CHAPTER 13

Consciousness as Differentiation

Presupposes: the structural dependence of content on contrast (Ch. 2); the zero-complexity constraint (Ch. 6).

Establishes: the categorical transition principle, and the hierarchy derived from it level by level; the two axes of the mental, and the directional difference as the genus of directive response; the reinterpretation of four rival theories of consciousness, two of them in the appendix; and, in the appendix, the uniqueness principle, which no later result uses. All four results are graded constructed; note 1 records the grading structure.[1]

In brief. A system can discriminate a million shades of colour and be no more advanced, in the relevant sense, than one discriminating two. What raises a mind's level is not more distinctions but a categorically new *kind* of distinction — and this chapter derives the ladder of kinds, from thermostat to moral agent, from a single principle: content requires contrast. Along the way it separates the two axes every mental state divides into — *content* (what is detected) and *direction* (the felt pull toward or away) — and locates the boundary of experience at availability to thought, and the boundary of *mattering* at valence. It also names the question every theory of consciousness must answer before the hard problem arises — which physical events constitute experience, and what distinguishes them from the billions that do not? — and answers it in an appendix, with a proposal the rest of the book does not use. The chapter does not solve the hard problem as it is posed; it offers in its place an identity — feltness as availability to thought — that a critic can doubt, explains why the doubt arises, and constrains any solution.

---

1 Introduction

What makes a thermostat different from a thinking mind? Not the presence of causal sensitivity: both respond to their environments. Not the number of distinctions registered: a thermostat with a billion sensors would still not think. What makes the difference is the *architecture* of the distinctions detected: the kinds of contrast a system is sensitive to, and how those kinds are related to one another.

Chapter 6 established the central negative result: a state of zero complexity, with no internal distinction of any kind, excludes cognition, representation, and experience. This chapter develops the positive counterpart, extending the foundational commitment that content requires contrast into a theory of consciousness.

It asks what structural conditions a system must satisfy to possess conscious content, what explains the qualitative differences between levels of mental complexity, and the question every competing theory must answer: what distinguishes the physical events that constitute consciousness from those that do not? This is the boundary question, which is prior to the hard problem.

The chapter states conditions that any theory of consciousness must respect. It does not solve the hard problem as Chalmers (1996) poses it, why there is experience at all. In its place it offers an identity, feltness as availability to thought, as a claim a critic can doubt, with an account of why the identity seems false to the mind that holds it; and it constrains any rival solution. It keeps its own answer to the boundary question apart from what it derives, in an appendix on which no later result depends.

The chapter's positive result is a ladder: the levels of mind, from the thermostat to the moral agent, each added by a new kind of distinction a system can detect. Section 2 derives it, saying what makes each step a new kind, and with it the two axes every mental state divides into, content and direction. Section 3 places the rival theories against the ladder, §4 states what would refute it, and §5 meets the objections. Section 6 closes with what the rest of the book takes from the chapter, which is less than the chapter. The appendix (§§A.1–A.8) holds the chapter's own proposal for where experience is located and why, with the rival theories and the test that bear on that proposal. The ladder itself comes first, in miniature, a line and an illustration to each level.

Level	**What is newly told apart**	**An illustration**
0 Registration	A difference is reacted to, and remains unfelt	The thermostat's click; the reflex arc's withdrawal
1 Sentience	A registered difference becomes available to thought, and so is felt	Warmth in a hand, felt whether or not it is thought about

Level	What is newly told apart	An illustration
2 Simple consciousness	What is present to the senses is told apart from what the mind generates	Seeing red against imagining red
3 Embodied feeling	The body that feels is told apart from bodies merely perceived	The hand that hurts when pinched, against the hands only seen
4 Self-consciousness	The feeling body and the thinking mind taken as one thing, against all else	Pain in this hand and worry about it belong to one self, before either is reflected on
4+ Somatic integration	The one self extended to its parts that do not feel	Hair and bone counted as one's own, though unfelt
5 Other-consciousness	Another body is detected as a subject, where before it was an object	Another's pain registered as pain, which is empathy; the same capacity run in reverse is cruelty
6 Reflection	The system's own states become objects it can detect	Noticing that one is angry
7 Moral consciousness	One's own conduct is evaluated by what it does to other subjects	"What I did hurt her, and that counts against it"

2 The Categorical Transition Principle

2.1 Detection and availability

A thermostat responds differentially to temperature, yet nothing about the response is felt. A nervous system responds differentially to temperature; at some point the response becomes an experience of

warmth or cold. The difference is not the number of states discriminated, nor the fineness of discrimination: a thermostat can be made arbitrarily fine-grained without becoming a subject. The difference is that in the one case a distinction is merely registered, while in the other it is registered *for the system*: undergone, where the thermostat's response merely occurs.

What "registered for the system" comes to can be stated in full. The distinction is available to thought — the kind of content the percept/thought system can take up, that system being characterised structurally, without phenomenal vocabulary, at §5.2 — whether or not any thought is in fact directed upon it. Availability is a standing property of the state; nothing has to fire for it to hold. A felt warmth is available to be thought about, and is felt whether or not it ever is.

The thermostat's registration has no such availability: it is sealed in the reflex arc, graspable by no thought even in principle. Blindsight is the instructive intermediate (Weiskrantz, 1986). The information is registered and drives behaviour, yet it is unavailable to the patient's thought, and for that reason unfelt. What divides the registered-for-the-system from the merely registered is availability in principle; the exercise of thought is not required.

The foundational commitment fixes where the line falls, because content requires contrast. A bare causal transition carries no content; it is a change, not a difference for anything. Registration becomes experience where the registered contrast is available to thought. Within experience, a registered contrast may further acquire a direction, the system drawn toward one pole and away from the other.

This directional pull is valence in its rawest form. Present before any actual thought and attached to no particular thought, it is the most primitive ancestor of emotion and the boundary of mattering. Where an available contrast is felt as a pull, the subject has a stake in it. Where available but uncoupled, it registers without caring. Where not available at all, there is only a mechanism.

2.2 The principle stated

Quantitative enrichment is not advance. A system can discriminate a million shades of colour and be no more advanced, in the relevant sense, than one discriminating two; it has accumulated more distinctions within a single category and crossed into no new one. A qualitative advance is the detection of a categorically new kind of distinction — new when its obtaining presupposes the items the prior categories make available while itself falling outside every prior category.

The principle. This is the categorical transition principle: increases in mental complexity are constituted by the detection of categorically new kinds of distinction, while quantitative increase within an existing category leaves the level unchanged. A distinction can be categorically new in one of three ways, the third a limiting case; the hierarchy uses all three.

The first is a new dimension: a contrast on an axis that did not previously exist. The percept/thought distinction is categorically new relative to red-versus-blue, because it contrasts two kinds of mental state, where red-versus-blue contrasts two instances within one kind. The second is a new composition: existing relata integrated into a unit that affords a contrast none of its components affords. The self/non-self distinction is not the conjunction “has sensation and has thought” but the recognition that the sensing thing and the thinking thing are one entity, which can then be set against everything it is not. The unity is a relatum that did not exist before the integration.

A limiting third form is a new order, in which the system takes its own states as objects, so detection itself becomes something detected. Marking the form of each transition is what makes the hierarchy a derivation, where a list would leave each step unmarked: at every step it can be said which kind of advance has occurred, and why it qualifies as categorical.

