Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Semantic Information Beyond Shannon: Meaning,Coherence, and Agency in Experiential Systems

Lolugu Vasudeva Prabhath

ud25011@students.iitmandi.ac.in

Indian Institute of Technology Mandi

Nitin Kumar

Indian Institute of Technology Mandi

Jyotiranjan Beuria

IKS Research Centre, ISS Delhi

Venkatesh H. Chembrolu

Indian Institute of Technology Mandi

Article

Keywords: Semantic information, Coherence, Agency, Meaning–Experience coupling, Information theory

Posted Date: August 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10284367/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Semantic Information Beyond Shannon: Meaning, Coherence, and Agency in Experiential Systems

Lolugu Vasudeva Prabhath[1*], Nitin Kumar[1], Jyotiranjan Beuria[2], Venkatesh H. Chembrolu[1]

[1*]IKSMHA Centre, Indian Institute of Technology Mandi, Mandi, Himachal Pradesh, India.
[2]IKS Research Centre, ISS Delhi, New Delhi, India.

*Corresponding author(s). E-mail(s): UD25011@students.iitmandi.ac.in;
Contributing authors: D24062@students.iitmandi.ac.in; jb@issdelhi.org; venkatesh@iitmandi.ac.in;

Abstract

Shannon's theory of communication achieved its generality by treating a message's meaning as irrelevant to the engineering problem of reproducing it at a receiver. That exclusion is harmless for coding and transmission, but becomes a liability once the same formalism is carried into fields that cannot proceed without some account of meaning, among them cognitive science, the philosophy of information, consciousness studies, and artificial systems that manipulate language while understanding none of what they manipulate. This paper addresses the question Shannon set aside: under what structural conditions does information become meaningful to the system that holds it? We propose a foundational framework in which semantic information is not a property added to syntactic information but emerges from a structural relation between representational organization and agentive alignment, formalized as Meaning–Experience Coupling (MEC) and expressed as a coherence functional $\mathbf{SI} = g(\boldsymbol{R}, \boldsymbol{A})$. Admissible functionals must satisfy five axioms - non-triviality, monotonicity, context sensitivity, continuity and normalization - which a weighted-sum form and a one-parameter power-mean family are both shown to satisfy. A numerical toy model illustrates the central consequence, that a fixed representational structure can carry different semantic information depending on the task in force. The account is situated relative to Shannon information and existing semantic, causal, and viability-based theories of information, as a complement rather than a replacement for each. Its contribution lies less in a new competing measure than in clarifying the conditions under which information becomes meaningful for an experiencing and acting system.

Keywords: Semantic information, Coherence, Agency, Meaning–Experience coupling, Information theory

1 Introduction

It is worth pausing over a situation that has become ordinary without becoming any less peculiar. A language model trained on a corpus that no person could read in many lifetimes will answer questions about bereavement, take a side in a moral dispute, and paraphrase the argument of a technical paper in the philosophy of mind, and yet, on almost every theory of content that has been seriously defended, none of this means anything to the model. The observation is not merely that present systems fall short of understanding, although they plainly do. It concerns the design. Systems of this kind are built to predict, to compress, and to reproduce their inputs faithfully, and meaning was excluded from the specification at the outset, in much the way Shannon excluded it when he set the engineering problem of communication apart from the semantic aspects of messages (Shannon 1948a,b).

There was nothing careless in that exclusion. It is the source of the theory's reach, and the reason its results transfer without modification from telegraphy to error-correcting codes to data compression and the treatment of noisy channels. The entropy of a source can be calculated with no information at all about what its symbols denote, and a channel capacity can be derived without attributing significance to any of the sequences that cross the channel. Stripping meaning out is what renders the mathematics at once tractable and universal. The cost of that universality is deferred rather than avoided, and it falls due when the same apparatus is imported into settings where meaning is not a layer one can remove.

Perception, cognition, and conscious life look different, though. Information there is not simply carried down a line and decoded at the far end - it gets sifted, weighted, pulled into shape, worked on. A perceptual cue does more than shave a little off somebody's uncertainty. It can arrive as a threat. It can arrive as an invitation, or as the one corner of a room a person recognizes, or as the very thing they had been straining to recall, or as the door they can finally leave by. No organism first receives a bare signal and then, in some second and separate act, works out whether that signal is worth anything. By the time any experience is present, it has usually been sorted already, with one feature standing out and that one receding, one contrast asking for a response while the next is left to itself. Granting as much, the pressing question stops being how many bits have arrived. It becomes the question of how the information came to be about anything in the first place, and about something for this particular subject rather than another.

The aim pursued here is accordingly narrower than a first glance might suggest, and it is worth stating plainly before the existing literature is surveyed. The paper is not chiefly in the business of proposing a new quantity to be measured, in the way an engineer might propose a new instrument reading. It is in the business of clarifying what has to be true of a system before any quantity earns the name semantic in the first place. Shannon's own achievement is instructive here: it lay not in measuring

communication more finely than his predecessors but in isolating, for the first time, the statistical structure that made a rigorous theory of communication possible at all. The question pursued below is the complementary one. Once the statistical structure of a signal has been isolated, what further structure must be present, and organized in what way, before that signal counts as meaningful to the system that receives it? This is a question about conceptual foundations before it is a question about formalism, and it is posed here in that spirit.

The distance between information and meaning has been narrowed from several directions, and a few of those efforts have narrowed it a long way. In the logical tradition the semantic content of a sentence is identified with the range of possibilities the sentence rules out (Bar-Hillel and Carnap 1953; Bar-Hillel 1964). Naturalistic theories tie content instead to causal covariation, to biological function, or to a history of selection (Dretske 1981; Millikan 1984; Papineau 1984). Floridi's philosophy of information treats semantic information as data that are at once well formed, meaningful, and truthful (Floridi 2010, 2011). The most operational of the recent proposals, due to Kolchinsky and Wolpert, identifies semantic information with that portion of an agent's information whose causal effect bears on whether the agent remains within the macrostates compatible with its own continued existence (Kolchinsky and Wolpert 2018). Each of these isolates something real. What none of them supplies, taken on its own, is an answer to the question being pressed here, which is how semantic information should be characterized from the standpoint of the very system for which information is information about anything at all.

The position advanced in this paper is that information is semantic to the extent that the distinctions present in experience are coherently coupled to the way an agent is oriented towards action. We call the coupling Meaning–Experience Coupling, abbreviated MEC, and the label is intended without metaphor. Meaning is not a tag attached to information once the processing has finished. It is not to be reduced to truth, to correlation, to reward, or to survival value. It is, on this view, a structural feature of organized experience, the feature by which distinctions become intelligible because they fall into line with what a system is set to do, to attend to, to anticipate, and to revise.

Stated formally the proposal is brief:

$$\mathrm{SI} = g(R, A). \tag{1}$$

In this expression R denotes the organization of the distinctions present in experience, A denotes agentive alignment, and g gauges how well the two are fitted together. Two consequences follow that can look counterintuitive at first. A state can be heavy with syntactic or statistical information and still carry next to nothing in the way of semantic information, provided it never enters the coherence at all. And a faint, unremarkable cue can turn out to be dense with meaning, so long as it reliably holds up orientation and action. So construed, semantic information will not behave like a fixed scalar. It comes in degrees; it shifts over time; it bends with context; and its anchor lies inside the system rather than in some reading taken from outside.

We are not putting forward a new universal yardstick to set beside entropy, mutual information, integrated information, or prediction error. The contribution is pitched

at the level of explanation. The classical measures describe relations among signals, states, probabilities, or causal structures, and they describe those relations well. What they do not describe is the relation through which distinctions in experience come to be meaningful for an agent, and it is this relation that is the target here. The same relation, we suggest, is the absent term in a surprising number of debates across consciousness science, cognitive science, the philosophy of information, and artificial intelligence, fields in which a great deal of information processing routinely takes place with no settled account of when, or whether, any of it rises to the level of meaning. MEC, in short, is not offered as a competitor to entropy or integrated information on their own ground; it specifies the explanatory conditions under which any of these quantities comes to matter to the system that carries them.

Five important aspects of this discussion are summarized below. The first is diagnostic, and amounts to insisting that semantic information is not information supplemented by truth, nor information supplemented by utility, nor information supplemented by correlation, and that running it together with any of these obscures the phenomenon. The second is conceptual, and develops MEC as a structural account of how meaning is sustained within experience rather than appended to it. The third is formal, and renders semantic information as a coherence functional defined over representational organization and agentive alignment. The fourth is axiomatic: a short list of conditions on admissible functionals, accompanied by a proof that a weighted-sum instance meets them and by a demonstration that a broader family does as well. The fifth is illustrative, and uses a numerical model together with a comparative analysis to place the proposal against the existing approaches while keeping it consistent with them.

The remainder is arranged as follows. Section 2 surveys the existing foundations and isolates the dimension along which they fall silent. Section 3 develops MEC. Section 4 sets out the formal framework, the axioms, the canonical instance, and the admissible family. Section 5 works through a context-sensitive toy model, including the cases of meaning failure, recovery, and drift. Section 6 relates the account to information theory, to consciousness research, and to artificial intelligence, indicates routes toward empirical work, and takes up objections. Section 7 concludes.

