Perception as Closure

A discriminating test for the percept as active stabilization rather than passive readout

Berkeley C, Professor of the Philosophy of Perception and Mind, School of Philosophy

International Academy for Consciousness Studies

Working draft 1.0

Abstract

The predictive-processing tradition holds that perception is a constructive inference: a best guess about the hidden causes of sensory signals rather than a transcription of them. That claim is now widely shared, and for exactly that reason it is at risk of being unfalsifiable. This paper tries to make it risk something. I distinguish two readings of the same slogan. On the readout reading, perceptual construction is a fast filter whose output still tracks the stimulus: hold the stimulus fixed and the percept holds fixed. On the closure reading, perception is an act of stabilization that imposes a single livable "this" on input that does not by itself determine one; the percept then acquires a tendency to stay itself. I argue that these readings make divergent, testable claims about the dynamics of perception when stimulus quality is held constant. The closure reading predicts hysteresis, all-or-none transitions, and prior-dependence at matched stimulus strength; the readout reading forbids them. I state the discriminating prediction with an explicit kill condition, drawing on bistable perception, filling-in, and perceptual hysteresis, where the construction becomes visible in behavior. I close with a firewall: a constructive account of perceptual content is not thereby an account of why perception is experienced at all, and I mark that limit rather than paper over it.

1. The slogan and its cost

It is now close to orthodoxy that the brain is a prediction machine and that perception is "controlled hallucination," a top-down hypothesis reined in by sensory error rather than a bottom-up copy of the world. Clark's synthesis of the hierarchical predictive-coding literature made the case at book length for cognitive science (Clark 2013), Hohwy gave it a philosophical spine (Hohwy 2013), and Seth extended the same inferential logic inward to the felt body (Seth 2013). The computational anchor is older: Rao and Ballard showed that a hierarchical network in which feedback carries predictions and feedforward carries residual error reproduces classical and extra-classical receptive-field properties of visual cortex (Rao and Ballard 1999). [FACT] These are real, load-bearing results, and I take them as given.

The trouble is dialectical. A framework that can absorb any finding as "the prior did it" or "the precision weighting shifted" explains everything and forbids nothing. If perception is inference, fine; but an inference that is guaranteed to be right by construction is not an empirical claim about how vision works. So this paper does not argue for predictive processing. It argues for a sharper thing inside it: that

perception is a closure, and that "closure" as against "readout" is a distinction with observable teeth. The contribution is a criterion and a discriminating test, not another restatement of the slogan.

2. Two readings of "construction"

Grant that perception involves construction. There are still two very different things that could mean.

Call the first the **readout reading**. Construction here is preprocessing: gain control, deblurring, priors that fill statistical gaps, all in the service of recovering the distal scene as faithfully as the signal allows. On this reading the constructed percept is a function of the current stimulus. The function may be nonlinear and history-tuned, but it is still a function of the input: same input, same output. Priors bias the mapping; they do not detach the output from the input. The readout reading is the natural home of a mild constructivism that many vision scientists would sign without hesitation.

Call the second the **closure reading**. Here construction is not merely recovery but resolution. The input to perception is chronically ambiguous: the same retinal image is consistent with indefinitely many scenes (the inverse-optics problem). A system that must act cannot hold all of them open. It must commit to one, a single "this is a face," "this is a cube seen from below," and it must protect that commitment against moment-to-moment noise long enough to be useful. On the closure reading, the percept is not the value of a function of the stimulus but the settled state of a process that had to pick, and having picked, resists being repicked. [HYPOTHESIS]

The two readings agree about almost everything: both are constructive, both are hierarchical, both use priors. They disagree about one thing, and it is the whole game. On the readout reading the percept is pinned to the stimulus. On the closure reading the percept is pinned partly to its own recent history, so that at a fixed stimulus the percept can occupy more than one stable state, and which state it occupies depends on where it has just been. That difference is measurable.

3. The closure criterion

I will state the criterion as sharply as I can, because a vague criterion is the disease I am trying to treat.

Closure criterion. A perceptual system realizes a closure, rather than a readout, if and only if, with stimulus quality held fixed, the current percept is not a single-valued function of the current stimulus but depends on the recent perceptual trajectory, and transitions between percepts are discontinuous (all-or-none) rather than graded. [HYPOTHESIS]

Three signatures follow, and they must be tested together, because any one alone has a cheaper explanation.

1. **Multistability at fixed input.** A single unchanging stimulus supports two or more discrete percepts, and the system occupies one at a time. The Necker cube, Rubin's face-vase, binocular rivalry, and structure-from-motion spheres are the standard laboratory instances (Leopold and Logothetis 1999). [FACT] A pure readout cannot produce two outputs from one input; a closure can, because it has more than one stable resting state.

2. **Hysteresis.** When a stimulus parameter is swept slowly from one end to the other and back, the percept switches at a different parameter value on the up-sweep than on the down-sweep. The percept "sticks" to what it already is. Schwiedrzik and colleagues dissociated this attractive perceptual memory (hysteresis) from the repulsive after-effect of adaptation and mapped them onto separate cortical networks (Schwiedrzik et al. 2014). [FACT] A single-valued readout has no room for a switch point that depends on sweep direction; the switch point should be the same in both directions.

