

The Puppeteer Mind

Dualism, Neural Dependence, and the Explanatory Redundancy of Immaterial Control

Phil Stilwell

Independent Scholar

ORCID: 0009-0009-3725-8682

Revised August 2026

Central claim

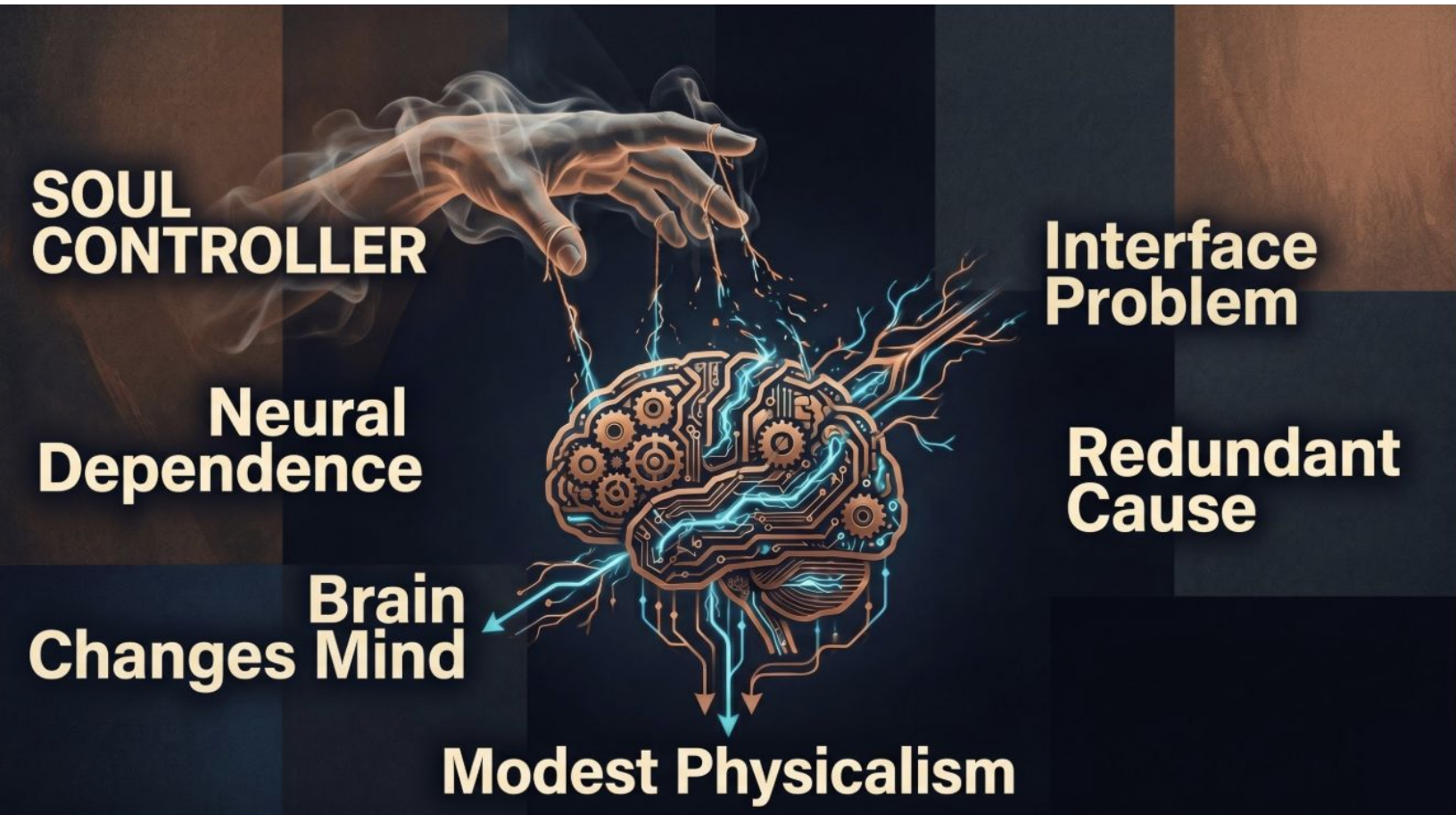
Interactionist substance dualism says that a nonphysical subject helps cause thought and action. Neural interventions establish brain dependence, but mediated control predicts that too. The added controller gains public causal-explanatory support only through a distinctive difference or independent evidence; otherwise observational equivalence leaves it unconfirmed, without settling consciousness or disproving every dualist view.