2 Existing Foundations and the Remaining Explanatory Gap

Theories of information and meaning disagree, at times very sharply, over what semantic content amounts to. Underneath the disagreement a good many of them share a commitment that is worth bringing into the open. They hold that the semantic status of a state is fixed by some relation that can be specified without any reference to the system's own lived or experiential organization. The relation in question is variously taken to be logical, causal, truth-conditional, teleological, thermodynamic, or predictive. Every one of these relations is genuine, and every one of them does explanatory work somewhere. The claim of this section is the more limited one that no single relation among them, considered in isolation, explains how information comes to be significant within the organization of experience. What follows should be read as

diagnosis rather than refutation. For each framework the aim is to locate one definite absence, the structural gap that turns out to be precisely what the rest of the paper is intended to supply.

2.1 Shannon Information and the Productive Exclusion of Meaning

Shannon information is a measure of uncertainty and of transmission structure. It is indifferent, and indifferent by design, to what a message means, to whether the message is true, to whether the receiver grasps it, and to whether it matters in any way to the receiver (Shannon 1948a; Cover and Thomas 2006; MacKay 2003). That indifference is what allows the theory to generalize across every medium and every code that has so far been devised.

For a wide class of scientific questions the indifference is exactly what one wants. A source entropy can be computed with no idea of what the symbols stand for. Mutual information between two variables can be estimated without first asking whether either variable is meaningful to anyone. The theory describes the structure of signals, and it describes that structure with great precision, but it does not describe semantic encounter and was never meant to do so.

Trouble shows up only where framework gets pressed into service in places that turn on the dimension it leaves out and familiarity in everyday life is enough to make the point. A room someone lived in for years holds almost no Shannon surprise, since nearly everything in it can be anticipated, and even so it stays thick with meaning. The half-seen glass on the table registers. So does the voice carrying in from the next room, and so does the quickest line to the door, each registering in its own key. Push the case the other way. A churning, chaotic display can brim with entropy and still leave the person watching it with almost nothing. Entropy and meaningfulness go separate ways here, and the reason is that they are answering two different questions. One asks how much a signal might have varied; the other asks what a given distinction comes to for a subject already in the middle of doing something. Newer refinements that split mutual information into redundant, unique, and synergistic parts sharpen the syntactic picture a good deal (Kolchinsky 2022), and yet by their own construction they stay within it, reporting how information is spread across variables and not what any of it means to anyone.

2.2 Logical and Truth-Conditional Semantics

Bar-Hillel and Carnap defined semantic information by way of logical probability, so that a sentence is informative in proportion to the possibilities it excludes, with $\inf(i) = -\log_2 m(i)$ (Bar-Hillel and Carnap 1953). Within the semantics of formal languages the definition has proved hard to better even after many decades. Its limits, though, are also obvious. Plenty of meaningful episodes have no sentence-shape to them at all. Think of a climber reading a rock face for where a hand will go next, or a musician who senses a chord turning a beat before any thought about it forms, or the bodily lurch of recognizing danger before there is a word for it. Not one of these passes through a proposition that could be marked true or false, and not one of them

needs to in order to steer what the agent does. On cases of this kind the framework has nothing to report, and the silence is not a flaw that sharper engineering would close, because semantic content there simply is, by stipulation, a relation that holds between sentences and possible worlds. Remove the sentences and the account is left with no object to describe (Bar-Hillel 1964).

Floridi's General Definition of Information adds in a third demand to well-formedness and meaningfulness: that the data be truthful (Floridi 2011). The truth clause buckles first. A soldier who built years of his life around a war that had, unknown to him, already ended held a belief that was false and, past any argument, meaningful, since it has ordered what he saw and did and braced for. To rule such content out as non-information because it fails to correspond inverts the order of explanation. Mistaken beliefs, misread cues, illusions, outdated maps all carry semantic load, and they carry it by arranging what an agent notices and how it then moves. Evolution may in fact tune perception to fitness payoffs rather than to literal truth (Hoffman 2019), in which case the gap between meaning and truth would be the rule rather than the exception. But the deeper problem sits earlier than that: the definition already assumes a semantic interpreter before it starts. Hand it one, and it can tell you what semantic information looks like; ask it how anything becomes an interpreter - a system with something at stake in meaning - and it has nothing to say, because that is exactly the question it took for granted from the start.

None of this makes truth beside the point. It makes truth less than the whole of the matter. A false distinction can be semantically powerful; a true one can be semantically thin; a statistically reliable one can be semantically unstable. The lesson is really one about order. Meaningful organization has to be accounted for first, and only then can truth come back in, as a norm laid over that organization rather than as the thing the organization is built from.

2.3 Causal, Teleosemantic, and Viability Accounts

Causal and teleosemantic theories naturalize content by tying internal states to environmental conditions, to biological functions, or to histories of selection (Dretske 1981; Millikan 1984; Papineau 1984). This was a real advance on the purely syntactic accounts. Representation came to be connected to world-involving relations, and misrepresentation became possible, so that a frog's prey detector can misfire, a perceptual state can signal a condition that is not present, and a sign can acquire its function through selection rather than by stipulation.

Even granting all of that, external tracking is still not meaning as meaning is lived. Set a domestic thermostat beside a visual cortex. Each tracks a physical quantity by way of a lawful dependency, the thermostat tracking room temperature and the visual cortex tracking luminance gradients, and the tracking relation, considered in the abstract, is the same sort of relation in the two cases. The thermostat, by ordinary assumption, means nothing to anyone, whereas the visual cortex means very nearly everything to the organism in which it sits. The causal and teleosemantic story has no principled means of marking this difference, for the simple reason that the difference does not reside in the external covariation. It resides in how the tracked distinctions are organized within the system's ongoing orientation. A cue becomes meaningful not

just because it covaries with something, however reliably and however anciently it does so, but because it enters a field of relevance within which it can be interpreted, put to use, and revised.

Kolchinsky and Wolpert offer the most operational proposal in this family, and also the proposal closest in spirit to the present one (Kolchinsky and Wolpert 2018). On their account semantic information is information whose causal effect on an agent makes the difference to whether the agent stays within its viable macrostates, that is, whether it maintains the nonequilibrium conditions required for its own persistence. The framework is substrate-neutral, it is properly graded, and it ties semantic status to autonomous agency rather than to an external interpreter, building on an older body of work that links the organization of a system to statistical mechanics (Atmanspacher and Scheingraber 1987). We have learned from it and borrowed from it. There are nonetheless three things it cannot say, and those three mark the boundary that interests us. It cannot say whether a glucose gradient is significant to a bacterium, as against merely causing the bacterium to swim. It cannot say how meaning shifts when an organism changes what it is trying to do, since on this account viability does not vary with the task. And it cannot say what meaning failure is like from within, in the case where a cue that once organized action quietly ceases to do so. All three demand access to the system's own encounter with information, whereas viability is assessed from an observer's vantage point: someone on the outside draws the boundary, fixes the macrostates, and traces the causal paths. The internal dimension is exactly what falls away. Viability, therefore, is better treated as an important constraint on meaning than as a theory of it.

2.4 Embodiment, Predictive Processing, and Consciousness

A long line of enactive and embodied work has argued that cognition is not passive representation at all but sense-making in action, the active making of a world of relevance by a system that holds itself together (Maturana and Varela 1980; Varela et al. 1991; Thompson 2007; Di Paolo et al. 2010). Phenomenology had reached a cousin of that thought earlier, coming at it from another side, holding that experience arrives already laid out as a field of salience and orientation and practical weight, and not as raw sense data over which meaning gets draped after the fact (Merleau-Ponty 1962; Zahavi 2019). Both traditions stand close to the view taken here, and closest of all in their refusal of the picture where neutral data come first and meaning is added on at the end.

Predictive processing and active inference supply formal machinery for treating perception and action as inference under a generative model, with prediction error as the quantity the system labors to reduce (Friston 2010; Hohwy 2013; Clark 2016; Friston et al. 2020). Integrated Information Theory characterizes consciousness in terms of irreducible causal integration, summarized in the quantity Φ (Tononi 2004; Tononi et al. 2016). Both programs have reshaped the study of mind, and we take both of them seriously. Neither, on its own, settles the semantic question. A system can drive its prediction error toward zero without anything being settled about why an inferred state should matter to it, or about what that state is even about for the system as opposed to merely correlated with. A network can register a high Φ, as a

suitably wired grid of resistors would, while not one of its integrated distinctions is about anything in particular. Deacon frames the residual worry as the absence of a properly semiotic dimension: a genuine sign refers to something for a subject that has a stake in the reference, whereas prediction is not yet reference and the minimization of error is not yet meaning (Deacon 2012). A semantic account has to specify not merely that some structure is integrated or predictive, but how distinctions become meaningful in relation to agency.

2.5 What the Survey Leaves Standing

The history here is older than it tends to look. Wiener insisted, as early as 1948, that information in purposive systems has a functional character which the bare statistics do not capture (Wiener 1948); Pattee asked how a molecule ever becomes a message and placed the answer in the context that reads the molecule rather than in the molecule itself (Pattee 1969); Roederer pressed much the same point for biological information several decades on (Roederer 2005); work on the algorithmic origins of life has argued that the transition to living matter is in part a transition in how information comes to exert causal control over its own substrate (Walker and Davies 2013); and careful taxonomies of information processing, computation, and cognition have shown repeatedly that these are distinct notions that get run together far too often (Piccinini and Scarantino 2011). Each of these gestured toward the dimension we are after. None of them grounded that dimension in the internal organization of experience, which is what the present account is meant to do.