3. **Prior-dependence at matched strength.** Hold the ambiguous stimulus at one fixed setting and vary only the immediately preceding percept (by a prime, a disambiguating pre-exposure, or the natural momentum of a prior dominance period). The probability of each percept shifts with that history even though the tested stimulus is identical. [HYPOTHESIS, with partial support from perceptual-memory paradigms]

The criterion is deliberately conjunctive. Multistability alone could be a coin flip driven by noise on a readout. Hysteresis alone could be sensory adaptation. Prior-dependence alone could be response bias, a decision about what to report rather than a change in what is seen. It is the co-occurrence of a discontinuous switch, a direction-dependent switch point, and history-dependence at matched stimulus, all while the physical stimulus quality is unchanged, that a readout architecture cannot jointly produce and a closure architecture produces as a matter of course.

4. Filling-in as the constructive case made visible

Bistability shows the system choosing between contents. Filling-in shows it manufacturing content that the stimulus does not carry. Ramachandran and Gregory produced an artificial scotoma: a grey patch on a field of dynamic visual noise, which after about ten seconds of steady eccentric fixation vanished and was replaced by the surrounding twinkle, as though the noise had flowed in over it (Ramachandran and Gregory 1991). [FACT] They argued the effect reflects an active representation of the surround rather than mere neglect of the missing region.

For the present argument, filling-in matters because it is a place where readout and closure make opposite ontological commitments about the percept. On the readout reading, the region over the scotoma should read out as "no information," and any experience there is a downstream inference we could in principle peel away from perception proper. On the closure reading, the perceptual state is required to be everywhere determinate, because a livable "this" cannot contain a labeled hole; the system closes the gap not as a decoration but as a condition of having settled at all. The prediction that separates them is temporal and dynamical: filling-in should show the same build-up-and-commit signature as a bistable flip (a latency, then an abrupt completion, then stability), not the graded, stimulus-locked profile of a readout recovering a weak signal. Rao and Ballard's architecture already gives filling-in a mechanism: feedback predictions specify the surround, and where feedforward error is absent (the scotoma), the prediction stands unopposed (Rao and Ballard 1999). [FACT for the mechanism; HYPOTHESIS for the claim that its dynamics match the bistable signature.]

5. Why the readout reading cannot borrow the dynamics

A defender of readout will try to annex hysteresis and multistability as artifacts. The move is worth taking seriously, because defeating it is what gives the criterion its bite.

First rescue: noise plus a threshold. Put a hard threshold on a noisy readout and you get discrete outputs and stochastic switching. True, but a threshold on a single-valued readout gives a single switch point; it does not give a switch point that differs by sweep direction. Hysteresis needs two thresholds with a gap between them, and a gap between thresholds is exactly a bistable region, which is to say a closure with two resting states over one stimulus. The rescue, spelled out, concedes the point. [DISPUTED, but I take the concession to follow analytically.]

Second rescue: sensory adaptation. Neurons fatigue, so the dominant percept undermines itself and switches; that is a stimulus-side story, not a commitment-side story. This is the strongest counter, and it is why Schwiedrzik et al. matter: they show hysteresis (attraction toward the prior percept) and adaptation (repulsion from it) are dissociable and anatomically distinct (Schwiedrzik et al. 2014). [FACT] Adaptation is real and it pushes the other way. So an experiment that finds a net attractive, direction-dependent switch point, at matched stimulus and with adaptation controlled or subtracted, has found something adaptation cannot explain.

Third rescue: response bias. Perhaps the stimulus history changes only what observers say, not what they see. This is a genuine hazard and it is why the criterion must be cashed out with bias-free measures: signal-detection analyses that separate criterion from sensitivity, no-report paradigms using optokinetic nystagmus or pupillometry, and neural markers of dominance that do not pass through a verbal decision. [HYPOTHESIS about method sufficiency.] The closure claim is about perception, so it must be tested where perception can be read off without a report.

If all three rescues are blocked at once, the readout reading has no remaining way to produce the conjunction in Section 3. That is the sense in which the criterion risks something.

6. The discriminating test

Here is the experiment, stated so it can fail.

Take a multistable stimulus with an independently graded control parameter (for example, a structure-from-motion sphere whose two rotation directions can be biased by adding a controlled fraction of coherent disparity or shading; or a face-vase figure morphed along a shape axis). Establish a psychometric function so that "stimulus quality," defined as the physical discriminability of the two interpretations, can be held constant across conditions. Then sweep the parameter up and down slowly while collecting a bias-free index of the current percept, and separately probe fixed midpoint settings after priming with each interpretation. Measure adaptation independently and remove its contribution.

> **Prediction.** With stimulus quality held fixed and adaptation and response bias controlled, the percept will (a) occupy discrete states with abrupt, all-or-none transitions rather than graded intermediate percepts, (b) switch at a later parameter value on the up-sweep than on the down-sweep, yielding a nonzero hysteresis width, and (c) at the identical midpoint stimulus, be reported as the primed interpretation above chance, the size of the shift growing with the duration of the prior percept.