Preview: The causal fork

Controller claim	What would distinguish it	Comparative result
Difference-making control	The subject predicts or constrains an outcome not supplied by the no-controller account.	A risky hypothesis: success could give the controller incremental support.
No controller-specific difference	For the tested conditions, both views make the same public predictions.	Shared evidence supports neural dependence, not the added controller.

The paper tests the evidence for additive immaterial control, not every theory of consciousness or every use of mental explanation.

Visual preview:



This heuristic—not a literal portrait of any named dualist theory—maps neural-dependence, interface, and redundancy burdens; its result is comparative, not a proof of exhaustive physicalism.

Abstract

Interactionist substance dualism treats a person as including a nonphysical subject that can affect the brain and body. The view promises more than survival or a private realm of experience: it presents the immaterial subject as part of the causal explanation of memory, deliberation, choice, and action. This paper asks what incremental support that extra controller receives once the neural machinery is specified. The governing test is ordinary difference-making. Mediation is not enough: a pianist acts through a piano. The discriminating questions are whether total physical conditions suffice for the outcome, whether the controller supplies a distinctive counterfactual constraint or successful prediction, and whether independent evidence supports its causal power.

The evidence does more than correlate minds with brains. Selective lesions alter particular forms of memory and judgment; stimulation changes neural states and can change experience and behavior; perturbing cortical activity helps distinguish conscious from unconscious conditions; changes in interhemispheric communication alter access and control; and the timing of voluntary movement can be modeled through ongoing neural accumulation. These methodologically diverse findings establish neural causal relevance and increasingly specific brain dependence. Yet an instrument dualist can expect neural dependence too, so the findings are largely shared evidence rather than a direct test of the two ontologies.

The strongest dualist reply is that the brain is an instrument: the subject acts only through it. That reply defeats any inference from mediation alone. It does not, without more, make the neural record evidence for the added controller. The interactionist accounts assessed here have not yet supplied a validated controller-specific difference or independent support for the controller's role in these public outcomes. The findings therefore give the controller no incremental causal-explanatory support over a no-controller brain-dependent account. Observational equivalence remains neutral between them; any provisional preference then rests on economy, testability, and research productivity, not on a disproof of dualism. This conclusion does not solve phenomenal consciousness, assume global physical sufficiency, or refute independently supported dualist theories.

Keywords: philosophy of mind; substance dualism; mental causation; neural dependence; causal exclusion; intervention; physicalism

1 The Controller Under Examination

The puppeteer picture is an intentionally simple heuristic. The brain is an instrument; the real thinker is a nonphysical subject using it. The brain receives sensory input, stores traces, and issues motor commands, but an immaterial mind supplies conscious thought, deliberation, and choice. Defenders develop this family of ideas in importantly different ways. Moreland and Rae place the soul inside a wider account of human nature, Hasker treats the self as an emergent individual, and Swinburne defends a soul that interacts with a body while retaining libertarian agency (Moreland and Rae 2000; Hasker 1999; Swinburne 2013). None should be reduced to a cartoon of a tiny person inside the skull.

The causal claim nevertheless needs precision. This paper targets *interactionist substance dualism*: the view that a person includes a nonphysical subject, distinct from the body, that can causally affect neural activity or bodily action. The target is not every position called dualist. An epiphenomenal view may hold that experience is nonphysical but never affects the body. Property dualism may treat phenomenal properties as irreducible without adding a second controller. An emergent view may make a subject wholly dependent on the organized brain even while denying standard physicalism. Those positions face their own problems, but a test of additive control cannot refute them all at once.

The central question is therefore narrow: when the immaterial controller is added to an account of human action, what explanatory work does it perform? The question is not answered by pointing out that consciousness remains difficult. Chalmers's hard problem concerns why physical processing is accompanied by experience at all, while Searle insists that the first-person reality of consciousness cannot

be erased by a third-person description (Chalmers 1996; Searle 1992). Both challenges can remain serious even if a second substance contributes no causal control.

