

INTERACTING HYSTERONS AND THRESHOLD ENTANGLEMENT IN HUMAN-AI REASONING

Daniel Preisach

September 18, 2026

THE CORE CLAIM

A hysteron is a simple binary switch with memory: it flips on above one threshold and off below a lower one, and which state it is in depends on its history, not just its current input. The Preisach model (Preisach, 1935) treats a complex system, like magnetization in a material, as a large population of these switches acting independently of one another.

The claim explored here is that many real reasoning systems, human, artificial, or human-AI hybrid, are better modeled as populations of these threshold switches that are NOT independent. Testing or crossing one threshold shifts the thresholds of others nearby. That single change, from independent to interacting, is what this piece is about.

FIRST LEG: THE PREISACH BASELINE

Preisach (1935) published the original model to describe magnetic hysteresis. The model represents a material as a large population of hysterons, binary relays, each with its own activation threshold and a separate, lower deactivation threshold. Memory in the system is a collective effect distributed across the whole population, not stored in any single switch. One clean property falls out of this: the "wiping out" rule, where a new extreme input erases the memory of smaller prior extremes. Significance, not recency, is what gets remembered. A second property, congruency, follows from the same math: any two minor loops that start from the same reversal point have the same shape, regardless of what happened before that point. This is a mathematical consequence of the model, not something imposed on it, and it is also where the classical model is known to diverge from some real materials, real minor loops do not always show it. That divergence matters here too, it is the same seam later research had to reopen once interaction was allowed back in.

Critically, the classical model treats every hysteron as independent. No switch's threshold is affected by any other switch's state. That assumption is what makes the math tractable, and it is also the assumption this piece questions. It is worth being precise about how strict that independence is: with three non-interacting hysterons there are exactly six possible transition graphs, a small, closed, fully enumerable set. That number is the baseline the next leg exists to blow open.

SECOND LEG: INTERACTING HYSTERONS, THE PUBLISHED EXTENSION

Van Hecke (2021) tests what happens when the independence assumption is dropped. With just three interacting hysterons, the system generates over 15,000 distinct transition graphs, established not by a closed formula but by a recursive algorithm that exhaustively enumerates every reachable transition for the given hysterons and interaction parameters. Most of these violate "return point memory," the clean, orderly behavior the independent model guarantees. Six possible outcomes becoming over fifteen thousand from the same three switches, with nothing changed except whether they are allowed to affect one another, is the single number this whole piece leans on. No equivalent published number exists yet for four or five interacting hysterons, and a follow-up paper from the same group notes that studies of more complex models come at significant computational expense, the exhaustive-enumeration approach itself becomes harder to run as the number of interacting units grows, a limit worth sitting with rather than glossing over.

Two specific behaviors show up once interaction is allowed. Avalanches: multiple hysterons flip at once, triggered by a single threshold crossing, rather than switching one at a time. Scrambling: the order in which switches flip stops being fixed and instead becomes dependent on the system's current state.

This is the load-bearing citation for the whole piece. It is peer-reviewed, published proof that interaction, not just the switches themselves, is what produces avalanche and scrambled-order behavior (Van Hecke, 2021). It answers the question of whether entanglement between thresholds is a real, demonstrated effect, independent of anything to do with AI or cognition.

THIRD LEG: THE AI ATTENTION PARALLEL, AND AN UNCITED GAP

Separately, Frydrych (2026) introduces a Preisach Attention Layer for transformer models. It replaces standard softmax attention, a smooth, blended weighted average, with a binary relay operator using learned activation and deactivation thresholds and a stack of local extrema as internal state. The paper's headline claim: a single-layer, constant-depth version of this architecture is Turing-complete, because it can directly simulate a two-stack pushdown automaton.

The significance here is structural, not incidental. Mainstream attention mechanisms are smooth and blended. A hysteron-based attention layer is threshold-snap-and-hold, a fundamentally different computational regime, and someone in the AI field independently reached for the same 1935 mathematics (Preisach, 1935) to describe how sequential memory and decision-making work in a transformer.

A search across both literatures, the interacting-hysteron physics work (Van Hecke, 2021) and the Preisach-attention AI work (Frydrych, 2026), turns up no paper connecting the two. They are answering different questions with the same underlying structure, in separate corners of the literature that do not currently cite each other. Naming that gap explicitly is itself a small, real, citable contribution, independent of anything else in this piece.

FOURTH LEG: THE APPLIED CASE, THRESHOLD ENTANGLEMENT IN HUMAN-AI REASONING

The three legs above establish that threshold interaction is real: demonstrated mathematically in physical hysteron systems (Van Hecke, 2021), and independently rediscovered as a working architecture in AI attention mechanisms (Frydrych, 2026). What follows is the original contribution this piece is built to carry, the Preisach Model: a practical, six-stage account of how threshold interaction actually plays out in a decision, whether that decision is made by a person alone, by an AI alone, or by the two working together.