2.3 Two axes: content and direction

The content axis. The progression tracks two things that must be held apart, since conflating them is the source of several persistent

confusions. The first is content: the kinds of distinction a system can detect or compose, which is what the numbered levels enumerate.

The direction axis. The second is direction: the pull toward or away, introduced at the first conscious level. Direction is not a level. It is valence, the felt toward-or-away, switched on with sentience and coupling to whatever content the higher levels make available. A thermostat and a sweating body both register heat unfelt, without valence, whatever electrical activity they carry. The levels build the content; valence runs through it.

What the separation fixes. This separates two categories of feeling that ordinary usage runs together. A sensation is localised — the warmth in a hand — and that localisation is part of its work, since it is what lets felt sensation demarcate the body. An emotion is not localised but a response of the organism as a whole, carrying a direction. Sensation answers what and where; emotion answers better or worse for the organism as a whole. Pain displays the composition coming apart: its nociceptive component is localised, so one can say where it hurts, while the suffering that gives it its negativity is not. The two dissociate, the sensation surviving intact when the suffering is removed.

The two axes are exhaustive; the exhaustiveness divides into two tiers whose grades differ. That every feature of a mental state is a determination or nothing is derived, and carries the derivation's strength (Chapter 6, §2.2). That a determination either carries a felt orientation or does not, and that orientation is the only further division a determination admits, is this chapter's own claim, carried at its constructed tier; the three cells — determination, orientation of one, or nothing — are what the derived partition yields once that claim is granted. Which cell a contested feature occupies, and what its occupancy consists in, can require substantive identification: the placement of feltness, below, is the availability thesis's work, and carries its register. The result, the two tiers together, is one later chapters lean on and Chapter 14 (§3.5) re-derives independently for its metaethical use.

The exhaustiveness argument. The derivation runs from a single question. Consider any feature of a mental state whatever. Either the

feature is a way the state is, a determination, one way rather than another, or it is not. If it is, it is content in the structural sense of Chapter 2: a diff, individuated by what it excludes, belonging to the first axis. If it is not — if the feature is no way the state is — then it is not a feature at all, by the dilemma of Chapter 6 (§2.2).

What remains to ask of any content is a single binary question: is it carried by valence or not, felt as a pull toward or away, or merely registered? Coupling is binary, and valence is not a further content but the orientation of one. There is therefore nothing for a third axis to be. A proposed third feature of the mental must be a determination (the first axis), the orientation of a determination (the second), or no way the state is at all, which is nothing.

One inheritance must be refused. The exhaustiveness does not descend from the partition of Chapter 3; the temptation to say it does should be resisted all the same, since the two divisions do not correspond. Chapter 3 divides on distinction or its absence, which yields determination or nothing — two cells, with no content/direction split among them. The trichotomy here is a sub-division of the first cell: a determination either carries valence or does not, and "the orientation of a determination" is a determination under a description, not a third thing beside determinations.

What the partition supplies is the exclusion of the third cell, since a feature that is no way the state is at all is nothing. What supplies the rest is the substantive claim that coupling is binary and that valence is the only sub-division a feature of the mental can take, which is carried at this chapter's tier. Chapter 14 (§3.5) does not rest on it. Its trilemma is re-derived there on a division of its own — transferable or not — so a reader who declines the two axes may still grant it.

One objection tests the partition where it is weakest, and meeting it fixes the sense in which the axes are complete. The objection points to subjectivity itself — that an experience is undergone, that there is something it is like, that the state is someone's — and claims this is a third feature, neither a content nor the orientation of one. The reply

separates two things the objection runs together: that experience is felt at all, and that it is felt as belonging to a self.

Take feltness first. The partition by itself forces only that a residue which is *no* way the state is at all is nothing (Chapter 6, §2.2). What it does not by itself deliver is the identification the account needs here: that feltness is *the availability of a content to thought* (§2.1). A felt state is content of the kind the percept/thought system can take up, whether or not it is taken up: availability in kind, an episode of access being a further matter. The unattended periphery of the overflow cases (Chapter 7, §2.3) is content the system can take up though it does not, which is what makes overflow a datum for the identification instead of a counterexample to it. Call this the **availability thesis**: a substantive identification, not a corollary of the partition, and carried in the Register as constructed. Granted, feltness is a property of the content and belongs to the first axis — a placement the partition guarantees is available, since no third cell exists for feltness to flee to, and which the thesis, not the partition, performs.

The residue and its two tests. The realist who resists presses a sharper residue: a determinate feature of available content, distinct from its availability — available content that could, she says, have been present unfelt. Against that residue the availability thesis has its reply, which should be weighed for what it is: a costing of the residue.

Two considerations support it. First, the residue must, by Chapter 2, carry a contrast. The realist owes a specification of what an available-but-unfelt content of the very same type would be, a single content held fixed with its feltness subtracted, a zombie drawn at the scale of a single state; no case exhibits one. Second, the equalisation evidence (Chapter 6) shows feltness and availability co-eliminating: strip a modality's contrasts and its felt field vanishes altogether, never thinning toward an unfelt-but-available residue. Identity predicts this; a further-determination view must record it as coincidence.

Neither consideration defeats the intuition that the residue is there; the chapter does not claim they do. What the projection mechanism of Chapter 8 explains is why a mind built from distinctions finds the

residue inevitable, which is an account of the intuition, short of a proof against it. One further instrument is available to a reader who accepts it elsewhere. Chapter 14 (§3.6) tests a contested clause by subtraction: remove it and ask what contrast remains. Remove "felt, over and above available" and nothing changes in any register, the subject's own reports included. By that test the residue is vacant, as the realist's clause is vacant there. A reader who refuses the test refuses it in both places at once, so the Register counts the refusal as one exit and not two.

The intuition that a fact is nonetheless left over is itself explained, on the projection pattern of Chapter 8 (§3.4). A mind whose exercise of its feltness-concept is itself a felt taking-up must find the identification incredible, since feltness is presented from the inside in a mode no structural description shares. The incredulity is therefore what the thesis predicts the realist would feel whether or not any further fact existed, so it cannot discriminate his view from the identity. What the literature calls the phenomenal concept strategy (Loar, 1990; Papineau, 2002) is here a consequence of the framework's own account of representation, with no auxiliary hypothesis needed.

The second half of subjectivity — that the experience is *mine* — is a content, and a specifiable one. Ownership is a composition: the system represents a body, represents itself, and represents the feeling-body as the seat of that self (Level 4, below); mineness is that composed representation, on the first axis by construction. This is why felt experience at the floor is undergone yet not owned: availability makes it felt before any self-model is present to make it personal. A sentient creature is already a subject (there is something it is like to be it) while representing no self for its experience to belong to.

What the axes settle. Neither half of subjectivity is a third axis. Feltness is availability of content; ownership is complex content; valence, where present, is the second axis, supplying mattering and not feltness. The realist's challenge is now definite: exhibit a single available content whose feltness can be subtracted while the content is held fixed, or a mineness that decomposed into no self/non-self contrast. Either would falsify the architecture here.

The same constraint binds Chapter 14's target in advance. A philosopher who posits a sui generis mode of the mental — a rational grasp that is neither the registration of structure nor the feeling of pull — must characterise it; every characterisation will attribute determinate character (the first axis), oriented character (the second), or no character, which is nothing.