The answer defended here is comparative. The relevant rival is a *no-controller brain-dependent account*: an account that assigns the public causal work to embodied and neural processes without adding a second substance. Such models identify structures whose alteration changes capacities, networks whose stimulation changes states, and organization whose disruption changes access and control. A controller hypothesis earns incremental support only if it improves on that work or has independent support. Merely redescribing the unexplained remainder as the activity of a soul does not do so.

2 What It Means to Make a Causal Difference

In ordinary inquiry, a cause is not merely something mentioned alongside an outcome. It makes a difference. If changing the position of a switch changes the lamp while the other relevant conditions remain stable, the switch has causal importance. Woodward develops this thought into an interventionist account: causal relationships support counterfactual claims about what would happen under suitable changes (Woodward 2004). The present argument does not require every detail of that theory. It uses a modest standard: a proposed controller should change, constrain, predict, or help manipulate some outcome.

This standard separates three questions that are often run together.

1. **Does a conscious subject exist?** The reality of experience is not in dispute.
2. **Are mental explanations useful?** Beliefs, intentions, and reasons can explain behavior at a higher level even if they are physically realized. Fodor's defense of special sciences shows why higher-level patterns need not disappear when a lower-level account is available (Fodor 1974).
3. **Does an additional nonphysical subject contribute a separate causal influence?** This is the distinctive interactionist claim.

The third claim cannot borrow support merely from the first two. A person can be conscious, and an intention can be a useful cause, without the intention being the action of a separate substance. Conversely, denying a separate controller does not make reasons unreal. It leaves open whether mental and physical descriptions pick out one process, whether mental properties are realized by physical organization, or whether another relation is best.

2.1 Mediation, sufficiency, and explanatory redundancy

The pianist analogy shows why three claims must be kept separate.

1. **Mediation:** the controller's influence reaches behavior through the brain. Neural dependence is expected, just as performance depends on an instrument.
2. **Physical sufficiency:** the total neural state, bodily state, and relevant physical inputs fix the public outcome. This is stronger than mediation and is not assumed here as a global theorem.
3. **Observational equivalence:** the controller view and the no-controller account make the same public predictions in the tested domain. This can occur even if the controller exists and acts.

Ontological denial says that the proposed controller does not exist. Explanatory redundancy is different: a controller could exist yet add no prediction, counterfactual constraint, intervention, or independently supported understanding to a particular comparison. The paper argues for this bounded evidential conclusion, not ontological elimination. If two views are observationally equivalent, the shared neural evidence does not disprove the controller; it also does not give the controller incremental support.

This distinction matters because a hypothesis can survive by retreating from risk. Suppose the controller initially explains why a person raises a hand. Neuroscience then traces the relevant sensory processing, valuation, motor preparation, inhibition, and muscle activation. The dualist may coherently say that the controller works through every stage. Mediation alone settles nothing. The evidential question is whether the controller view constrains that sequence differently, predicts some feature the no-controller account does not, or has independent support for adding its cause.

2.2 The causal fork

The controller therefore faces a fork.

1. **Controller-specific difference:** the controller predicts or constrains some neural transition or behavioral outcome beyond what the no-controller account supplies. The theory is then empirically risky. It should identify a stable pattern, even if present measurement cannot yet detect it.
2. **No controller-specific difference:** the controller may act through the brain, but the two views make the same public predictions. The shared evidence then supports neural dependence without favoring the added controller. Any provisional preference for the no-controller account rests on economy, testability, or research productivity rather than on an experimental disproof of the controller.

This is not the claim that every true entity must be directly manipulated or easy to measure. Historical and latent causes are often inferred from the patterns and counterfactual relations they explain. The weaker claim is that a public causal hypothesis must constrain the world somehow. If no possible outcome counts against a controller, no outcome is more expected because of it, and no independent argument establishes its relevant power, the neural findings have not made that controller an evidentially better explanation.

