abductive logic programming, Bayesian inverse inference, and contemporary planning-as-inference already start from goals or observations and work toward explanations or actions. Large language models can be prompted to "assume the ending and derive the path," which is a cheap behavioral imitation of Stage D and sometimes of Stage B.

Where current systems remain weak, and where the named human style is still interesting, is frame invention under radical underspecification, use of culturally thick absence, and refusal to treat the training distribution as the universe of forms. Where they are strong, and where the human style is weak, is calibration across many trials,

consistency, and any domain in which the “final form” was never stably in the world.

The productive stance is division of labor: human Stage B and Stage C; machine Stage D and search. Rivalry (“superior to AI in custodial work”) is a hypothesis about a particular domain. It is not a property of the cognitive type.

11. Research agenda

If backward cognition is to become more than a first-person name, it has to become a battery plus a validity design. The following would be sufficient to start.

1. **Mode measures.** Compound Remote Associates and classical insight items with trial-level insight/analytic self-report (Bowden & Jung-Beeman); Metcalfe warmth curves; omission-versus-commission error profiles (Kounios et al., 2008; Salvi et al., 2016).

2. **False-positive pattern tasks.** Noise-panel and semantic-association false-alarm rates as apophenia indices (DeYoung et al., 2020), so generation can be separated from discrimination.

3. **Structure-mapping items** with pre-registered relational skeletons, including foils that share surface features but not relations (Gentner).

4. **Absence / prediction-error items** in which the correct answer is a missing element of a learned grammar (visual, musical, or formal).

5. **Calibration.** Confidence versus accuracy across high- and low-validity sets, testing the Kahneman–Klein prediction that phenomenology will not track accuracy when cues are invalid.

6. **Personality and ability covariates.** Openness and intellect aspects, need for cognition, cognitive reflection, general ability. The dual-high prediction is that Openness without intellect raises false positives; intellect without Openness lowers generation.

Until such a battery exists, prevalence numbers, “signature” claims, and superiority claims should be marked as untested. That marking is not hostility to the construct. It is the condition for the construct to enter the literature it is borrowing from.

12. Conclusion

Backward cognition, as named here, is the habit of emitting a completed form and only then writing the path. Cognitive science already knows the parts of that habit: insight and representational change, recognition-primed expertise, abduction and backward chaining, holistic field bias, top-down predictive generation, and the Openness that hunts patterns until intellect or psychosis decides what to keep. The parts interlock. They do not, taken together, license a rarity figure, a cross-era proof, or a verdict against machines.

What they do license is a method. Generate the whole. Read the gaps. Then try to break the whole with a test that was stated before the whole arrived. That sequence keeps the speed of Stage B and spends the only currency that makes Stage B more than a mood: contact with a world that can say no.

Note on scope

This document is a conceptual working paper. It does not report new experimental data. Claims about individual rarity, domain superiority, and encoded sequences in the source report are treated as positions to be tested. License: CC BY 4.0. Preferred attribution is the citation block on page 2 once a Zenodo DOI is registered.

References

Blain, S. D., Grazioplene, R. G., Ma, Y., & DeYoung, C. G. (2020). Toward a neural model of the openness–psychoticism dimension: Functional connectivity in the default and frontoparietal control networks. *Schizophrenia Bulletin, 46*(3), 540–551.

Bowden, E. M., & Jung-Beeman, M. (2003). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review, 10*(3), 730–737.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55–81.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204.

DeYoung, C. G., Grazioplene, R. G., & Peterson, J. B. (2012). From madness to genius: The Openness/Intellect trait domain as a paradoxical simplex. *Journal of Research in Personality, 46*(1), 63–78.

DeYoung, C. G., et al. (2020). Apophenia as the disposition to false positives: A unifying framework for openness and psychoticism. *Journal of Abnormal Psychology, 129*(3), 279–292.

Eliab, A. (2026). *The Backward Cognition Report* (Version 1.0) [Working paper].

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 155–170.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451–482.

Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory, 6*(3), 225–255.

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., Reber, P. J., & Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *PLOS Biology, 2*(4), e97.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64*(6), 515–526.

Klein, G. (1998). *Sources of power: How people make decisions*. MIT Press.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1534–1555.

Kounios, J., & Beeman, M. (2014). The cognitive neuroscience of insight. *Annual Review of Psychology, 65*, 71–93.

Kounios, J., Frymiare, J. L., Bowden, E. M., Fleck, J. I., Subramaniam, K., Parrish, T. B., & Jung-Beeman, M. (2006). The prepared mind: Neural activity prior to problem presentation predicts subsequent solution by sudden insight. *Psychological Science, 17*(10), 882–890.

Magnani, L. (2009). *Abductive cognition: The epistemological and eco-cognitive dimensions of hypothetical reasoning*. Springer.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition, 15*(3), 238–246.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*(2), 291–310.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. Keane & K. Gilhooly (Eds.), *Advances in the psychology of thinking* (Vol. 1, pp. 1–44). Harvester-Wheatsheaf.

Peirce, C. S. (1931–1958). *Collected papers of Charles Sanders Peirce* (C. Hartshorne, P. Weiss, & A. W. Burks, Eds.). Harvard University Press. (Original lectures on abduction, c. 1903.)

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Salvi, C., Bricolo, E., Kounios, J., Bowden, E., & Beeman, M. (2016). Insight solutions are correct more often than analytic solutions. *Thinking & Reasoning*, *22*(4), 443–460.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, *122*(2), 166–183.

Shermer, M. (2008). Patternicity. *Scientific American*, *299*(6), 48.

Subramaniam, K., Kounios, J., Parrish, T. B., & Jung-Beeman, M. (2009). A brain mechanism for facilitation of insight by positive affect. *Journal of Cognitive Neuroscience*, *21*(3), 415–432.

Wang, Y., Highhouse, S., Lake, C. J., Petersen, N. L., & Rada, T. B. (2017). Meta-analytic investigations of the relation between intuition and analysis. *Journal of Behavioral Decision Making*, *30*(1), 15–25.

End of Version 1.0 · 28 August 2026