The quarantine extends to the whole chain by a fork the realist is owed. Grant the phenomenal residue arguendo, a feltness over and above availability. Either it acts — affecting report, behaviour, or orientation — in which case it is a determination absorbed into content or direction. Or it idles, a pure spectator touching nothing, changing no downstream result, since value, morality, purpose, and persuasion all run on registered pulls. Every result from Chapter 14 onward holds with or without the glow, so the residue dispute, wherever it settles, is a local dispute about this chapter's positive theory and never a lien on the chain.

Because valence is separable from content, any content can be taken up in either of two modes. Coupled to valence, it is evaluated: felt as mattering, oriented as better or worse. Uncoupled, it is felt but neutral where available, and merely registered where not — computed, or checked against a standard, without mattering — orthotic processing, in the term Chapter 14 (§4.3) fixes. Call a content of the first kind, a difference carried by valence, a directional difference.

The vocabulary can now be fixed. Valence is the bare pull. An emotion is valence bound to content, a directional response about something, of which human moral emotion is the high-content case. The felt bindingness of a rule is a further, misattributed instance (Chapter 16). This single fork, recurring at every level, is the structural root of the distinction between evaluative and cognitive operation developed there: what appears in that chapter as a feature of moral judgement is a general feature of mind, present wherever content can be processed with or without valence engaged.

2.4 Structural containment

The levels are ordered by structural containment: a higher level presupposes and incorporates the categories the lower levels supply. A self-model (Level 4) cannot be assembled without the feeling-body and mind categories it integrates (Level 3). The recognition of others as subjects (Level 5) presupposes the self whose structure it extends. Damage to a lower capacity degrades the higher capacities that depend on it, never the reverse.

This is a substantive structural claim, since it carries a falsifiable prediction, which no stipulation does. The levels are separable in one direction only: a system may possess a lower level without the higher, but not a higher while wholly lacking the lower one it composes. A system shown to possess an integrated self-model while wholly lacking the percept/thought distinction would falsify the ordering. Containment is an ordering of structural dependence, not of temporal sequence, so the levels need not be acquired in order; a mature system runs many concurrently. The claim is only that the higher cannot obtain without the lower.

2.5 The hierarchy

With the principle and the two axes in place, the hierarchy can be set out. Each level names the categorically new distinction it adds, the form of transition it represents, and the coupling of valence it makes available. The table compresses the derivation; the notes after it record only what later chapters use.

Level	Content axis: the new distinction	Form of transition	Direction axis: valence
0 Registration	Differential causal response unavailable to thought (thermostat, reflex arc)	— (pre-conscious floor)	none; unfelt

Level	Content axis: the new distinction	Form of transition	Direction axis: valence
1 Sentience	Registration become available to thought, and therefore felt	availability crosses mechanism → experience	valence may couple: neutral if not, directional if so
2 Simple consciousness	Percept vs thought — the present-to-sense vs the internally generated	new dimension	couples to thought: one representation = instinct; linked representations under one direction = reasoning; uncoupled = orthotic computation
3 Embodied feeling	Feeling-body vs merely-seen body; thought recognised as a category	new dimensions	localised sensation marks the body as the system's own; valence merely couples more or less strongly
4 Self-consciousness	The feeling body and the thinking mind are one entity, set against the non-self	new composition	valence becomes personal — the pull is owned
4+ Somatic integration (parallel branch)	Connected-non-feeling (bone, hair) vs unconnected-non-feeling	new composition, branching	— (no later level depends on it)

Level	Content axis: the new distinction	Form of transition	Direction axis: valence
5 Other-consciousness	Other bodies are subjects	new composition, by extension	valence takes others' feelings as object = empathy; reversed = cruelty (one capacity, two directions); uncoupled = cold theory of mind
6 Reflection	The system's own states become objects	new order	—
7 Moral consciousness	Empathic evaluation (5) brought to bear on the system's own conduct, taken as object (6)	new composition	concordant direction = moral goodness; reverse = cruelty toward an end

Three notes carry forward.

At Level 1, availability is the on-switch of experience and valence is the on-switch of mattering. One dependency in the ordering should be recorded here. Availability is availability in kind to the percept/thought system; that system's contrast with the percept is the dimension Level 2 introduces, so the first level is characterised by reference to a format whose full articulation arrives one level up. The ordering is one of containment; a format can be characterised before the contrast that displays it is detected (§5.2); the relation is a dependency to be read in that direction, not a circle. Nor is a Level 1 creature felt by courtesy of a consumer it lacks. The consumer is present at Level 1: an internally generated system that takes registrations up is what makes them available in kind, and the three marks of §5.2 fix it without any mention of feltness. What Level 1 lacks, and Level 2 adds, is the detection of the

contrast between what that system takes up and what it generates. Availability at Level 1 is availability to a system the creature has, whose difference from the percept it does not yet register.

A sensation available to thought but carrying no pull is felt yet neutral — a grey field, a bare touch — so the account affirms neutral felt experience as the ordinary case and owes no strained treatment of it. That valence, and not fineness of discrimination, is what turns a felt sensation into an emotion is fixed here.

At Level 5, empathy and cruelty are one capacity run in opposite directions, with no second mechanism, which is the result Chapter 16 (§6.3) redeploys against the persecutor. The cold-theory-of-mind case is Level 5 content processed without valence, which the analysis of psychopathy requires (Chapter 16, §5.2).

At Level 7, that valence constitutes evaluative difference, and does not register values obtaining independently of it, is argued in Chapter 14; the present chapter places the capacity, borrows the name “moral” from the identification Chapter 16 argues, and leaves its metaphysical analysis to that argument. The progression from 0 to 7 is generated by a single principle, content requires contrast, applied through the two forms of categorical transition and threaded by a single valence. Each step is a determinate advance over the one below, its place fixed by what it presupposes. A being that detected every category the principle permits would instantiate every level at once.

The derivation the table compresses can be run in full, one level at a time, by the test §2.2 states: does the level’s distinction presuppose the items the prior categories make available, while itself falling outside every prior category? At each level the test returns the same verdict, so the form of each transition can be read off. The levels themselves follow.

Level	The new contrast	What it presupposes, and why no lower level draws it	Form
2 Simple consciousness	The present-to-sense against the internally generated	Felt registration, since both relata are felt; Level 1 contrasts registrations within one kind, never two kinds of mental state	New dimension
3 Embodied feeling	The body that feels against bodies merely seen; thought marked as a category	The percept/thought dimension: only a system that separates the two can locate a sensation in one perceived body among others, and recognise the internally generated as a kind	Two new dimensions
4 Self-consciousness	The feeling body and the thinking mind as one entity, set against everything it is not	Both relata; the unity is a relatum that did not exist before the integration, and the self/non-self contrast is available to nothing below	New composition
4+ Somatic integration	The unity extended to parts that do not feel — bone, hair — against unconnected non-feeling things	The self-model; no later level depends on it, which is why it branches	New composition, branching

Level	The new contrast	What it presupposes, and why no lower level draws it	Form
5 Other-consciousness	Another body detected as a subject	A self to extend; a subject other than oneself is a relatum no prior level affords	New composition, by extension
6 Reflection	The system's own states as objects	A self whose states they are; the contrast between a state and that state as an object is of a different order from every contrast below	New order
7 Moral consciousness	Empathic evaluation brought to bear on chosen conduct	Levels 5 and 6; the object — the system's own conduct as it bears on other subjects — is a composite relatum neither affords alone	New composition

Two natural candidates test the count and not the ordering; running the criterion on them shows why the hierarchy derives its levels instead of listing them. Episodic memory looks level-like, since a mind that re-lives its own past seems transformed in kind. Apply the test: does its obtaining presuppose the prior categories while itself falling outside every one of them? Episodic recollection re-presents contents of kinds already on the table — percepts, feelings, the self who had them — ordered in time, and temporal ordering is detected wherever change is detected, from the lowest levels up. Its distinctive signature, ownership of a past experience, is the Level 4 self-model applied along a dimension it already carries, which makes episodic memory a

spectacular quantitative enrichment on the wrong side of the quantitative/categorical line.