3 From Correlation to Intervention

Neural dependence is often summarized too loosely: damage the brain and the mind changes. That slogan invites the instrument reply. Damage a violin and the music changes, but the player is not made of wood. The better case looks at several kinds of causal evidence and asks what each contributes.

The ledger below prevents two opposite errors. The first is triumphalism: no experiment in the list proves that only physical things exist. The second is evidential flattening: interventions are not mere correlations. When carefully interpreted, they identify necessary components, manipulable levers, state-dependent dynamics, and organized routes through which capacities are expressed. Those are positive causal findings even when the metaphysical comparison remains underdetermined.

The cumulative force comes from methodologically diverse convergence. Memory systems decompose in repeatable ways, moral judgment changes after focal damage, circuit manipulation changes neural and sometimes experiential states, responsiveness tracks evoked neural complexity, access changes with connectivity, and voluntary-action timing follows evolving neural dynamics. Together these results strongly support neural causal relevance and an increasingly detailed brain-dependent research program. They do not accumulate as six independent disproofs of a mediated controller, because an instrument view can expect dependence in every case. Their comparative pressure is instead incremental: the neural program supplies detailed, testable structure, while the controller accounts assessed here have not yet added a successful controller-specific constraint or independent causal support.

Table 1: What neural interventions establish, and how far they discriminate between the two accounts.

Evidence type	What it establishes	Comparative force
Selective lesion	Removing or degrading a neural component selectively impairs a capacity, identifying structured dependence and possible necessity.	Both accounts can expect instrument damage; the pattern discriminates only if one view predicts its structure better.
Direct stimulation	Manipulating tissue changes neural activity and sometimes mood, perception, movement, or report, identifying a usable lever.	Manipulability supports a neural causal role, but therapeutic efficacy and controller discrimination require further evidence.
State perturbation	The complexity and spread of an evoked response vary across wakefulness, sleep, anesthesia, and severe injury.	A measure can track conscious level while remaining neutral about ontology and phenomenal character.
Connectivity change	Altered communication changes access, report, and coordinated control.	It supports organization-dependent unity; it does not by itself decide whether one or two subjects exist.
Temporal modeling	An accumulating neural process explains averages once treated as evidence of an early unconscious decision.	It weakens one dualist rescue without deciding responsibility or making conscious reasons causally idle.

4 Memory, Judgment, and the Compositional Person

Memory is not one container hidden behind the brain. Experimental and clinical work distinguishes declarative recollection from several nondeclarative systems, with different roles for the medial temporal lobe, amygdala, striatum, cerebellum, and distributed cortex (Squire 2004). A person can lose the ability to form new episodic memories while retaining skills, habits, or emotional learning. This is more informative than a generic claim that a damaged instrument performs badly: the organization of what can be learned, retained, retrieved, and consciously reported changes with the organization of the brain. An instrument view remains compatible with the pattern, but the detailed dissociation comes from the neural account.

This matters to the controller hypothesis because memory is often treated as a possession of the enduring subject. If the subject stores complete memories independently, the theory must explain why the public person loses precisely the memory operation supported by the damaged system. The dualist can reply that the memory remains in the subject but can no longer pass through the instrument. Yet the reply does not predict the detailed dissociation until the neural account supplies it. The neural model tells us which operation fails and often why; the controller is added afterward.

Focal brain damage can also change evaluative performance. In a small lesion study, patients with bilateral ventromedial prefrontal damage gave an unusually utilitarian pattern of judgment in a selective class of emotionally charged moral dilemmas (Koenigs et al. 2007). The study does not show that morality is located in one spot, that the judgments were simply wrong, or that six patients reveal a universal law. It shows something more careful: altering a specific neural system can selectively change the balance of emotion and judgment while leaving other responses comparatively intact.

The rational conclusion is not that persons are unreal collections of modules. An orchestra remains one performing group even though different instruments carry different parts. Likewise, a person can be a genuine agent whose capacities are composed and coordinated. What the evidence challenges is the claim that unity requires an indivisible controller whose central powers remain causally independent of the components. The observed person changes in structured ways when those components change.