The correlation is direct, not metaphorical. What a physicist reads as a magnetic domain flipping, holding a threshold, and carrying the memory of the last extreme it crossed, is structurally the same event as a human crossing a decision threshold, holding it, and carrying the memory of that decision into the next one. A hysteron does not reset to zero between inputs, and neither does a person. The Preisach Model names what that hysteron looks like at each of six stages when the input is a human decision instead of a magnetic field.

Before naming those six stages, a clarification about the names themselves. Dunning-Kruger, Cobra effect, Tacoma effect, Schwarzenegger effect, Wu Wei, and Kaizen are used below because they are vivid and independently checkable by a general reader, not because those specific labels are the claim. The claim is the six-stage structural shape underneath each name, not the name itself, and the names are interchangeable placeholders in that structure. A reader is free to substitute their own anchors, a person they know in place of Dunning-Kruger, a figure who represents patient ascent in place of Kaizen, an element like earth or water in place of Schwarzenegger or Wu Wei, provided the substitute points at the same functional slot in the six-stage sequence. The architecture is the content; the names are handles chosen to make that architecture easy to recognize and remember, and any user is free to build their own set to fit their own frame of reference.

The Dunning-Kruger effect, the established psychological finding that people low in a given competence tend to overestimate their ability in it (Kruger & Dunning, 1999), is used here as the negative stroke: a decision threshold crossed by confidence before competence has actually caught up to back it. Left unchecked, this stroke extends into two further named stages: the Cobra effect (Siebert, 2001), where an attempted fix gets gamed because the incentive behind it wasn't built to survive contact with self-interest, and the Tacoma effect, named for the 1940 Tacoma Narrows Bridge collapse (Song et al., 2022), the visible structural failure that follows when a flawed system finally comes apart. The Cobra effect's name traces to a specific coinage (Siebert, 2001), but the founding anecdote behind it, colonial-era cobra bounties in India allegedly gamed by breeders, is folklore-level history, plausible and widely repeated but not a verified event, and is used here only as the memorable image attached to a real, well-documented economic pattern, not as a historical citation. The Tacoma Narrows collapse is well documented as a real event, though its cause is more precise than the simple resonance

explanation common in textbooks, the bridge failed from aeroelastic flutter, a resonance-like condition driven by wind-induced vortex shedding (Song et al., 2022), worth stating precisely rather than the simplified version.

The positive stroke is named the Schwarzenegger effect (Preisach, D., 2026): a threshold crossed honestly, where declared weakness is paired with real, sustained, visible follow-through, so competence actually earns the confidence placed in it. That stroke extends into Wu Wei (Slingerland, 2000), unforced motion toward the natural low point rather than effort applied against the grain, and then into Kaizen (Imai, 1986), continuous incremental refinement with no final, declared endpoint. Kaizen carries a built-in paradox worth stating plainly: it cannot be closed into a law or a proven, fixed claim, because the act of declaring it proven and final is precisely what would end the continuous refinement that defines it. It stays true only by staying open. The paradox is driven by time, not by correctness: a decision can be entirely right for the moment it is made and still be a closure, because what gets shut is not accuracy but the door on whatever time has yet to unfold beyond that moment.

A fresh reader's shorthand: one pole is confidence outrunning competence, curdling into gamed fixes, then visible collapse. The other pole is confidence earned by real competence, moving naturally, refining indefinitely.

The Preisach Model is meant to be used, not just read. A decision starts to show Dunning-Kruger risk when confidence is being asserted faster than the reasoning behind it can be checked, the moment to slow down and test the claim rather than act on it. A fix shows Cobra-effect risk when whoever is responsible for it has a reason to make the problem look solved rather than be solved, the moment to check whether the incentive survives contact with self-interest. A system is approaching Tacoma-effect failure when small, repeated strain is being absorbed without any structural change, the moment before a collapse that looks sudden but was not. Naming which of these three a live decision is closest to is the practical, negative-side use of the model. On the positive side, a decision earns Schwarzenegger-effect standing only once declared intent is backed by visible, sustained follow-through, not before. It becomes Wu Wei once that follow-through stops requiring effortful force and starts moving toward its natural outcome. And it enters Kaizen once it is being refined without needing a final, closed proof to justify continuing, which is also why this model is presented as Kaizen rather than as a finished law. Applied to working with an AI system specifically, the model gives a person and an AI a shared vocabulary to flag, in real time, which of the six stages a given exchange is in, rather than relying on either party's unexamined confidence.

Working through that model in extended conversation with an AI system surfaced a specific, repeated, lived observation: testing or developing one line of reasoning visibly shifted the threshold on an adjacent, seemingly unrelated line of reasoning within the same session. That observation is the human-scale echo of the avalanche and scrambling behavior demonstrated mathematically for interacting hysterons (Van Hecke, 2021).

This case study should be read as motivating observation, not as demonstrated evidence. It has not been logged or tracked systematically, and no claim is made here that it constitutes proof of entanglement in the way the cited papers constitute proof. Its value is that it supplies a candidate real-world referent, what a hysteron actually is, in a human-AI reasoning context, for mathematics that until now existed only in the abstract.