Language invites the same verdict by a different route. Its representational content rides on distinctions the prior levels supply; its new contribution, the public labelling act, is what Chapter 7 (§3.1) files as the third grid's work — a loop within the grids, adding no new kind of detected distinction. What language transforms is transmission and combination; what a level requires is a new category of detectable difference. Both candidates fail the criterion for stateable reasons; the stateability is the point: a critic who thinks a level is missing knows what she must exhibit, a kind of distinction that presupposes the existing furniture and falls outside every category on the table.

A third candidate is agency, the chosen against the merely occurring; the criterion turns it away for a reason that bears on the whole hierarchy. Every level is a level of registration: what the system detects. Acting is not a further thing detected. It is what the system does on what it detects; it is present from the floor. The thermostat's click and the reflex arc's withdrawal are output conditioned on registration unavailable to thought; the shout of fear is output conditioned on a felt registration; a choice is output conditioned on alternatives the field has registered, which the percept/thought distinction of Level 2 first supplies, an internally generated course set beside the perceived one; deliberation is output conditioned on alternatives and on the system's own states taken as objects, which is Level 6. What changes across the hierarchy is what the output is conditioned on; the complexity of the act scales with the complexity of the field that feeds it, so "choice" names the act at the first level where alternatives are in the field and adds no category of detected difference of its own. The rule that turned away episodic memory turns away agency: the chosen/merely-occurring contrast is composed from the alternatives Level 2 first supplies and the reflection of Level 6; a system with both has the contrast without a further level. The criterion is registration, not anatomy — a spinal reflex is central nervous tissue; what places it at the floor is that nothing in it is available to thought.

Level	Output conditioned on	The act, named
0	differential causal response unavailable to thought	The reflex, on an unfelt registration
1	registration available to thought, and therefore felt	Felt reaction: the shout of fear, on the felt registration alone
2	the present-to-sense distinguished from the internally generated: the first alternatives in the field	Instinct, where valence couples to one representation; choice, where representations are linked and the act runs through the alternatives
3	the body registered as the system's own	A choice about the body registered as one's own
4	the feeling body and the thinking mind as one entity, set against the non-self	A choice for oneself, the options held as a self's own
5	other bodies registered as subjects, so that their perspectives are among the alternatives	A choice conditioned on another's perspective: the threshold of sufficiently complex choice (Chapters 14 and 16)
6	the system's own states taken as objects	Deliberation: the alternatives, and one's own states taken as objects
7	the system's own conduct as it bears on other subjects	Moral choice: one's own conduct, evaluated by what it does to other subjects

The output axis is also the axis of choice; instinct is a route on it and no rung. The instinctive act is the one conditioned on the felt registration alone, bypassing whatever alternatives the field contains; the route stays open at every level from the first: a dog bites on instinct though its field has alternatives, a person flinches though hers has more. Choice is the act that runs through the alternatives; its complexity scales with the field that feeds it. What Chapters 14 and 16 call a sufficiently complex choice has its address here: an act that runs through the alternatives and not by the instinctive route, in a field in which the affected party is registered as a subject (Level 5), with reflection (Level 6) supplying the agent's own conduct as an object for the evaluation Level 7 performs.

One thing the output axis does carry, and Chapter 18 draws on it. What a system does is never among the things it registers: the producing of an act is not a content of the field that produced it, at any level. Where the act is choice, this is the asymmetry a chooser meets from inside.

Where the argument stands: the categorical transition principle and its hierarchy are derived (§2), on the two axes and the availability thesis; they are what the chapters after this one draw. The boundary question and the chapter's own answer to it — the uniqueness principle, the interface, and the equalisation constraint — are set out in the appendix (§§A.1–A.8), where the chain does not use them. The rival theories can now be placed against the hierarchy.

3 Engagement with the Literature

The rival theories of consciousness stand in determinate relations to the hierarchy: two are complemented, the nearest neighbour of the availability thesis is placed, illusionism is met; the affective-neuroscience convergence corroborates the valence claim. The interface account's own rivals, integrated information theory and predictive processing, are met in the appendix (§A.6).

Higher-Order Thought theory (Rosenthal, 2005) holds that a state is conscious when it is the object of a suitable higher-order thought. The

account reinterprets HOT as a theory of one categorical transition, the transition to reflective self-consciousness (Level 6). It captures that transition correctly while misidentifying its scope, since Level 1 registration involves phenomenal character without higher-order awareness; the hierarchy supplies the levels below and above the HOT transition that the theory lacks.

Global Workspace theory (Dehaene & Changeux, 2011) holds that consciousness arises when information is broadcast globally, becoming available for report and control. The categorical transition principle complements it by supplying what is accessed. GWT identifies the mechanism of access but not what makes the broadcast content conscious as well as accessible: the content is conscious when it involves the detection of categorically distinct contrast, which the broadcast then makes available for integration.

The nearest neighbour of the availability thesis in this literature is Prinz's (2012) account, on which a representation is conscious when attention makes an intermediate-level perceptual representation available to working memory. The two agree that a consumer fixes the boundary of experience and differ on the consumer — attention to the intermediate level there, availability in kind to the percept/thought system here — and the difference is the one §A.7 stakes on the phosphene case. Tye's (1995) poisedness condition is the older ancestor: a content is phenomenal when poised for use by the cognitive system, whether or not used. The availability thesis keeps the poise and drops the abstractness and nonconceptuality Tye required of the content, since the format the percept/thought system consumes is fixed structurally at §5.2 and by no further mark.

Illusionism (Frankish, 2016) holds that phenomenal properties do not exist and that a theory of consciousness owes only an account of why introspection represents them; the meta-problem programme (Chalmers, 2018) makes the second demand general, asking every theory to explain the intuition of a further fact. The availability thesis is the realist neighbour of the first and an instance of the second. It keeps feltness, identified with availability in kind, and denies only the residue; the projection mechanism of Chapter 8 is its answer to the meta-

problem, explaining why a mind built from distinctions finds the residue inevitable (§2.3). Illusionism and the thesis share the debunking and part on what survives it, the illusionist retaining the seeming and the thesis the availability; what tells them apart is whether the equalisation evidence, which co-eliminates feltness and contrast, is read as a fact about experience or about reports of it.

Affective neuroscience corroborates the chapter's most distinctive structural claim, valence, from a different quarter. The account holds that what converts registration into the lowest grade of experience is the switching-on of valence at sentience (§2.3). Panksepp (1998) reaches a structurally parallel claim from lesion and stimulation evidence: the primordial form of consciousness is affective, phylogenetically prior to the cortical machinery of perception and thought, with cognition layered above it. That is the ordering the two-axes architecture predicts, since valence is the first thing that makes a registered difference a difference for the system.

The convergence is undesigned. Panksepp proceeds from the neuroscience, the present account from the structural dependence of content on contrast; the two reach the same ordering. One divergence remains: Panksepp treats affect as a natural kind delivered by specific neural systems, while the account treats valence as a structural role those systems realise in the mammalian case without exhausting, since the exhaustiveness argument of §2.3 makes valence a feature any possible sentience must instantiate, whatever its substrate.