5 Stimulation and Conscious State

Intervention becomes stronger when investigators add or alter activity rather than merely observe loss. Deep brain stimulation has been studied as a way to modulate dysfunctional circuits. In an early uncontrolled six-patient study of treatment-resistant depression, stimulation near the subgenual cingulate region was associated with improvement in four participants and with changes across connected limbic and cortical sites (Mayberg et al. 2005). A later multisite randomized trial confirmed feasibility but found no statistically significant antidepressant advantage for active over sham stimulation during its six-month blinded phase (Holtzheimer et al. 2017). A 2024 systematic review and meta-analysis found long-term improvement across studies but also found larger effects in open-label than randomized trials and could not adequately distinguish the time course of active from sham improvement in the limited controlled data (Reddy et al. 2024).

The distinction matters. Implanted stimulation is plainly a physical manipulation of circuit activity; whether it produces a durable therapeutic benefit beyond placebo and study effects is a further question. Basic manipulability can establish a neural causal lever without proving a robust treatment effect, explaining phenomenal character, or deciding between materialist and instrument-dualist ontologies.

A different line of work perturbs the cortex with transcranial magnetic stimulation and records the resulting electrical response. Casali and colleagues introduced a perturbational complexity index and tested it across wakefulness, dreaming, several anesthetic conditions, and patients recovering from coma (Casali et al. 2013). Conscious conditions tended to support responses that were both integrated and differentiated, whereas unconscious conditions produced responses that were more local or stereotyped. The measure does not read the content of a mind, and later work may refine its interpretation. Its importance here is methodological: a controlled physical perturbation reveals a state-dependent property closely related to the capacity for consciousness.

An immaterial-controller view can accommodate these results. It can say that the subject needs a functioning instrument to express mood or consciousness. But the accommodation does not yet discriminate the view. The intervention tells investigators where to stimulate, which network response to measure, and how the result varies across states. The controller supplies no additional tested constraint on those relations.

This is the central asymmetry. Brain-dependent accounts remain incomplete, but they produce a research program in which intervention changes prediction and control. A controller theory may be metaphysically richer while being causally poorer. Richer ontology does not automatically produce a better explanation.

6 Unity Without a Simple Split-Brain Story

The source of personal unity is often offered as a reason for an indivisible soul. Split-brain research seems at first to reverse the argument: sever communication between the hemispheres and one person appears to divide. That simple story is too strong.

Classic findings show striking dissociations in perception, report, and action after surgical separation of the hemispheres. More recent work complicates the inference from divided processing to two conscious subjects. Pinto and colleagues reported divided visual perception alongside evidence they interpreted as one conscious agent, and a later multidisciplinary review emphasized both the robust disconnections and the continuing disputes about consciousness (Pinto et al. 2017; Haan et al. 2020). The responsible conclusion is therefore not that each hemisphere houses a separate person.

The narrower result is enough. Access, coordination, and report depend on communication routes. Change those routes and the functional unity of the person changes in systematic ways. A theory

that treats unity as primitive still needs to explain why the supposedly unified subject cannot make information available across the damaged channel.

The instrument reply again remains possible: the subject is unified, but the divided instrument prevents unified expression. The cost is familiar. Every discriminating feature of unity is assigned to the instrument, while the controller's unity makes no independent public difference. The reply protects a hidden fact about the subject by giving the brain the entire explanatory role in behavior.

7 Agency Without Libet Theater

Voluntary action is a tempting place to locate the puppeteer. A person considers reasons, forms an intention, and acts. Early readiness-potential experiments were often presented as though the brain decides first and consciousness arrives late. That interpretation was always too simple, and it encouraged an equally simple dualist rescue in which consciousness retains a last-second veto.

Schurger, Sitt, and Dehaene modeled the readiness potential as the average result of spontaneous neural fluctuations accumulating toward a movement threshold (Schurger, Sitt, and Dehaene 2012). On this account, the familiar rising signal need not mark one unconscious decision made long before awareness. It can emerge from noisy preparation whose threshold crossing determines when a self-initiated movement occurs. A later review by Schurger and colleagues argues that this and related advances require reassessing what the readiness potential can show about volition and free will (Schurger, Hu, et al. 2021). The model does not establish determinism, eliminate conscious reasons, or settle responsibility. It shows why the evidence should be interpreted within a layered control process rather than a contest between a brain and a ghostly veto.