Drawing the threads together, four requirements stand out. A theory of semantic information ought to be:

1. *Graded*. Meanings strengthen, weaken, fracture, and reassemble, so that semantic status is a matter of degree rather than an on/off flag, and a binary measure loses the phenomenon.
2. *Context-sensitive*. One and the same distinction can matter very differently under different tasks, different histories, or different modes of attention; alter what an agent is doing and a cue can travel from richly meaningful to almost empty without itself changing at all.
3. *Dynamic*. Meaning is held in place over time through action, feedback, and reinterpretation. It is maintained rather than stored, and a single snapshot is unable to register either its erosion or its recovery.
4. *Internally grounded*. Semantic information must be characterized in terms of organization within the system for which the information matters, even though that organization is continually constrained, and on occasion overruled, by the world.

MEC is constructed so as to meet these four requirements together, and meeting them together, as the survey suggests, is just what none of the existing accounts manages.

Two clarifications are in order before the positive account begins. The first is that the point is not that the surveyed approaches go wrong within their proper domains. Shannon information remains the correct tool for a very wide range of questions about signal structure; the causal and teleosemantic theories remain indispensable

for explaining how representational systems come to be world-involving; the viability accounts remain powerful wherever autonomous persistence is the explanatory target. The claim is only that not one of them, taken alone, delivers a satisfactory account of semantic information as information that is organized for an experiencing and acting system. The second clarification is that the account developed below is not a private introspective theory, and it does not equate meaning with whatever a subject happens to report. Its object is a structural relation that can in principle be probed through behavior, through perturbation, through adaptation, and through neural or computational modeling. The reason for beginning with experience is methodological. If the phenomenon to be explained is meaning for a subject, the theory has to retain the level at which relevance, intelligibility, and the guidance of action first appear. Removing the level of experience in the beginning makes the later reappearance of meaning puzzling to the theorist's own making.

3 Meaning–Experience Coupling

One assumption has to be made explicit before any formalism is introduced, because the entire account rests on it: *experience is already organized.*

What reaches us is not the neutral spread of sensation waiting to be put in order. From the very start it is a field of distinctions - some features push forward while others drop out of view, one contrast demanding a response while the next just sits there in the background. By the time any deliberate interpretation gets going, a great deal has already gone into settling what counts as relevant: the swing of attention, the cast of mood, what memory hands up, the posture of the body, the pull of the surrounding situation. One does not take in a patch of photons and only later, at some downstream stage, arrive at a face; one meets a face, and the photon-level story is an abstraction laid back over that meeting after the event (Merleau-Ponty 1962; Zahavi 2019). The organization on offer may be partial, it may be shaky, it may be flatly wrong. It is organization nevertheless.

It is worth being clear about what the assumption does and does not commit us to. It does not require commitment to any particular metaphysics of consciousness. It asserts only the following: for a deeper understanding of semantic information in relation to subjects, we require a level of description which is centered in the experience associated with it. Removing that level too early leaves us unable to raise the question we care about in a reasonable manner.

3.1 Experience Is Already Structured

We write R for the organization of the distinctions present in experience and call it the representational structure. The word "representational" is to be read loosely; it does not deal with inner pictures, explicit propositions, sentences in a head, or anything else of a linguistic cast. It points to the ordered pattern the distinctions fall into at a given moment: whatever is prominent and whatever stands in contrast to it, the relations that hold, the tensions that do not resolve, the points that stay put. An edge in the visual field, a voice one happens to know, or a route carried in memory, or a

charged glance met across a room. Any of these enters R, as long as it belongs to the structured field of experience.

However, on its own, R does not yet amount to meaning. Examples include sensory overload - flood a system with distinctions and it can still fail to gather a single one of them into anything worth calling it meaningful. Differentiation runs high, coherence has collapsed, the structure is all there and the sense-making is nowhere. The reverse teaches the same lesson from the far side. A sparse structure can carry real semantic weight, provided the few distinctions it does hold sit squarely on whatever the moment happens to be asking for. Differentiation is a precondition of meaning, although by itself it is not sufficient for meaning.

3.2 Agentive Alignment

What the overload case is missing is not more structure but orientation. Agents do not merely register distinctions; they are directed toward possibilities. They search, avoid, approach, ask, test, repair, conserve, signal, or wait. Even in perception that looks entirely passive, experience is shaped by readiness, by what counts as relevant, by which outcomes are treated as mattering, by which possibilities stand open, and by which errors would compel a revision.

Call this orientation A, agentive alignment. It takes in task structure, practical relevance, attentional weighting, policy tendencies, and the felt or functional organization of possible action. Like R, the alignment A belongs to the agent's own organization rather than being an externally imposed objective function. A reward signal in a model may approximate some aspect of A, but agentive alignment is wider than any reward function, since it is the manner in which the agent's current mode of engagement makes certain distinctions count while letting others fade. How that orientation is itself anchored is a substantive question on which the account stays neutral: self-model theories locate the anchor in a transparent model the system builds of itself as a whole (Metzinger 2003), somatic accounts root it in feeling and in the continuous regulation of the body (Damasio 2010), and MEC asks only that, whichever story is right, the orientation be the agent's own.

And A, taken by itself, is no more sufficient than R. A reactive system can be oriented toward a stimulus without interpreting anything whatsoever. A thermostat responds to temperature, a reflex arc responds to pressure, a sea slug withdraws from a touch, a simple controller drives an error signal toward zero. Each of these has functional organization, and none of them, so far as we can tell, has semantic information in the sense at issue here. One can pile on sensitivity, response types, and environmental coupling without ever crossing the gap by accumulation. What is missing is not action. What is missing is the coherent relation between action-orientation and experiential differentiation.

3.3 Why the Two Terms Must Not Be Collapsed

There is a standing temptation to read semantic information off R alone, to suppose that once a system represents the world in fine enough detail, meaning just is that

representational richness. The temptation fails because organization can outrun intelligibility. Picture a traveler stranded in a crowded foreign station, a beginner squinting at a circuit diagram that means nothing to her, a classifier handed an adversarial image: each of them is holding, or registering, a great many distinctions, and each is at a loss as to what to do with any of them. Differentiation with no organizing constraint behind it is not meaning.

The opposite reduction tempts just as much, and misleads just as badly. It reads semantic information off A alone, so that meaning becomes whatever ends up serving action, or reward, or survival, or task success. The proposal catches something real and then throws out the internal structure of meaning altogether. A piece of behavior can come out right by luck, or by reflex, or by sheer optimization, with nothing in the situation having become intelligible to the system that produced it. A controller can sit at zero error forever without ever once possessing a distinction that means anything to it. Acting correctly, in other words, is not the same thing as acting semantically, at least not in the sense this paper cares about.

Keeping R and A apart earns its keep on exactly the cases either reduction would flatten. Dreams pair vivid representational structure with an alignment that has gone weak or adrift. Reflexes carry plenty of action-orientation on top of almost no representational organization at all. Skilled action is the case where both show up together: the situation is differentiated in just the shape the action calls for, and the action keeps those same distinctions live as it goes. That is really the whole argument for treating meaning as a coupling rather than as representation on its own, or utility on its own.

3.4 Coupling as the Site of Meaning

Semantic information arises where R and A are coupled. A distinction becomes meaningful when its place in representational organization supports the agentive alignment that is currently structuring engagement. It follows that the very same physical signal can carry different semantic information in different contexts. A sound is irrelevant background while one reads, an urgent warning while one crosses a street, and a cue to enter on time during a performance. Nothing about it changes acoustically. What changes is the relation between the distinction and the orientation.

MEC is our name for the process that keeps this relation in place. Perception introduces or updates distinctions. Integration stabilizes them into a coherent organization, and where differentiation outpaces integration the result is fragmentation, while integration that runs ahead of fresh experience hardens into rigidity (Maturana and Varela 1980). Action tests the organization and, in testing it, alters the conditions under which the next distinctions will arise. Feedback may reinforce an alignment, weaken it, or reorganize it across several timescales (Varela et al. 1991; Thompson 2007). Meaning lives inside this loop, and it is maintained there provisionally and never once and for all, rather than being written down a single time as a fixed label.

Do not read Figure 1 as a circuit with SI sitting in it as one more processing stage. SI is evaluative - it registers how well R and A are keeping track of each other, nothing more mechanical than that. Matched, the distinctions are intelligible and action

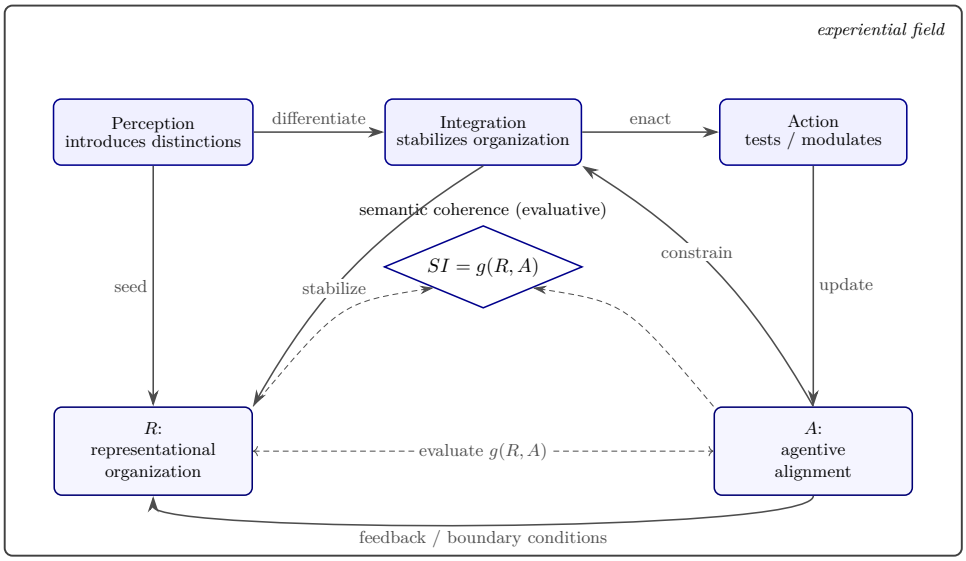


Fig. 1 The MEC loop. Perception feeds new distinctions into experience; integration works them into a representational organization R; and action checks whether R still backs up whatever the current agentive alignment A calls for, reshaping A in light of what it finds. $SI = g(R, A)$ is a gauge of how well R and A are keeping track of one another – not a further stage the causal chain passes through, but an evaluative index laid over the loop.

follows them. Pulled apart, experience can stay perfectly structured, action can keep happening, and meaning can nonetheless be quietly coming apart underneath both.