4 Falsification Conditions

The categorical transition principle generates predictions that could falsify it. Level 4 (self/non-self) without Level 2 (percept/thought) would fell the containment ordering. A qualitative increase in complexity through quantitative accumulation within one category, with no categorical transition, would fell the principle. A Level 0 system possessing any consciousness would fell the zero-complexity constraint. The availability thesis adds a condition of its own, stated at §5.2: a content shown to be phenomenal while unavailable in kind, of a format

the percept/thought system could not take up even in principle, would refute the identification. The interface account adds conditions of its own, stated with the stake they carry in the appendix (§A.7).

5 Objections and Replies

5.1 The hard problem objection

Objection. The chapter identifies structural necessary conditions without explaining why physical processes give rise to subjective experience at all.

Reply. Conceded as the objection poses it: the chapter offers no account of why there is experience at all. What it offers in place of one is an identity — feltness as availability to thought (§§2.1, 5.2) — which a critic can doubt, the residue of that doubt being the hard problem in the form it keeps here. Beyond the identity, the claim is that any solution must be compatible with the structural conditions identified here, and that the equalisation constraint (§A.5) narrows the space of viable solutions by ruling out theories on which experience and contrast separate. The residual sense of a gap, the explanatory gap of Levine (1983), is itself addressed at §5.2: it may be the representational limitation of Chapter 8, since a mind constituted by distinctions cannot picture the identity it theoretically endorses.

5.2 The access/phenomenal objection

Objection. Block (1995) distinguishes access consciousness, a state's availability for reasoning, report, and control, from phenomenal consciousness, there being something it is like to be in the state, and argues the two are dissociable. The availability thesis identifies feltness with availability to thought, so it appears to define the phenomenal into a species of access — the very conflation Block's distinction exposes. A circularity compounds the charge: "availability to thought" explains consciousness by appeal to thought, itself a conscious kind.

Reply. The identification is drawn at a different joint from the one Block targets; the difference is the one §2.3 marked: availability in kind

versus episode of access. Block's access is episodic and functional, so a state is access-conscious when its content is in fact poised for use. The availability thesis concerns a standing property: content of the type the percept/thought system can take up, whether or not it is taken up.

The dissociation evidence then falls into place on both sides of the thesis. Overflow — the unattended periphery, the uncued rows of the partial-report grid (Sperling, 1960; Block, 2007) — is phenomenality without episodic access, which Block adduces against identifying the phenomenal with access. On the availability-in-kind reading it is a datum for the thesis: the periphery is felt because it is available in kind; its going unretrieved does not make it unfelt. Blindsight is the converse, registration from which availability in kind has been removed, and with it the feltness. What Block's cases show is that feltness is not episodic access, which the thesis never claimed. A critic holding that they show feltness floats free of availability in kind as well owes the state §2.3 demanded: a single content of the very same type, available in kind, with its feltness subtracted. No dissociation exhibits it.

Availability is categorical, not dispositional. A dispositionalist misreading must be blocked. Availability in kind is not a bare disposition to be accessed but a categorical property of the state: its content is of the representational format the percept/thought system consumes, individuated by the contrastive structure of Chapter 2 and delivered at a terminus (§A.3). The state possesses that format whether or not the consuming capacity is free. Overflow dissociates the two grounds cleanly, since the working-memory bottleneck limits the consumer while the periphery's contents keep their format throughout; their feltness tracks the format and survives the bottleneck.

The circularity charge mistakes the register in which "thought" enters. What the thesis invokes is a structural role — the internally generated, inferentially linked representational system whose contrast with the percept constitutes Level 2 — and the role can be specified, and need not be gestured at, in three marks, none of which mentions feltness.

Origin: the system's states arise internally, from no sensory terminus, an architectural fact statable in the pathway vocabulary of §A.3, where origins and termini are already distinguished without phenomenal remainder. Linkage: its states compose inferentially, each content constraining and constrained by others in chains, a relational property of contents individuated contrastively (Chapter 2). Format: its states consume and produce a common representational format, individuated the same way.

Each mark is attributable from the third person, to a system about which the phenomenal question is still open. The analysis is therefore recursive but grounded, as a recursive definition is grounded by its base clause. Felt states are those available to a system some of whose states are themselves felt, while the system is fixed by the three marks and not by its feltness, so the definition can be entered from outside.

The discharge exposes where the thesis's substance lies and conceals none of it. With both relata structurally specified, the identification between availability so defined and feltness is a substantive claim a critic can coherently doubt, the residue of that doubt being the hard problem (§5.1); the Register's "constructed" records this (§2.3). What the discharge removes is the charge of circularity. What it leaves is the claim, standing where the Register always placed it.

Chalmers's (2007) dilemma for phenomenal-concept strategies is taken at its first horn. The dilemma runs: either the concepts are structurally explicable, and cannot explain our epistemic situation, or they are not, in which case the gap reopens. The projection mechanism of Chapter 8 is a structural account of representation; what it explains is not the truth of the identity but the inevitability of the intuition of its falsity, which is the epistemic situation to be accounted for. What is conceded is what §5.1 conceded: the demand for a further fact lapses only if the availability thesis holds, which the Register carries as constructed, its support and its limit stated where they are given (§2.3).

The empirical wing of Block's case has since been sharpened by the no-report paradigms (Tsuchiya, Wilke, Frässle, & Lamme, 2015); the refinement should be registered, because it bears on which theories the

dissociation evidence actually pressures. In these designs the contents of experience are tracked without asking the subject to report them — percepts in binocular rivalry read from reflexive eye movements, for example — and the neural correlates so isolated shrink, since much frontal activity previously credited to consciousness tracks the act of reporting instead.

The result presses hardest on accounts that tie feltness to access episodes, global broadcast for report included; the availability thesis stands on the other side of the cut it makes. Availability in kind is poise, never performance. The rivalry percept in a no-report design retains the format the percept/thought system consumes, since the subject can report it the moment she is asked; the reflex the decoder reads is downstream of the registered content; what the paradigm subtracts is the access episode the thesis never required. The paradigms also supply the falsification shape §2.3's challenge calls for: a content demonstrated phenomenal while unavailable in kind, of a format the percept/thought system could not take up even in principle, would refute the identification — and the no-report literature, for all its subtractions, exhibits none.

6 Conclusion

The chapter yields four results, all constructed, over an established constraint.

The categorical transition principle. Increases in mental complexity are constituted by the detection of categorically new kinds of distinction, ordered by structural containment; the hierarchy, Levels 0–7 with 4+ as a parallel somatic branch, provides the architecture. What generates the qualitative difference between a thermostat and a self-conscious being is a different architecture of the distinctions detected, not a different kind of relation.

The two-axes architecture and the availability thesis on which it turns are what Chapters 14–16 inherit — the two axes and the existence of valence, and nothing else from this chapter's positive theory. The inventory can be given by name, so the quarantine is checkable at a

glance. The trilemma of Chapter 14 (§3.5) draws the derived tier of the two-axes result — determination or nothing — together with a division of its own, transferable or not, so that it uses no result from this chapter's constructed tier. The anchoring of Chapter 14 (§6.4) draws the existence of valence as the orientation of a content. The grip argument of Chapter 16 (§3.1) draws the derived tier again, while the chapters beyond draw the modality's existence through those results. No downstream result draws the availability thesis in full, so a reader who rejects the placement of feltness loses this chapter's positive theory and nothing that later chapters use.