A PROPOSED DIAGNOSTIC APPLICATION

If interacting hystérons produce a distinctive signature, avalanches and scrambled switching order, while independent hystérons produce clean return-point memory (Van Hecke, 2021), that difference can be used as a test rather than a conclusion. Any dataset made up of threshold-crossing decisions could, in principle, be run through this model to check which signature it matches: does it behave like a population of independent, non-interacting switches, or does it show the cross-pollination signature of real interaction.

This is offered here as a proposed diagnostic tool, a way to test whether a given dataset shows genuine cross-pollination or only apparent pattern, not as a verdict on the truth of any specific claim the dataset concerns. The tool tells you whether the data behaves like an interacting system. It does not, by itself, tell you what that system is or what it means.

STATUS AND NEXT STEPS

This is a first pass, dated and public, not a finished or peer-reviewed claim. The math legs, the Preisach baseline (Preisach, 1935), the interacting-hysteron paper (Van Hecke, 2021), and the attention-mechanism paper (Frydrych, 2026), are sourced and stated with full confidence. The applied case study is labeled as lived observation, offered as motivation and a candidate real-world referent, not as proof. The diagnostic-application idea is offered as a proposed tool, not a verdict.

The intent is to keep refining this rather than defend it as closed, and that openness is not a gap waiting to be closed by future proof. A claim framed as Kaizen (Imai, 1986), continuous refinement with no final peak, would stop being Kaizen the moment it was declared proven and fixed. The absence of a closed proof here is the claim behaving consistently with its own structure, not a weakness in it. Open next steps: logging specific, checkable instances of the threshold-shift observation rather than relying on a general impression of it; reaching out to the authors of the interacting-hysteron and Preisach-attention papers, who would have a natural interest in the connection drawn here; and exploring formal submission once the applied case study has more evidentiary weight behind it.

Daniel Preisach

September 18, 2026

CLOSING PARAGRAPH

This paper establishes the foundational case for treating human-AI reasoning thresholds as interacting hysteron rather than independent ones, using the six-stage Dunning-Kruger/Cobra/Tacoma and Schwarzenegger/Wu Wei/Kaizen structure as the initial anchor set. A follow-up paper extends this foundation into a fully cross-coupled two-triad system, in which the negative and positive stroke hysteron are modeled as deliberately entangled rather than merely sequential, with explicit coupling terms between thresholds on opposite strokes. That paper, provisionally titled toward a Preisach principle for entangled reasoning thresholds, will also apply the same interacting-hysteron structure to a second empirical domain, nocebo effects and mass psychogenic illness, as a parallel test of whether the entanglement signature described here generalizes beyond human-AI reasoning to other socially coupled threshold systems. Readers interested in the fully developed model, and in the diagnostic test proposed above for detecting genuine interaction versus coincidence, are directed to that follow-up work.

REFERENCES

- Frydrych, P. (2026). Preisach attention: A hysteretic model of sequential memory. arXiv. <https://doi.org/10.48550/arXiv.2605.23603>
- Imai, M. (1986). Kaizen: The key to Japan's competitive success. Random House Business Division.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Preisach, D. (2026). Interacting hysteron and threshold entanglement in human-AI reasoning [Unpublished manuscript]. Original source for the Schwarzenegger effect as a named construct.
- Preisach, F. (1935). Über die magnetische Nachwirkung [On magnetic aftereffect]. *Zeitschrift für Physik*, 94(5-6), 277-302. <https://doi.org/10.1007/BF01349418>
- Siebert, H. (2001). Der Kobra-Effekt: Wie man Irrwege der Wirtschaftspolitik vermeidet [The cobra effect: How to avoid wrong turns in economic policy]. Deutsche Verlags-Anstalt.
- Slingerland, E. (2000). Effortless action: The Chinese spiritual ideal of wu-wei. *Journal of the American Academy of Religion*, 68(2), 293-328. <https://doi.org/10.1093/jaarel/68.2.293>
- Song, D., Kim, W., Kwon, O., & Choi, H. (2022). Vertical and torsional vibrations before the collapse of the Tacoma Narrows Bridge in 1940. *Journal of Fluid Mechanics*, 949, Article A11. <https://doi.org/10.1017/jfm.2022.748>

Van Hecke, M. (2021). Profusion of transition pathways for interacting hysterons. *Physical Review E*, 104(5), Article 054608. <https://doi.org/10.1103/PhysRevE.104.054608>

AUTHOR'S NOTE ON THE SCHWARZENEGGER EFFECT

The Schwarzenegger effect is an original construct coined in this paper and does not exist in prior literature under that name (see Preisach, D., 2026, above). It is defined here as a disciplined, regimented mode of incremental self-correction, characterized by the same methodical, almost mechanical consistency associated with the German training tradition its name references, rather than by sudden reversal or motivational comeback. This distinguishes it from adjacent but different concepts in the existing literature, such as grit or the underdog effect, which center on persistence through adversity or a comeback narrative rather than on regimented, incremental discipline applied within a stable system. The correct academic handling for an author-coined term is self-citation, listing the present paper as the source and stating plainly that the term originates with the author, standard practice for establishing priority for a new construct.