That same picture gives the framework a vocabulary for how meaning fails, instead of treating every failure as its own separate mystery. Distinctions can multiply and compete without ever settling - fragmentation. An alignment can outlast whatever once justified it - rigidity. Or a distinction can stay perceptually sharp while it stops backing the action it used to support - misalignment. Three names, but not three unrelated disorders; on this picture they are three regions of one space, the space of possible R–A coherence.

3.5 A Working Criterion

Put the conceptual core of the argument as a single criterion:

> An informational distinction has semantic information for an agent to the degree that it participates in a stable, context-sensitive coupling between representational organization and agentive alignment.

This criterion is built to block both reductions at once. Internal complexity will not do, since R without an alignment behind it falls short; utility will not do either, since an outcome can succeed while the organization behind it never comes together. Neither one, taken alone, is where the phenomenon lives - it lives in the coupling. The next section tries to say exactly what that means.

4 A Foundational Formal Framework for Semantic Information

The formalization below is kept deliberately thin - it is not offered as a finished computational model of semantic cognition, only as the structure any such model would have to instantiate to count as representing semantic information in the MEC sense.

4.1 Representational and Agentive Spaces

Let $\mathcal{R}$ be a space of representational structures. An element $R \in \mathcal{R}$ is a finite indexed family of experiential distinctions,

$$R = \{r_1, \ldots, r_n\}. \tag{2}$$

The r_i may stand for perceptual contrasts, bodily cues, affective tones, remembered relations, or any other structured distinctions; nothing here assumes them to be linguistic or propositional.

Let $\mathcal{A}$ be a space of agentive alignments. An element $A \in \mathcal{A}$ specifies a finite family of policies, tendencies, or action-relevant orientations,

$$A = \{\pi_1, \ldots, \pi_m\}. \tag{3}$$

Once again "policy" is to be read broadly: a policy may be an explicit plan, an attentional set, a regulatory tendency, or an embodied readiness to act.

For a given MEC instance, define a compatibility assignment

$$C : \{r_i\}_{i=1}^{n} \times \{\pi_j\}_{j=1}^{m} \to [0, 1], \tag{4}$$

where $C(r_i, \pi_j)$ records how well distinction r_i supports orientation or policy π_j in the context at hand. The value is not a truth value; it is a coherence score, and it is open to revision.

Finally, let

$$w : \{1, \ldots, n\} \times \{1, \ldots, m\} \to \mathbb{R}_{\geq 0}, \qquad \sum_{i,j} w_{ij} = 1, \tag{5}$$

be a relevance weighting. The weights encode which distinction–policy pairings are salient under the current agentive alignment, and they are the channel through which a change in A makes itself felt.

4.2 Compatibility and Relevance

The assignment C and the weighting w perform different jobs, and keeping them apart is essential. What $C(r_i, \pi_j)$ asks is whether a distinction can support a policy or orientation. What the weight w_{ij} asks is whether that pairing matters at the moment. A cue can be highly compatible with a policy that happens not to be in play; a

landmark is excellent for navigation, but if the agent is listening hard for a threat then the landmark's compatibility with navigation contributes almost nothing to present semantic information. Run the case the other way and a weak cue can carry great weight whenever the task at hand confers high relevance on it.

The model in this way separates available meaning from active meaning. Compatibility concerns the structure that could become meaningful under the right orientation; weighting concerns the orientation that is active now. There is nothing exotic in the distinction; everyday life is full of it. A word in a language one knows carries its latent structure whether or not one is attending to it. A faint twinge in the body goes unregistered until some worry, or some task, or a doctor's question hauls it into the foreground. A noise in the street slides from background to foreground with no change in its acoustics at all. Pattee's old question about how a molecule becomes a message has the same shape: the sequence is fixed, and whether it functions as a message turns on the cellular context doing the reading (Pattee 1969).

None of this makes g a performance score in disguise, either. An agent can stumble onto a winning policy while the compatibility structure underneath stays poor, and performance will climb anyway. Semantic information climbs on a different basis - the relation between distinctions and orientations actually getting more coherent - so the two quantities can and do pull apart, which is exactly what makes the account testable rather than circular. Robust semantic information ought to hold up under perturbation, carry over into a neighboring task, and find its footing again after a moderate disruption; bare performance owes you none of that.

4.3 Semantic Information as Coherence

Definition 1 (Semantic information) For a MEC instance (R, A, C, w), semantic information is a coherence functional

$$\mathrm{SI} = g(R, A). \tag{6}$$

The canonical weighted-sum instance is

$$g(R, A) = \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}\, C(r_i, \pi_j). \tag{7}$$

Equation (7) is not put forward as the only admissible measure. It is the simplest one. Other functionals will be better suited to hierarchical, nonlinear, probabilistic, or temporally extended systems, and a family of such functionals is given below. What every admissible functional has to have in common is that it measures coherence between the organization of distinctions and the organization of agency.

Figure 2 renders the intuition in geometric terms. The structures R and A overlap partially within experiential organization, and semantic information grows as distinctions come into closer alignment with task and policy constraints. The overlap is a graded affair, not a binary one: there is no threshold that is either cleared or not cleared, only more coherence or less of it.

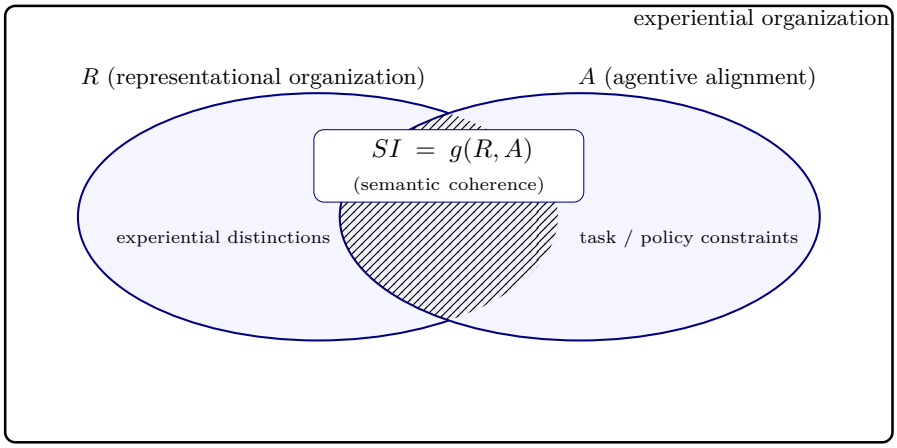


Fig. 2 Semantic information as coherence within experience. Representational organization R and agentive alignment A overlap partially within the experiential field, and the degree of that overlap, quantified by $SI = g(R, A)$, captures how well the experiential distinctions support stable, action-guiding interpretation. The graded overlap shown by the hatched region is meant to picture graded meaning rather than a semantic threshold that is either crossed or not.

4.4 Axioms for Admissible Coherence Functionals

An admissible coherence functional should satisfy five conditions. Each is given first as a formal statement and then in a plain reading; the two are paired in Table 1.

Axiom 1 (Non-triviality) *$g(R, A) > 0$ if and only if at least one active distinction–policy pairing has positive compatibility:*

$$\exists (i, j) : w_{ij} > 0 \text{ and } C(r_i, \pi_j) > 0.$$

Axiom 2 (Monotonicity) *If compatibility does not decrease for any active pairing and increases for at least one active pairing, then semantic information increases.*

Axiom 3 (Context sensitivity) *For some fixed representational structure R, there exist agentive alignments $A \neq A'$ such that*

$$g(R, A) \neq g(R, A').$$

Axiom 4 (Continuity) *Small changes in compatibility produce small changes in semantic information. Formally, g is continuous in C under pointwise convergence.*

Axiom 5 (Normalization) *For comparable MEC instances, semantic information is bounded:*

$$g(R, A) \in [0, 1].$$

Proposition 1 (Admissibility of the weighted-sum instance) *The functional in Equation (7) satisfies Axioms 1–5.*

Proof Each $w_{ij} \geq 0$, each $C(r_i, \pi_j) \in [0,1]$, and $\sum_{i,j} w_{ij} = 1$; the weighted sum is therefore nonnegative and bounded above by $\sum_{i,j} w_{ij} = 1$, which gives normalization. It is positive exactly when at least one positive-weight pairing has positive compatibility, which gives non-triviality. Should some active compatibility value increase while none decrease, the sum increases by the corresponding positive weighted amount, which gives monotonicity. A change in A changes the weights and the policy set, and whenever compatibility is not uniform across policies, which is the generic situation, a change in the weights changes the value of the sum, which gives context sensitivity. The weighted sum is linear in C and so continuous, which gives continuity. □

The axioms are not a circuitous way of singling out one formula. They carve out a class, of which the weighted sum is merely the most transparent member. To make the point concrete, consider the weighted power-mean family

$$g_p(R, A) = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}\, C(r_i, \pi_j)^{\,p} \right)^{1/p}, \qquad p > 0, \tag{8}$$

which recovers the canonical instance at $p = 1$, bends toward conjunctive behavior as $p \to 0^+$, rewarding alignments in which many pairings are jointly compatible, and bends toward disjunctive behavior as $p \to \infty$, rewarding a single strong pairing.