The uniqueness principle. What makes certain physical events constitute experience is their structural uniqueness at an interface with the modality's absence, identified through three ordered features and constrained by equalisation (Appendix, §§A.1–A.5). This rules out theories on which experience is intrinsic to particles, persists without contrast, or diminishes gradually. The panpsychist's proto-consciousness cannot answer the boundary question, because a property that marks no boundary can constitute no content.

The reinterpretation of rival theories. The hierarchy supplies the structure IIT lacks, reinterprets HOT as one categorical transition, complements GWT with the architecture of what is broadcast, contests predictive processing on a discriminating prediction (§A.6), and is corroborated on valence by affective neuroscience.

If consciousness is constitutively dependent on detected difference, the question becomes: how do conscious beings evaluate the differences they detect? Not all differences are alike — red-versus-blue is a chromatic difference; better-versus-worse is an evaluative one. What modality constitutes that last category? And what follows if it is not reason but emotion? Chapter 14 develops the answer.

Grades and Dependencies

1. **Grades.** All four results are *constructed* — positive proposals with live rivals — over the *established* zero-complexity constraint (Chapter 6). Two

gradings are internally structured and the Register keeps them apart: the two-axes exhaustiveness has a derived tier (that a feature is a determination or nothing) and a constructed tier (that a determination's only further division is by orientation, together with the availability thesis's placement of feltness), and the uniqueness principle separates its established boundary constraint (§A.2) from its constructed criteria (§A.3). The downstream-use inventory lives in §6 by design, where a critic can check the quarantine at a glance; the availability thesis is used by no later result; the Register marks it so. The neural vocabulary of §§A.2–A.4 is read realistically for convenience and carries no commitment the framework lacks: "terminus" and "interface" name stable patterns of registered distinction; every claim made in them survives translation into that idiom, the phosphene prediction included, which becomes the prediction that a registered signature meeting a registered absence yields a registration of the modality's kind (Chapter 1, §5.3).

Appendix: The Boundary of Experience

This appendix holds the chapter's own proposal for the boundary question, which no later result uses: the uniqueness principle and the interface at which it locates experience (§§A.1–A.4), the equalisation constraint any theory of consciousness must respect (§A.5), the two rival theories whose target is the proposal itself (§A.6), the predictions the proposal is staked on (§A.7), and the one objection aimed at it (§A.8). Its results are carried in the Register at their own tier; rejecting them costs the chain no result.

A.1 The boundary problem

Inside a human brain, billions of neural events occur every second, and most are not conscious. A motor command, a homeostatic signal, an immune-related response: each involves activation against surrounding non-activation, each has a physical boundary with its inactive neighbourhood, yet none constitutes experience. A pattern of activation in the visual cortex, triggered by photons transduced through the retinal pathway, is also a neural event — and this one is a conscious experience. What distinguishes it? Not anatomical location, which merely relocates the question. The question is structural; every competing theory must answer it. Physicalism gestures at "the right

organisation" without specifying what constitutes the crossing; panpsychism inherits the question at the system level (§A.5); dualism locates the boundary between substances without explaining interaction. Each offers a metaphysics; none characterises the boundary.

A.2 From the structural dependence thesis to uniqueness

The foundational principle answers directly. Content requires contrast, so any determinate content is individuated by at least one distinction (Chapter 2). A property shared by all physical events — mass, energy, charge, or "proto-consciousness" — cannot constitute the content of consciousness, because it is present on both sides of the boundary and therefore contrasts with nothing. Only features unique to the consciousness-constituting events can do the work.

A critic may charge that the structural dependence thesis concerns content individuation, so applying it to phenomenal experience equivocates. The charge would bite if the connection were stipulated. The equalisation principle (Chapter 6) supplies it as evidence instead: phenomenal experience and detected contrast co-vary as the thesis predicts, vanishing together categorically. The thesis applies to phenomenal content because phenomenal content behaves in accordance with it.

A.3 Three levels of uniqueness

The consciousness-constituting events are distinguished by three features, ordered by the kind of uniqueness each contributes.

Positional uniqueness: they are the terminal output of a dedicated processing pathway carrying a signal inward to a terminus, not restricted to external input, since visceral pain terminates in the insula, emotional experience in affective centres, proprioception in the somatosensory cortex. A carbon atom at the visual-cortex terminus and one in the kneecap are chemically identical; what differs is position. This is necessary, not sufficient.

Functional uniqueness: the events are the endpoint of a chain dedicated to carrying a specific input type inward, "pathway" meaning

the entire dedicated architecture, loops included, in place of a linear feedforward chain. Function narrows the field beyond position. A motor command is an origin, not a terminus; a homeostatic signal circulates regulatory information and carries no input type inward.

Structural uniqueness: the specific activation pattern, the triggered signature, that a dedicated architecture generates in response to a specific input type, defined by causal-structural role, whatever the physical composition. It preserves no copy of the input: what arrives at the higher visual areas bears little resemblance to the entering light, having been recoded through opponent processes, lateral inhibition, and cortical remapping. This is the deepest level and the one that constitutes qualitative character. Two events could share position and function yet, with different pathway architectures, produce different signatures — and the experience would differ.

The cochlear-implant case checks the inference against report. The implant transduces sound through an architecture unlike the intact cochlea's, and recipients describe a qualitatively unfamiliar mode of sound — a report the audiology literature attributes largely to the implant's coarse spectral resolution, which the account reads as one dimension of the altered signature: different architecture, different signature, different character, reported from the inside, with the inference from the second to the third the one the account stakes. The honeybee's tetrachromatic vision extends the same inference to where report is unavailable; there the different-character claim is a prediction the architecture licenses, with no datum the subject can confirm.

The order of discovery should be stated openly, because a reader will suspect it. The three criteria were found by asking what distinguishes the events that constitute experience from those that do not — reverse inference. The reply separates discovery from justification. What the framework derives is the constraint the criteria must satisfy: consciousness-constituting events must be distinguished by features unique to them (§A.2), since a shared feature contrasts with nothing.

Which uniqueness does the work is an inference to the best explanation of the dissociation evidence (§A.4); its warrant is

prospective, resting on the falsification conditions of §A.7, including the prediction that discriminates the interface account from its predictive-processing and global-workspace neighbours. The principle is carried as constructed, a positive proposal with live rivals, while the boundary constraint it answers to is established. The criteria could be wrong in a way the constraint could not, so the account stakes them on the predictions, whatever fit first suggested them.

A.4 The interface

The terminus activation occurs against the modality's absence: the condition obtaining when no activation of the relevant type is present — the blind field for vision, silence for hearing, numbness for affect. A distinction matters here. Darkness is a visual experience, minimal light detected against the visual field; the blind field is not a visual experience at all, since the congenitally blind person does not see darkness but has no visual phenomenology. The modality's absence is the structural zero of the detection system.

The experience-constituting events are located at a specific interface: the boundary where a terminus activation carrying a unique triggered signature meets the modality's absence. The quality question — why this character? — is answered by what differs across modalities, the signature, and not by what is shared, the absence and the particles. The absence is not a fourth ingredient. It is the contrast pole the signature needs in order to be a diff at all, which is Chapter 2's requirement applied at the interface. The condition holds of any mind.

One class of cases lets the interface reading do discriminating work beyond reinterpreting the conjuncts: those where a signature meets an absence that has just been created. After deafferentation the somatosensory terminus can fire into a limb that is gone, so the phantom is felt vividly at a location the body no longer occupies. The sudden silencing of a cochlear input can be followed by tinnitus.