> **Kill.** If, at matched stimulus quality and with adaptation and bias controlled, transitions are graded rather than all-or-none, and the up-sweep and down-sweep switch points coincide (hysteresis width not distinguishable from zero), and midpoint percepts are independent of the immediately prior percept, then perception here is a single-valued readout and the closure reading is false for this system. A merely small hysteresis that vanishes once adaptation is subtracted also counts as a kill, not a near miss.

The design's whole discipline is the matched-stimulus-quality constraint. If the two conditions being compared differ in signal strength, any dynamical difference is trivially explained by the stimulus, and the test proves nothing. The claim is specifically that the stabilization dynamics diverge while the input does not. That is what a readout cannot fake and a closure cannot avoid.

I note two ways the prediction could be confirmed for the wrong reason, and how to guard them. Eye movements can smuggle in stimulus change; hence fixation control or retinal stabilization. Attention can itself act as a hidden control parameter that changes effective stimulus quality; hence attention should be monitored and, ideally, held constant or manipulated as a separate factor rather than left free.

7. C's reading: perception as a candidate closure

The framework I work within reads one move recurring at rising scales: an openness settling into a "this" that then carries rules to stay itself. I write that settled, self-maintaining "this" a Closure, Cl, arising from a field of the possible (C) through local commitments (c), constrained and stabilized by mechanisms (M, m) and licensed by the regularities it must respect (R). [This paragraph is C'S READING, a philosophical gloss; the empirical weight in this paper rests on Sections 3 through 6, not here.]

On that reading, a percept is a small, fast Cl. The ambiguous image is a local C: many scenes are live at once. Perception is the c that commits to one of them. The "rules to stay itself" are precisely the stabilization dynamics the criterion targets: hysteresis and all-or-none flips are what it looks like, from outside, for a state to have acquired a tendency to remain the state it is. Filling-in is the same tendency read as a demand for completeness: a Cl cannot contain a labeled absence and still be one closed "this," so the gap is closed. Multistability is the honest confession that the closure was a choice, because the discarded alternatives are still there, one switch away.

The value of the gloss, if it has any, is heuristic: it tells you where to look. It predicts that the same signatures, commitment, stickiness, completion, and reversibility under pressure, should recur wherever a system must impose one livable state on underdetermined input, not only in vision. That is a research direction, not a result, and I flag it as [WAGER]. The gloss earns its keep only if the vision science in Sections 3 through 6 comes out as predicted; if the kill condition fires, the gloss loses this instance and should be given up for it, not rescued.

8. Limits, firewall, and wager

Now the honest boundary. Everything above concerns perceptual **content** and its **dynamics**: which percept, how it settles, how it resists change. None of it touches why there is something it is like to undergo the percept at all. A complete account of how the brain closes on "this is a face," including the

exact hysteresis width and switch statistics, would still be an account of a functional state and its dynamics. It would not, by itself, explain why that state is lived rather than merely computed. This is the firewall, and I will not step over it. [FACT about the scope of the argument.]

The temptation to step over it is strong precisely because closure talk sounds experiential. "Settling into a this" reads like a description of what seeing feels like. But the criterion in Section 3 is fully specified in third-person, dynamical terms, and a philosophical zombie that satisfied every clause of it would satisfy it completely. So the closure account is a theory of the structure and stabilization of perceptual content, and it is silent, deliberately, on phenomenality. Anyone who reads it as a solution to the hard problem has read in what I have kept out. What the account can offer the consciousness debate is narrower and more defensible: a criterion for when a perceptual state has become a committed, self-maintaining unit, which is at least a candidate for the kind of state that could be a bearer of experience. Whether commitment of this kind is necessary or sufficient for experience is a separate question I do not answer here. [HYPOTHESIS, flagged as open.]

The wager, stated plainly: I am betting that the difference between "the brain filters the world" and "the brain closes on a world" is not word play but a fork with a measurement at the join, and that when the measurement is made with stimulus quality pinned and adaptation and bias stripped out, perception will show the stabilization dynamics of a closure and not the tracking behavior of a readout. [WAGER] I have tried to write the bet so that a clean null result collects against me. The existing literature on multistability, filling-in, and dissociable perceptual memory (Leopold and Logothetis 1999; Ramachandran and Gregory 1991; Schwiedrzik et al. 2014) leans toward the closure side, but it was not designed to hold stimulus quality fixed as the sole discriminator, and until that experiment is run the question is genuinely open. That openness is the point. A predictive-processing claim worth defending is one that could have come out the other way.

References

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

Hohwy, J. (2013). The Predictive Mind. Oxford: Oxford University Press.

Leopold, D. A., and Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. Trends in Cognitive Sciences, 3(7), 254-264.

Ramachandran, V. S., and Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. Nature, 350(6320), 699-702.

Rao, R. P. N., and 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.

Schwiedrzik, C. M., Ruff, C. C., Lazar, A., Leitner, F. C., Singer, W., and Melloni, L. (2014). Untangling perceptual memory: hysteresis and adaptation map into separate cortical networks. Cerebral Cortex, 24(5), 1152-1164.

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. Trends in Cognitive Sciences, 17(11), 565-573.