Proposition 2 (Admissibility of the power-mean family) *For every $p > 0$, the functional g_p in Equation (8) satisfies Axioms 1–5.*

Proof Write $S = \sum_{i,j} w_{ij} C_{ij}^p$. Since each $C_{ij} \in [0,1]$ we have $C_{ij}^p \in [0,1]$, so $S \in [0,1]$ by the same convex-combination argument as before, and because $S \in [0,1]$ its root $S^{1/p} \in [0,1]$, which gives normalization. The maps $x \mapsto x^p$ and $S \mapsto S^{1/p}$ are both strictly increasing on $[0,1]$, so increasing any active C_{ij} while decreasing none strictly increases S and hence g_p, which gives monotonicity. The functional is positive if and only if $S > 0$, that is, if and only if some active pairing has positive compatibility, which gives non-triviality. Both $x \mapsto x^p$ and $S \mapsto S^{1/p}$ are continuous on $[0,1]$, so g_p is continuous in C, which gives continuity. A change in A alters the weights, and for non-uniform C this alters S and therefore g_p, which gives context sensitivity. □

Remark 1 The exponent p earns its keep. It sets the trade between breadth and depth that the functional strikes: a small p suits a task that has to knit several weak cues together, while a large p suits one where a single decisive cue should carry the day on its own. Below, in the toy model, we simply fix $p = 1$ so the arithmetic stays easy to follow.

4.5 Dynamics

Semantic information hardly ever holds still. Perception keeps turning distinctions over; goals, attention, and bodily readiness keep nudging the agentive alignment around;

Table 1 Structural requirements for semantic information in the MEC framework. The five conditions define the admissible class; the weighted sum and the power-mean family of Eq. (8) both belong to it.

Property	**Formal role**	**Interpretive role**
Non-triviality	Positive SI requires at least one active compatible pairing	Meaning is absent only when no salient distinction supports any active orientation
Monotonicity	Increasing active compatibility increases SI	A better distinction–action fit strengthens semantic information
Context sensitivity	Fixed R can yield different values under different A	The same content can matter differently across tasks and situations
Continuity	Small compatibility shifts produce small SI shifts	Meaning can erode or recover gradually rather than by jumps
Normalization	$\mathrm{SI} \in [0, 1]$	Values are bounded and comparable within a model class

compatibility itself climbs or slides as action succeeds, fails, or runs into some constraint nobody had flagged in advance. Given all that movement, a time index stops being a nicety and becomes something closer to a requirement:

$$\mathrm{SI}(t) = g(R(t), A(t)) \,. \tag{9}$$

None of this has to be read as optimization. Compatibility shifts through habituation, through learning, through a reweighting of affect, through correction from other people, or simply through the body settling into a new adjustment. What MEC insists on is only that semantic stability be treated as something achieved rather than assumed. Take three cases that feel nothing alike from the inside. A cue that once organized action can lose its grip once outcomes stop matching it - in the formalism, nothing more dramatic than the relevant C value decaying, with no error term required to explain it. A cue that meant nothing can start to matter once a new task raises its relevance. And a cue that is traumatic or ambiguous can end up carrying far more weight than its reliability would justify, swamping the rest of the alignment. Different as these feel, the formalism treats them identically: motion in C, w, R, or A.

Figure 3 sketches one such trajectory - stable coherence, then disruption, then recovery. Nothing about it should be read as a universal law, and the numbers on the axis are illustrative rather than measured. What the picture earns its place by showing is simply this: meaning can start to slip before any behavioral failure is visible from outside, and it can come back later without the original interpretation returning intact. None of that would show up in a measure of SI taken once and left frozen. Deacon and Koutroufinis call the slow reconstitution of coherence across timescales dynamical depth (Deacon and Koutroufinis 2014); Equation (9) is little more than what is needed to get any grip on it at all.

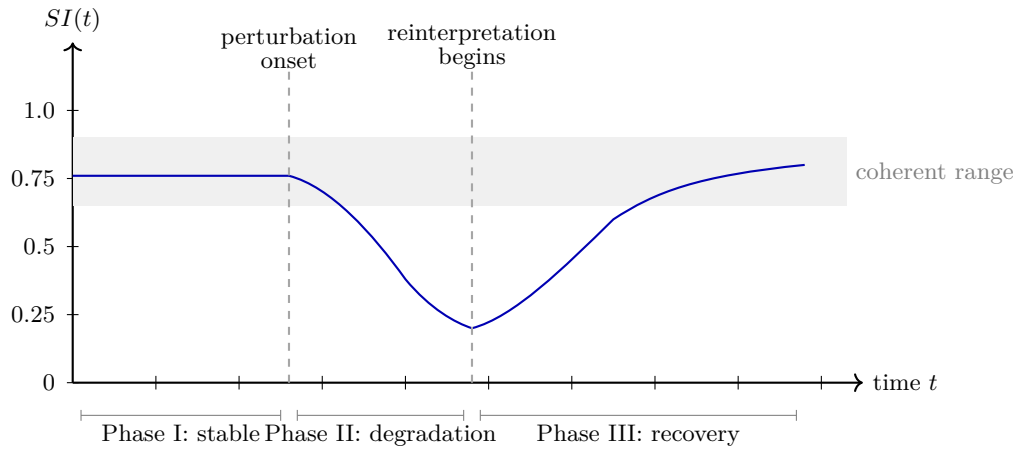


Fig. 3 Schematic evolution of semantic information $SI(t) = g(R(t), A(t))$ across three phases of engagement. *Phase I*: coherent alignment between R and A is stably maintained, with SI lying within the coherent range (shaded band). *Phase II*: an environmental or contextual perturbation disrupts alignment; SI falls below the coherent threshold as interpretations cease to guide action reliably. *Phase III*: reinterpretation under a revised agentive alignment gradually restores coherence, and SI recovers. The trajectory is qualitative; actual values depend on the specific choice of g and the dynamics of $R(t)$, $A(t)$.

5 Toy Model: Context-Sensitive Semantic Information

Picture a small agent feeling its way through a place it does not know. Hold the moment still. Its representational organization comprises three distinctions,

$$r_1 = \text{spatial cue}, \quad r_2 = \text{edge pattern}, \quad r_3 = \text{ambient tone},$$

and its agentive alignment comprises three policy orientations,

$$\pi_1 = \text{navigation}, \quad \pi_2 = \text{landmark identification}, \quad \pi_3 = \text{background monitoring}.$$

Let the compatibility matrix be

$$C = \begin{pmatrix} 0.90 & 0.20 & 0.05 \\ 0.30 & 0.75 & 0.10 \\ 0.10 & 0.15 & 0.80 \end{pmatrix}, \tag{10}$$

a near-diagonal pattern in which each distinction strongly supports one policy and supports the others only weakly.

Under uniform weighting, with $w_{ij} = 1/9$,

$$\mathrm{SI}_{\text{uniform}} = \frac{1}{9} \sum_{i,j} C_{ij} = \frac{3.35}{9} \approx 0.372. \tag{11}$$

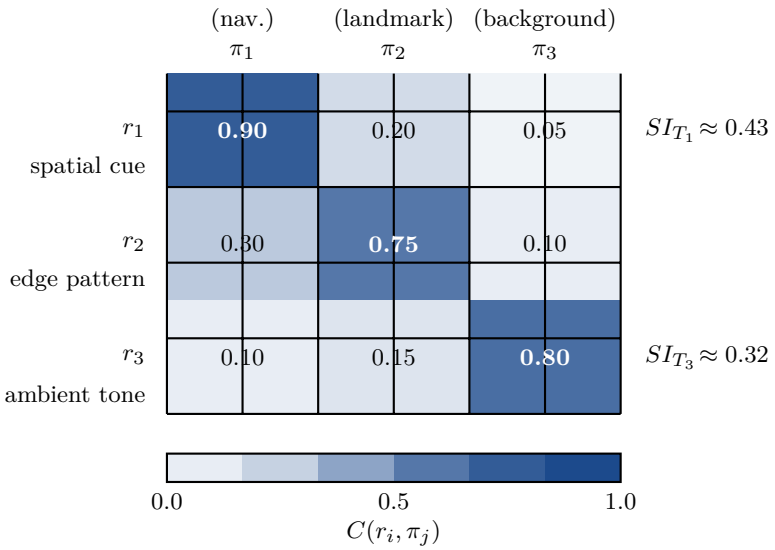


Fig. 4 Compatibility matrix $C(r_i, \pi_j)$ for the toy model, $n = m = 3$. Darker cells mean higher compatibility, and the bold values trace the dominant pairings along the near-diagonal. Concentrate the alignment on π_1 (task T_1) and semantic information comes to $SI_{T_1} \approx 0.43$; concentrate it on π_3 instead (task T_3) and the same R yields only $SI_{T_3} \approx 0.32$ – a 27% gap produced purely by the change in A, with R left untouched.