On the interface account these are expected, since vividness follows from a signature meeting the modality's absence, whatever the peripheral source. The clinical picture is consistent with the reading and

does not select it, because accounts that locate the phantom in the reorganisation of the deafferented cortex (Ramachandran & Hirstein, 1998) predict persistence too; what it excludes is only a view on which experience tracks the intact input alone, which no live theory holds. That phantom and tinnitus are often more insistent than the veridical percept is what the interface reading expects; the test that separates it from its live rivals is the phosphene design of §A.7. Dreams are the complementary limit: internally generated activation acquires the characteristics of the live terminus signal and produces experience indistinguishable from perception, which the account predicts, since phenomenal character is constituted by the interface conditions, whatever the causal origin of the activation.

The harder test is the dissociation in which perceptual information guides behaviour while remaining unconscious; here the conjunction discriminates. In blindsight, damage to the primary visual cortex abolishes acknowledged visual experience across a region of the field, yet patients reliably detect and discriminate stimuli they deny seeing — residual capacity carried by routes that bypass the damaged area (Weiskrantz, 1986). This is the predicted profile. The residual pathway sustains a functional response, while the dedicated route that would generate the triggered signature at the visual terminus is the very thing the lesion removes; it is that conjunct, and not the bare registration, on which visual phenomenology depends. Blindsight is registration surviving the loss of the terminus signature.

The dorsal stream supplies the converse. On the two-visual-systems account (Goodale & Milner, 1992), the ventral pathway subserves conscious perceptual identification while the dorsal pathway transforms visual information into the real-time control of action. Dorsal information can scale a reaching grasp to an object’s dimensions without entering awareness: the visual-form-agnosic patient DF scales her grip accurately to a target whose size she cannot consciously report. In the present terms the dorsal output stands on the side of an origin of action and not of a perceptual terminus, so the account predicts its unconsciousness. DF’s converse pattern — ventral terminus damaged, conscious form perception lost, dorsal action-guidance spared — locates

the conscious character at the terminus signature, where the uniqueness principle places it.

The same analysis dissolves two apparent counterexamples. Motor imagery is felt because rehearsed movement arrives at the sensory termini where executed movement is felt, while the efferent command that drives real movement, being an origin, is silent (Jeannerod, 2006). The sense of agency is carried by the reafferent signals following the act, matched against prediction (Frith, Blakemore, and Wolpert, 2000), which is why it can be induced where no command issued; the readiness-potential (Libet, 1985), intentional-binding (Haggard, 2017), and illusion-of-control (Wegner, 2002) findings are the expected consequence of feeling the act, like everything else, at its terminus.

A.5 The equalisation constraint

The equalisation principle (Chapter 6) functions as a constraint on theories of consciousness, not only as confirmation of the zero-complexity constraint. The co-elimination of experience and contrast is categorical and uniform across modalities sharing no common mechanism. The person born blind has no visual phenomenology and no visual imagery, since an absent pathway entails an absent modality and not an empty one; the cochlear-implant recipient reports not silence replaced by sound but an unfamiliar dimension of experience opening.

The convergence across independent modalities with entirely different substrates confirms that the categorical character is a fact about the phenomenon, which no taxonomic convention could supply. Within a modality richness varies by degrees, while the modality itself is either present or absent, so the equalisation principle operates at that boundary.

Configuration-dependence and the argument against panpsychism. The constraint rules out theories on which experience is intrinsic to particles and not a property of configurations. If experience were intrinsic to particles, disassembling the consciousness-constituting configuration should produce gradual diminishment as each particle takes its micro-experiential contribution with it. Equalisation shows the

opposite: experience vanishes categorically when the interface conditions are no longer met; the same particles rearranged have no experience whatever. Experience belongs to the configuration; the particles carry none of it.

The panpsychist (Strawson, 2006; Goff, 2019) eliminates the boundary question by stipulating that all matter is proto-experiential, then inherits it again as the combination problem (Chalmers, 2017): what distinguishes the proto-conscious particles constituting visual experience from those constituting a rock? A sophisticated panpsychist replies that proto-consciousness is not a static shared property but combinatorial, interacting differently by configuration (Goff, 2017; Mørch, 2019).

Grant the combinatorial reading in full; two arguments then need separating. The categorical-versus-gradual argument reaches only the naive panpsychist who locates the experience equalisation removes in the particles; the combinatorial panpsychist keeps micro-experience in the particles and locates the macro-experience at the combination, so its categorical loss under disassembly is what she predicts too, so the equalisation constraint does not discriminate against her. What does reach her is the boundary question, which survives the combinatorial reading intact. What distinguishes the configurations in which proto-experiential properties combine to produce visual experience from those in which they produce a rock must cite features of the configuration — pathway, terminus, signature, interface — while proto-consciousness, present in both, contributes nothing to identifying which is which. The parallel is close: hydrogen behaves differently in H_2O than in H_2SO_4, but we need not attribute "proto-wetness" to hydrogen to explain why water is wet, since the combinatorial behaviour is fully explained by the physical properties plus the configuration. So here, the work is done by structural uniqueness, and adding proto-consciousness to the particles does no further work on the boundary question.

The panpsychist may reply that proto-consciousness answers the hard problem and not the boundary question — why there is experience at all. The chapter takes no position on that, confining its claim to the boundary question, where a property present everywhere marks nothing.

The experience was never in the particles; it was in the unique situation the particles occupied.

The identity hypothesis. The constraint favours an identity reading without establishing it: that experience and the structurally unique configuration at the interface are the same thing. In superconductivity, resistance drops to zero at a critical temperature while the system persists with rich internal structure. In the experience/contrast case, when contrast reaches zero the modality itself ceases to exist. The blind person is not a seer in a special state; she is not a seer at all. This is what identity predicts, since removing one side makes both vanish — though whether it suffices to establish identity or only to motivate it is left open.

A.6 The interface account among its neighbours

Two rival theories have the interface account as their target and no part of the hierarchy; they are met here.

Integrated Information Theory (Tononi, 2004; Tononi et al., 2016) identifies consciousness with integrated information and selects among candidate systems by maximum Φ. The two accounts address related but distinct problems. The exclusion postulate selects a system by a quantitative criterion, while the uniqueness principle identifies, by a qualitative one, what makes any system's activity experiential and not computational — and could explain why the system IIT selects has the specific character it has, which Φ alone cannot.

IIT instantiates the zero-complexity constraint and does not confirm it independently. $\Phi > 0$ requires internal differentiation, which is the presupposition the constraint identifies, so IIT's agreement displays that presupposition and supplies no evidence for it from outside (Chapter 6, §5). The point holds regardless of the major objections to IIT (Aaronson, 2014; Bayne, 2018), since it depends only on that formal requirement.

Predictive processing (Seth, 2021) holds that experience is "controlled hallucination" constituted by the brain's generative models. The accounts are compatible where predictions are pathway-specific

and prediction errors are detected contrasts, but two tensions divide them.

The first is about what is constitutive. For Seth it is the predictive model's best guess; for the interface account it is the triggered configuration at the interface, whether or not it functions as a prediction. A newborn's first visual experience, formed before experience has trained any predictive model, should on Seth's account be degraded or absent, while on the interface account it is present whenever the pathway is active against the modality's absence — a divergence the equalisation constraint resolves in the interface account's favour. The second is about quality. Identical signatures under different prediction contexts should on Seth's view differ in character and on the interface view not, which is a substantive empirical disagreement and the discriminating test of §A.7.

A.7 Falsification conditions for the interface

The uniqueness principle carries a falsification condition of its own: experience persisting when the interface conditions are absent — a motor command, homeostatic signal, or immune response constituting phenomenal experience — would fell it. The equalisation constraint adds another: experience diminishing gradually, and not categorically, as contrast is removed within a modality would fell it.

One prediction discriminates the account from its neighbours, beyond exposing it, so it is worth stating because the interface account and the predictive-processing account diverge on it. Where a dedicated pathway delivers a triggered signature at its terminus against the modality's absence, the account predicts phenomenal character even when no predictive model is in play and no global broadcast has occurred.

The clean test is direct cortical stimulation. Stimulation of the primary visual cortex produces phosphenes, experienced flashes reported as seen, with no external stimulus, no inference, and no predictive model about the world's causes (Penfield & Rasmussen, 1950; Winawer & Parvizi, 2016); dynamic stimulation traced across the

cortical surface yields form percepts without training (Beauchamp et al., 2020). On the interface account this is what a triggered signature against the modality's absence should yield, whatever the origin of the activation.

The predictive-processing theorist can reply that standing priors are engaged even by a bare flash. But a prior general enough to be always already present predicts nothing in particular, so it cannot constitute this phosphene instead of none. The global-workspace theorist has the harder case, since stimulation-evoked phosphenes can be reported under restricted attention, where the interface account predicts experience and the workspace account its absence. Recurrent-processing theory (Lamme, 2003) predicts the same dissociation of awareness from attention and report, locating feltness in local recurrence within sensory cortex; the interface account agrees on where the experience is and differs on what constitutes it: the signature against the modality's absence, which the loops of a dedicated architecture may realise without being it.

The evidence is not yet decisive, because stimulation studies vary in the attentional controls the adjudication needs. The case is nonetheless the right shape to adjudicate, where a lesser case would only embarrass; it is in that shape, and not as a settled result, that the General Conclusion's falsifiability inventory (§2) enrols it among the framework's live exposures. The equalisation constraint indicates which way it should fall: what vanishes when contrast is removed is the activated configuration at the interface, not the model or the broadcast.

Two further exposures sharpen the adjudication. Sensory substitution supplies one. Devices that deliver visual information through tactile arrays route the signal to a somatosensory terminus, so the account predicts that the resulting experience carries the somatosensory signature however visually competent the behaviour becomes — spatially informative touch, never seeing — because the signature belongs to the terminus and its contrast class. Practised users do achieve the behavioural competence at issue (Bach-y-Rita & Kercel, 2003), who read those reports as seeing with the brain and not as touch.

A substitution user who developed visual phenomenology without recruitment of the visual terminus would tell against the account.

The attentional-control gap in the stimulation literature supplies the other; the design that would close it can be specified: phosphene induction under a concurrent attentional load, with detection probed by no-report measures alongside verbal report. The interface account predicts that the signature survives the withdrawal of attention and broadcast; the workspace account predicts its elimination. Either outcome would be an adjudication, which is what enrols the design, alongside the phosphene case itself, among the framework's live exposures.

A.8 The multiple realisability objection

The objection is that identity with a specific neural configuration is chauvinistic: an octopus or an artificial system with a different substrate could not be conscious.

In reply, the uniqueness principle identifies structural conditions, not substrate conditions. Any system — biological, silicon, or otherwise — with a dedicated pathway generating a unique triggered response at an interface with absence satisfies them, because the triggered signature is defined by structural role, not physical composition. Whether "functionally equivalent" can be specified sharply enough is a question the chapter flags and does not resolve (Polger & Shapiro, 2016).

References

Aaronson, S. (2014). *Why I am not an integrated information theorist (or, the unconscious expander)*. Shtetl-Optimized, 21 May 2014. https://scottaaronson.blog/?p=1799

Bach-y-Rita, P., & Kercel, S. W. (2003). Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences*, 7(12), 541–546.

Bayne, T. (2018). On the axiomatic foundations of the integrated information theory of consciousness. *Neuroscience of Consciousness*, 2018(1), niy007.

Beauchamp, M. S., Oswalt, D., Sun, P., Foster, B. L., Magnotti, J. F., Niketeghad, S., Pouratian, N., Bosking, W. H., & Yoshor, D. (2020). Dynamic stimulation of visual cortex produces form vision in sighted and blind humans. *Cell*, 181(4), 774–783.e5.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227–247.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30(5–6), 481–499.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford University Press.

Chalmers, D. J. (2007). Phenomenal concepts and the explanatory gap. In T. Alter & S. Walter (Eds.), *Phenomenal concepts and phenomenal knowledge: New essays on consciousness and physicalism* (pp. 167–194). Oxford University Press.

Chalmers, D. J. (2017). The combination problem for panpsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary perspectives* (pp. 179–214). Oxford University Press.

Chalmers, D. J. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies*, 25(9–10), 6–61.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200–227.

Frankish, K. (2016). Illusionism as a theory of consciousness. *Journal of Consciousness Studies*, 23(11–12), 11–39.

Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society B*, 355(1404), 1771–1788.

Goff, P. (2017). *Consciousness and fundamental reality*. Oxford University Press.

Goff, P. (2019). *Galileo's error: Foundations for a new science of consciousness*. Pantheon.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20–25.

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4), 196–207.

Jeannerod, M. (2006). *Motor cognition: What actions tell the self*. Oxford University Press.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, 7(1), 12–18.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64(4), 354–361.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*(4), 529–539.

Loar, B. (1990). Phenomenal states. In J. E. Tomberlin (Ed.), *Philosophical perspectives 4: Action theory and philosophy of mind* (pp. 81–108). Ridgeview Publishing Company.

Mørch, H. H. (2019). The argument for panpsychism from experience of causation. In W. Seager (Ed.), *The Routledge handbook of panpsychism* (pp. 269–284). Routledge.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.

Papineau, D. (2002). *Thinking about consciousness*. Oxford University Press.

Penfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man: A clinical study of localization of function*. Macmillan.

Polger, T. W., & Shapiro, L. A. (2016). *The multiple realization book*. Oxford University Press.

Prinz, J. J. (2012). *The conscious brain: How attention engenders experience*. Oxford University Press.

Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs: The D. O. Hebb lecture. *Brain*, 121(9), 1603–1630.

Rosenthal, D. M. (2005). *Consciousness and mind*. Clarendon Press.

Seth, A. (2021). *Being you: A new science of consciousness*. Faber & Faber.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74(11), 1–29.

Strawson, G. (2006). Realistic monism: Why physicalism entails panpsychism. *Journal of Consciousness Studies*, 13(10–11), 3–31.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5, 42.

Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17(7), 450–461.

Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-report paradigms: Extracting the true neural correlates of consciousness. *Trends in Cognitive Sciences*, 19(12), 757–770.

Tye, M. (1995). *Ten problems of consciousness: A representational theory of the phenomenal mind*. MIT Press.

Wegner, D. M. (2002). *The illusion of conscious will*. MIT Press.

Weiskrantz, L. (1986). *Blindsight: A case study and implications*. Clarendon Press.

Winawer, J., & Parvizi, J. (2016). Linking electrical stimulation of human primary visual cortex, size of affected cortical area, neuronal responses, and subjective experience. *Neuron*, 92(6), 1213–1219.