Where navigation dominates the alignment, so that $w_{i1} = 1/3$ and the rest are zero,

$$\mathrm{SI}_{T_1} = \frac{1}{3}(0.90 + 0.30 + 0.10) = \frac{1.30}{3} \approx 0.433. \tag{12}$$

Where background monitoring dominates, so that $w_{i3} = 1/3$,

$$\mathrm{SI}_{T_3} = \frac{1}{3}(0.05 + 0.10 + 0.80) = \frac{0.95}{3} \approx 0.317. \tag{13}$$

Nothing forces the alignment to concentrate on a single policy. A blended task that divides attention evenly between navigation and monitoring, so that $w_{i1} = w_{i3} = 1/6$, gives

$$\mathrm{SI}_{T_{1\&3}} = \frac{1}{6}\big[(0.90 + 0.30 + 0.10) + (0.05 + 0.10 + 0.80)\big] = \frac{2.25}{6} \approx 0.375, \tag{14}$$

an intermediate value, just as gradedness requires.

Look down the column of results and ask what actually moved. Not R - that holds fixed across every line, unchanged from row to row. What moves is the semantic information, and it moves quite a bit: about 0.317 at one end, about 0.433 at the other, roughly a 27% swing, traceable entirely to A. Call this context sensitivity if you like the formal name (it is Axiom 3, worked out in real numbers rather than asserted); the plainer way to put it is that the agent leans on the spatial cue while navigating and

Table 2 Compatibility matrix for the toy model. Bold values mark the dominant distinction–policy pairings along the near-diagonal.

	π_1 (navigation)	π_2 (landmark)	π_3 (background)	
r_1 (spatial cue)	**0.90**	0.20	0.05	
r_2 (edge pattern)	0.30	**0.75**	0.10	
r_3 (ambient tone)	0.10	0.15	**0.80**	
SI under navigation-dominant task T_1:				≈ 0.433
SI under background-monitoring task T_3:				≈ 0.317

on the ambient tone while monitoring, without R ever having to change to make that happen.

5.1 Meaning Failure

Suppose now that the corridor the spatial cue used to point down has been sealed off, so the cue itself turns unreliable. Nothing about r_1 has to change on the perceptual side - it can stay just as crisp as it ever was - but its compatibility with navigation collapses, $C(r_1, \pi_1) \to 0$, and under T_1 the semantic information falls toward

$$\frac{1}{3}(0 + 0.30 + 0.10) \approx 0.133.$$

Nothing here says the cue has stopped carrying information in the syntactic sense; it is still a perfectly distinguishable state. What it has lost is narrower and more specific: its semantic force for navigation, lost precisely because it no longer holds up the alignment that used to give it meaning. Most people have lived through a version of this without ever putting a number on it - a sign that once told you something goes quiet, not because anything about the sign changed, but because the action it used to underwrite stopped paying off.

5.2 Recovery Through Reorientation

Suppose instead the agent's task changes rather than the world - it stops navigating and starts monitoring ambient conditions. Nothing about R has moved, and yet r_3, which barely registered under T_1, becomes the distinction that matters most under T_3. Semantic information recovers here with no new information arriving and no representational change at all; what does the work is reorientation, A redrawing which pairings count. Pattee's old question about the molecule and the message has the same shape (Pattee 1969): the structure never moves, but the surround that reads it does, and meaning goes wherever that surround goes.

5.3 Semantic Drift

Compatibility rarely fails all at once. More often it erodes: a route grows a little less dependable each time it is used, a word picks up a connotation nobody intended, a

once-neutral gesture starts to carry a charge it never had before. None of these is a sudden relabeling. Each is what we are calling semantic drift, a slow reworking of the relation between distinction and orientation, and MEC represents it the same way it represents everything else that moves - as motion in C, w, R, and A over time, the dynamical depth Deacon and Koutroufinis describe (Deacon and Koutroufinis 2014). The particular update rule will differ from one domain to the next, and nothing here commits to any single one of them. What matters is only that the quantity tracking meaning can register gradual loss, partial recovery, and slow transformation - three things a measure taken once and then frozen simply cannot see.

6 Discussion

6.1 Relation to Classical Information Theory

The account is complementary to Shannon information rather than a correction of it. Shannon's question was specific and technically clean: given a source and a channel, how much does the received signal reduce the receiver's uncertainty? Our question is a different one, namely how well the distinctions in a subject's experiential field sustain the action-guidance on which the subject depends, and the difference is not cosmetic. Shannon's answer is measured in bits and ours in coherence, and neither denomination converts into the other.

The split is easiest to catch at the extremes. Take a practiced driver on a road she covers every day. Almost nothing surprises her, in Shannon's sense, because she can call it all before it arrives; and yet the lane markings and the set of the mirrors, the note of the engine, the pedestrian beginning to drift toward the curb, every bit of it is organized, and every bit of it is bound tight to what she will do in the next second. Now reverse the case. A high-entropy display can throw distinctions at a viewer by the dozen and still leave her nothing steady to steer by. Shannon information climbs with uncertainty reduced or with statistical dependence; MEC semantic information climbs with coherent, action-guiding organization. That the two can move in opposite directions is no fault in either of them. It is a sign that they are measuring different things, and that to collapse them together is to end up with something poorer than either on its own.

The same point accounts for why heaping on more data need not add an ounce of meaning. More data lifts semantic information where it firms up the compatibilities that matter. It depresses semantic information, just as easily, where it splinters attention or floods integration past what the system can hold together. The semantic question is never how much information is on hand; it is how well the information that matters is coupled to agency. Wiener had the intuition back in 1948, and what he lacked, and at that date could not have had, was any way to ground the functional character of information in the internal organization of experience instead of laying it down from outside (Wiener 1948). The R–A coupling is meant to be that grounding.

6.2 Relation to Theories of Consciousness

We are not offering a theory of consciousness, and it is best to say so plainly. The hard problem, the question why any physical process should be accompanied by experience at all (Chalmers 1996), is not our concern, and nothing said here bears on why physical systems should have experience or on what the neural correlates of phenomenal character might be. The question we take up is narrower and, we think, prior to that one: given an experiential field, what is it that organizes some of its distinctions as meaningful rather than as merely present? That target is not the target of IIT. The quantity Φ measures irreducible causal integration, and a system can be highly integrated while not one of its distinctions is coherently aligned with any agentive orientation; conversely, meaningful organization may depend on forms of embodiment, salience, and action that a single scalar integration measure does not capture (Tononi 2004; Tononi et al. 2016). Nor is it the target of predictive processing. Predictive models explain perception and action through the minimization of error under a generative model, and MEC can sit comfortably alongside them, since prediction error is a plausible mechanism by which compatibility gets updated (Friston 2010; Clark 2016; Hohwy 2013; Seth 2021). But minimizing error is not the same as meaning. A prediction becomes meaningful when it is taken up into the agent's field of relevance and action, and a perfectly calibrated prediction machine could be semantically empty. What Thompson calls sense-making, the capacity that marks living, minded systems off from merely reactive ones (Thompson 2007), is what $\mathrm{SI} = g(R, A)$ is meant to give a formal shape, without committing to any particular neural or computational substrate.

6.3 Relation to Artificial Intelligence

The framework also gives a fairly precise shape to a worry that keeps coming up about artificial systems. Contemporary models carry highly structured internal representations and turn in impressive performance across a wide range of tasks; in this vocabulary they plainly have plenty of R, and the objectives engineered into them approximate something like A. What is not settled is whether the two are actually coupled, in the sense of internally grounded experience, or merely coordinated from outside, the way any well-engineered piece of computation is coordinated.

None of this argues that artificial semantic information is impossible - only that it would take more to establish than symbol manipulation, statistical prediction, or reward optimization on their own. What a system would need is an internally sustained organization where distinctions are meaningful because they guide, and are in turn reshaped by, the system's own agentive orientation - and agentive orientation here is a structural feature of how the system exists, not a parameter someone bolts on from outside. The point connects to the old symbol grounding problem (Harnad 1990) and to the embodiment-oriented critique of disembodied computation that Dreyfus pressed, against the consensus of his day, more than half a century ago (Dreyfus 1972). What computation was said to miss was never raw processing power; it was engaged, situated involvement with a world. MEC gives that intuition a formal address in the A term of $g(R, A)$, together with the conditions under which the term could become non-trivial. Whether any artificial system could in principle sustain a genuine R–A

coupling is a further question, and not one we are positioned to settle here. What the framework does is make explicit what would have to hold were the answer to be yes.

6.4 Operational Interpretation and Empirical Routes

A theory that stays purely conceptual will not earn its keep. MEC is built so that its components can be approximated in empirical settings, though different domains will call for different operational choices, and Table 3 sets out candidate proxies across three of them.

In behavioral experiments, R can be approximated by discriminable stimulus dimensions, by reported salience, by attentional allocation, or by representational geometry inferred from patterns of response. The alignment A can be approximated by task instructions, by policy constraints, by motivational state, or by the structure of the affordances available (Gibson 1979). Compatibility C can then be estimated from how reliably a given distinction supports successful and interpretable action under a given alignment.

In neuroscience, candidate proxies for R include population-level representational structure, multivariate decoding geometry, and the stability of state-space trajectories across conditions. Candidate proxies for A include task-set signals, attentional control, affective salience, and motor preparation. MEC does not stake itself on any single neural signature. It predicts a relation: semantic information should be high where representational distinctions and agentive constraints co-vary in a stable and task-sensitive way, and a system that encodes many distinctions but fails to align them with task structure would show high representational richness together with low semantic coherence.

In artificial systems the same logic becomes a diagnostic. One holds the internal representations fixed and varies the task alignment, and then tests whether the system's use of those representations changes in a structured, recoverable, context-sensitive way or merely degrades. One perturbs the relevant cues and observes whether the system loses accuracy or instead reorganizes its interpretation. None of this would, by itself, demonstrate experience, but it would go some way toward separating brittle information processing from something closer to semantic organization. The underlying difficulty never really goes away, though: SI is not the kind of thing you read off a system the way you read a dial. All you can really do is triangulate it - hold R fixed and vary A, hold A fixed and vary compatibility, and so on - much as fitness, robustness, or the free-energy quantities get studied elsewhere: as constructs inferred from how a system responds to being systematically disturbed, not as objects sitting in plain view.

6.5 Comparative Position

Table 4 collects the contrast, and Figure 5 renders it schematically. The MEC column is not a claim that the other approaches fail. It is a claim that the target has moved. Rather than asking what information excludes, what it tracks, what it verifies, or what it contributes to viability, MEC asks how distinctions come to be coherently organized as meaningful for an agent. That is a question the other columns do not pose, which is why they leave its answer blank.

Table 3 Candidate operational proxies for the MEC components across three empirical domains. None is canonical; each is a starting point for triangulating SI through controlled variation rather than direct measurement.

Component	**Behavioral proxy**	**Neural proxy**	**Artificial-system proxy**
R (distinctions)	Discriminable stimulus dimensions; reported salience	Population geometry; multivariate decoding structure	Hidden-state / embedding geometry
A (alignment)	Task instructions; motivation; affordance structure	Task-set and attentional-control signals; motor preparation	Objective, prompt, or policy specification
C (compatibility)	Reliability of cue-guided successful action	Co-variation of representation with task demand	Cue-conditioned change in model behavior
w (relevance)	Attentional weighting under instruction	Gain / salience modulation	Attention or gating weights
Signature of high SI	Stable, transferable, recoverable performance	Task-sensitive, stable representation–demand coupling	Structured reorganization under task shift

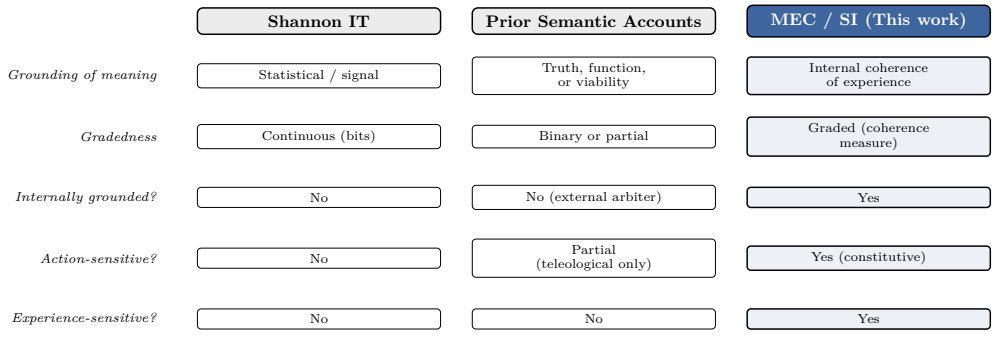


Fig. 5 Schematic comparison of three approaches to information and meaning across five explanatory dimensions. "Prior Semantic Accounts" aggregates the logical probability (Bar-Hillel & Carnap), causal/teleosemantic (Dretske, Millikan), truth-conditional (Floridi GDI), and viability-based (Kolchinsky & Wolpert) traditions. The MEC framework (right column, shaded) grounds semantic information internally, makes it graded and experience-sensitive, and treats agency as constitutive rather than external.

6.6 Implications for the Foundations of Information Science

The comparisons just drawn are not only about where MEC sits next to five rival accounts. They point to something about how information science treats its own vocabulary, and that is worth drawing out on its own.

Table 4 Comparison of information frameworks across explanatory dimensions. "Partial" indicates that the dimension is addressed only under restricted assumptions. The final column (shaded) is the present account.

Dimension	**Shannon**	**Logical sem.**	**Causal / teleo.**	**Floridi GDI**	**Viability**	**MEC / SI**
Primary grounding	Statistical structure	Logical probability	Tracking / function	Truthful meaningful data	Viability contribution	**Internal coherence**
Gradedness	Yes, in bits	Yes, in content measure	Partial	Limited	Yes	**Yes, as coherence**
Context sensitivity	No semantic context	Limited to language model	Limited	Limited	Partial	**Constitutive**
Agency	Not included	Not included	Often externalized	Presupposed interpreter	Autonomous viability	**Agentive alignment**
Experience	Not included	Not included	Not required	Not required	Not required	**Central domain**

Start with the status of semantic information itself. Most existing frameworks bolt meaning onto information after the fact, as a predicate a message either earns or fails to earn once it is already well formed. Here coherence between representational organization and agentive alignment is not bolted on; it is treated as an explanatory primitive in its own right, standing beside rather than beneath the statistical, logical, and causal primitives that already do work elsewhere in information theory. That $g(R, A)$ might not survive in its present form is beside the point. What changes, if the primitive is accepted at all, is what a theory of information now owes an explanation of.

There is also the older question of how syntax relates to semantics. Shannon's formalism tells you how a signal is structured. MEC tells you how a structured signal comes to matter to whatever carries it. It is tempting to picture these as two stops on one line, structure first and meaning after, but that picture is probably wrong; they read more like two different projections of the same information, answering two different questions that neither cognitive science nor artificial intelligence nor consciousness research can really afford to collapse into one.

A further consequence touches the old split between observer-independent and subject-dependent information. Shannon information does not need an interpreter: hand someone a source and a channel and the entropy falls out without anyone asking who, if anyone, is listening. Semantic information will not fall out that way, because it is defined over a system's own representational and agentive organization, and that organization is not something a third party can read off the channel. None of this makes semantic information private in any worrying sense; R, A, and their compatibility can still be probed from outside, through behavior, through perturbation, through modeling. But a science built to explain meaning has to leave room, somewhere in its apparatus, for the system's own vantage point, rather than averaging that vantage point away as noise.

One more point, closer to method than to metaphysics. The formalism in Section 4 was kept thin on purpose, and the thinness is not a placeholder waiting on some future tightening. A model of semantic information that closed every gap would likely have

closed the gap where the experience it is supposed to formalize used to sit, which is not progress so much as substitution. Some explananda, semantic information among them, are only captured by formalisms built to leave room for what they describe rather than to replace it.

None of these four observations stands alone, and none is meant to. Together they point toward treating semantic information as its own explanatory target, not a further quantity to be squeezed out of the information-theoretic machinery Shannon left behind, since that machinery, by design, was never built to hold it.

6.7 Objections

Is the account circular?

The charge is that meaning gets explained by leaning on experience, while experience is itself explained by leaning on meaning, so that in the end nothing has been explained. The charge would land if the account actually ran that way, and it does not. Experience is here taken as a given, something to be characterized and not derived. What gets characterized is the manner in which informational distinctions are structured within experience, which is a wholly different question from why there should be any experience to structure in the first place. Blur the two questions together and both get harder to pose, not easier. Experience is the domain we start from here; the coupling is what the analysis concludes. It was never meant to be the other way around.

Is internal grounding a form of relativism?

"Internally grounded" can sound, at first hearing, like license to assign SI however one pleases. That is not how the account works, because action is the check. Let a distinction keep failing to guide action and its compatibility decays; let an alignment ignore some standing feature of the environment and its coherence goes with it, not cheaply regained. The world still pushes back here - just not as a verdict on meaning handed down from outside, but as outcomes the system either absorbs or does not. Call it neither solipsism, nor pure internalism, nor ordinary externalism, since the constraint doing the work runs outward from inside the coupling rather than arriving as some external truth-condition.

Is semantic information measurable?

Not yet, and not directly, and there is no use pretending otherwise. No off-the-shelf protocol reads SI straight off a system. Then again, no such protocol reads off fitness, or biological free energy, or robustness, and not one of those has lain idle for the want of one. A great many scientific quantities start life as structural constructs and only get operationalized later. The candidate proxies, meanwhile, are entirely real: how stable an interpretation stays under perturbation, how long coherence takes to recover after a disruption, how the very same stimulus is processed differently as the agentive context shifts, how well action-guidance survives a change of setting. Wiring any of that to particular experimental designs is work that still has to be done.

Does the account reduce meaning to utility?

It does not. Utility is a matter of outcomes weighed against a goal or a preference ordering. Semantic information is a matter of how well distinctions cohere with agentive orientation. The first can shape the second, certainly. But a useful piece of behavior may be semantically shallow, if it never gets woven into organized experience, and an interpretation may be semantically rich even as it leads its holder straight into failure, because it goes on shaping what gets noticed and what gets expected and what gets revised. The two quantities are free to come apart. In practice they do.

Is the framework anthropocentric?

It trades in experience and agency, but it does not call for human language, human reflection, or human culture. It applies anywhere a system holds organized experiential distinctions together with an agentive alignment. Which biological or artificial systems actually meet that condition is an empirical and philosophical question, and a hard one, that runs past the boundaries of this paper.

6.8 Limitations and Future Work

The framework as it stands is foundational, and a fair amount of work is left undone. The dynamics of $R(t)$ and $A(t)$ want a fuller treatment than the sketch of Section 4.5 provides, one that would set down explicit update rules for compatibility, stability conditions for MEC, and a characterization of the passages among coherent, rigid, fragmented, and recovering regimes. The account is also, as it has been developed, largely intrapersonal, whereas human meaning is forever shared, negotiated, and propped up by institutions, so that carrying MEC over to intersubjective meaning would call for models of coupled agents whose alignments constrain one another (Di Paolo et al. 2010). The framework is substrate-neutral, which is a virtue at the conceptual stage and a debt at the empirical one, since real progress will turn on concrete bridges to neuroscience, to behavioral science, to artificial agents, and to embodied robotics, with attention, affordance perception, active inference, and dynamical-systems models of sense-making among the likeliest. The formal side is open as well: functionals beyond the power-mean family, and nonlinear ones in particular, may be needed to take in threshold effects, hierarchical relevance, conflict among competing policies, and long-range temporal dependence.

7 Conclusion

Most attempts to bridge syntax and semantics settle, in the end, into one of two resting places. In the first, meaning is dissolved downward into something sub-semantic, into signal statistics or causal tracking or evolutionary fitness. In the second, it is left hovering as a primitive the theory can gesture toward but never take apart. Neither resting place is ours, or so we have tried to see to.

Shannon information remains one of the genuine achievements of twentieth-century science, and it is great precisely because it leaves meaning out. The omission becomes a problem only when the same framework is asked to explain systems for which

meaning is central. For those systems we have argued that semantic information is best understood as a distinct kind of coherence, the alignment between representational organization and agentive orientation within experience. R captures how distinctions are organized, A captures how action, attention, and relevance are oriented, and $\mathrm{SI} = g(R, A)$ is just a measure of how well the two are tracking each other. Graded, context-sensitive, dynamic, internally grounded - the account is built to be all four at once. A weighted-sum instance satisfies the axioms, an entire power-mean family does as well, and a small toy model shows semantic information shifting with the task while the representational structure is held fixed.

Where the proposal is modest and where it is ambitious can be said in a breath. It does not solve consciousness, it does not reduce meaning to physics, and it hands over no finished metric; and the problems left standing are not minor, given that the full dynamical treatment of $\mathrm{SI}(t)$ has only been sketched, that the move to shared meaning is barely under way, and that the mapping of R, A, and g onto particular neural or computational processes is wide open. The claim we do press is narrower than any of that, and prior to it: to have isolated a target that the harder questions about mind, meaning, and machine intelligence routinely take for granted, without pausing to ask what would have to be true of a system before information could count as meaningful to it at all. The significance of the present work lies less in introducing another quantitative measure than in clarifying that foundational question. Between syntactic information and meaning as it is lived lies the organization that makes distinctions matter, and semantic information, as set out here, is our name for that organization and our attempt to give it a shape, one we hope can furnish a common conceptual language through which information theory, cognitive science, artificial intelligence, and consciousness studies may investigate semantic phenomena without reducing meaning either to statistical structure or to an interpretation imposed from outside. Whether $g(R, A)$ proves to be the right formalization or only a first approximation, we hope it contributes toward a more coherent foundation for treating semantic information as a scientific matter.

Acknowledgments

We thank the IKSMHA Centre, Indian Institute of Technology Mandi, and the Institute for Science and Spirituality, New Delhi, for institutional support and for discussions across both groups that shaped this work, and we thank colleagues who commented on earlier drafts. Any errors that remain are our own.

Statements and Declarations

Funding. No funding was received for this study.
Competing interests. The authors declare that they have no competing interests.
Data availability. All data generated or analyzed during this study are included in this paper.
Ethics approval. Not applicable; the study involved no human participants, animal subjects, or associated data.

Author contributions. All authors took part in conceiving and developing the framework. J.B. led the theoretical formalization and axiomatic treatment. V.H.C., L.V.P., and N.K. worked on the comparative analysis, the toy model and its operationalization, and the writing. All authors read, revised, and confirm the final manuscript.

References

Atmanspacher, H., Scheingraber, H.: A fundamental link between system theory and statistical mechanics. Foundations of Physics **17**, 939–963 (1987) https://doi.org/10.1007/BF00734321

Bar-Hillel, Y.: Language and Information: Selected Essays on Their Theory and Application. Addison-Wesley, Reading, MA (1964)

Bar-Hillel, Y., Carnap, R.: Semantic information. The British Journal for the Philosophy of Science **4**(14), 147–157 (1953) https://doi.org/10.1093/bjps/IV.14.147

Chalmers, D.J.: The Conscious Mind: In Search of a Fundamental Theory. Oxford University Press, New York (1996)

Clark, A.: Surfing Uncertainty: Prediction, Action, and the Embodied Mind. Oxford University Press, New York (2016)

Cover, T.M., Thomas, J.A.: Elements of Information Theory, 2nd edn. Wiley-Interscience, Hoboken, NJ (2006)

Damasio, A.: Self Comes to Mind: Constructing the Conscious Brain. Pantheon Books, New York (2010)

Deacon, T.W.: Incomplete Nature: How Mind Emerged from Matter. W. W. Norton & Company, New York (2012)

Deacon, T.W., Koutroufinis, S.: Complexity and dynamical depth. Information **5**(3), 404–423 (2014) https://doi.org/10.3390/info5030404

Di Paolo, E.A., Rohde, M., De Jaegher, H.: Horizons for the enactive mind: Values, social interaction, and play. In: Stewart, J., Gapenne, O., Di Paolo, E.A. (eds.) Enaction: Toward a New Paradigm for Cognitive Science, pp. 33–87. MIT Press, Cambridge, MA (2010). https://doi.org/10.7551/mitpress/9780262014601.003.0003

Dreyfus, H.L.: What Computers Can't Do: A Critique of Artificial Reason. Harper & Row, New York (1972)

Dretske, F.I.: Knowledge and the Flow of Information. MIT Press, Cambridge, MA (1981)

Floridi, L.: Information: A Very Short Introduction. Oxford University Press, Oxford

(2010)

Floridi, L.: The Philosophy of Information. Oxford University Press, Oxford (2011). https://doi.org/10.1093/acprof:oso/9780199232383.001.0001

Friston, K.: The free-energy principle: a unified brain theory? Nature Reviews Neuroscience **11**(2), 127–138 (2010) https://doi.org/10.1038/nrn2787

Friston, K.J., Wiese, W., Hobson, J.A.: Sentience and the origins of consciousness: From cartesian duality to markovian monism. Entropy **22**(5), 516 (2020) https://doi.org/10.3390/e22050516

Gibson, J.J.: The Ecological Approach to Visual Perception. Houghton Mifflin, Boston (1979)

Harnad, S.: The symbol grounding problem. Physica D: Nonlinear Phenomena **42**(1–3), 335–346 (1990) https://doi.org/10.1016/0167-2789(90)90087-6

Hoffman, D.D.: The Case Against Reality: Why Evolution Hid the Truth from Our Eyes. W. W. Norton & Company, New York (2019)

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

Kolchinsky, A.: A novel approach to the partial information decomposition. Entropy **24**(3), 403 (2022) https://doi.org/10.3390/e24030403

Kolchinsky, A., Wolpert, D.H.: Semantic information, autonomous agency and non-equilibrium statistical physics. Interface Focus **8**(6), 20180041 (2018) https://doi.org/10.1098/rsfs.2018.0041

MacKay, D.J.C.: Information Theory, Inference, and Learning Algorithms. Cambridge University Press, Cambridge (2003)

Metzinger, T.: Being No One: The Self-Model Theory of Subjectivity. MIT Press, Cambridge, MA (2003)

Millikan, R.G.: Language, Thought, and Other Biological Categories: New Foundations for Realism. MIT Press, Cambridge, MA (1984)

Merleau-Ponty, M.: Phenomenology of Perception. Routledge & Kegan Paul, London (1962). Translated by Colin Smith; originally published 1945 (French)

Maturana, H.R., Varela, F.J.: Autopoiesis and Cognition: The Realization of the Living. Boston Studies in the Philosophy of Science, vol. 42. D. Reidel Publishing Company, Dordrecht (1980). https://doi.org/10.1007/978-94-009-8947-4

Papineau, D.: Representation and explanation. Philosophy of Science **51**(4), 550–572 (1984) https://doi.org/10.1086/289205

Pattee, H.H.: How does a molecule become a message? Developmental Biology Supplement **3**, 1–16 (1969)

Piccinini, G., Scarantino, A.: Information processing, computation, and cognition. Journal of Biological Physics **37**(1), 1–38 (2011) https://doi.org/10.1007/s10867-010-9195-3

Roederer, J.G.: Information and Its Role in Nature. Springer, Berlin (2005)

Seth, A.: Being You: A New Science of Consciousness. Dutton, New York (2021)

Shannon, C.E.: A mathematical theory of communication. The Bell System Technical Journal **27**(3), 379–423 (1948) https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shannon, C.E.: A mathematical theory of communication. The Bell System Technical Journal **27**(4), 623–656 (1948) https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

Tononi, G., Boly, M., Massimini, M., Koch, C.: Integrated information theory: from consciousness to its physical substrate. Nature Reviews Neuroscience **17**(7), 450–461 (2016) https://doi.org/10.1038/nrn.2016.44

Thompson, E.: Mind in Life: Biology, Phenomenology, and the Sciences of Mind. Harvard University Press, Cambridge, MA (2007)

Tononi, G.: An information integration theory of consciousness. BMC Neuroscience **5**, 42 (2004) https://doi.org/10.1186/1471-2202-5-42

Varela, F.J., Thompson, E., Rosch, E.: The Embodied Mind: Cognitive Science and Human Experience. MIT Press, Cambridge, MA (1991)

Walker, S.I., Davies, P.C.W.: The algorithmic origins of life. Journal of The Royal Society Interface **10**(79), 20120869 (2013) https://doi.org/10.1098/rsif.2012.0869

Wiener, N.: Cybernetics: Or Control and Communication in the Animal and the Machine. MIT Press, Cambridge, MA (1948)

Zahavi, D.: Phenomenology: The Basics. Routledge, London (2019)