Agency includes attention, memory, valuation, inhibition, forecasting, self-monitoring, and sensitivity to reasons. These capacities operate at different levels of description. Saying that an intention caused an action can be true and useful even when the intention is realized by organized neural activity. The paper's target is not mental causation as such. It is a separate controller whose causal role remains after the full agent-level and neural-level process is described.

If the controller changes a threshold, interrupts a motor plan, or selects between options, the theory should say what difference this intervention makes. If the controller's choice is always perfectly mirrored by the evolving neural process, the public evidence cannot distinguish two causes from one process described at two levels. In that case, the separate substance has not yet earned causal credit.

8 The Interface and Exclusion Burdens

The traditional interaction problem asks how a nonphysical subject can affect a physical body. The question is sometimes posed badly, as if every cause had to strike another object like a billiard ball. Modern theories of causation need not require contact or the transfer of a visible substance. The real burden is more general: what connects this subject to this brain, and what difference does the connection make?

Kim's work on mental causation sharpens the second part (Kim 1998). If a physical event already has a sufficient physical cause, adding an independent mental cause appears either unnecessary or systematically overdetermining. The problem also pressures nonreductive physicalists, not only substance dualists. Anyone who treats a mental cause as wholly distinct from a sufficient physical cause must explain why the outcome has not simply been caused twice.

This paper does not assume that physical causal closure is an exceptionless theorem. Pitts argues that familiar conservation-law objections often smuggle in the very closure that interactionists deny; on his account, the serious issue should be investigated empirically in the fine operation of the brain rather than settled by a slogan about energy (Pitts 2020). That objection is well taken. A dualist should not be dismissed because an elementary conservation story says so.

Granting the objection strengthens the present test. If interaction is a live empirical possibility, then interactionist dualism can expose itself to empirical risk. A nonphysical influence might alter a transition probability, break an otherwise stable dependence, support a novel counterfactual pattern, or improve prediction beyond a well-specified neural model. An unexplained residual by itself would not be enough: it might reflect measurement error, omitted physical variables, or a poor model. The controller theory must motivate the pattern in advance or explain it better than those alternatives. No law guarantees in advance that investigators must fail to find such evidence.

The absence of a current signature does not prove that none exists. It changes the allocation of incremental support. Neural models have repeatedly identified manipulable dependencies, while the assessed controller hypotheses have not identified a validated distinct effect or constraint. Where the views are observationally equivalent, the experiments themselves are neutral between them. The no-controller program's provisional advantage then comes from doing the predictive and manipulative work with fewer independent commitments, clearer tests, and an established research trajectory—not from evidence that directly reveals the controller's absence.

9 The Strongest Dualist Replies

9.1 The instrument is supposed to mediate everything

The most direct reply says that the dependence pattern is exactly what dualism predicts. A pianist cannot play through a broken piano, and a subject cannot think or act through a damaged brain. The controller should not be expected to bypass the instrument.

The reply defeats a crude inference from damage to nonexistence. It does not yet produce confirmation. If the instrument determines every capacity-specific change, then a brain-dependent account and a completely mediated controller account expect the same public pattern. The data support the shared neural-dependence claim, not the extra controller. To gain comparative support, the dualist needs evidence or argument that is more expected if the controller exists.

9.2 The controller explains first-person subjectivity

A defender may say that the controller is not introduced mainly to predict lesions. It explains why there is a subject of experience at all. Chalmers's arguments against reductive explanation and Searle's defense of first-person ontology keep that issue open (Chalmers 1996; Searle 1992).

This reply changes the target in a legitimate way. A theory of phenomenal subjectivity may have value even if it adds no public control channel. It should then be assessed as a theory of consciousness, not advertised as an improved causal explanation of memory, choice, or movement. The hard problem does not by itself show that a nonphysical subject pushes neurons.

9.3 The self emerges and then acts

Hasker's emergent dualism ties the self to the complex organization from which it arises, avoiding the picture of a preexisting soul merely inserted into a machine (Hasker 1999). This view can expect deep neural dependence while retaining a nonphysical subject.

The view is a stronger comparator than a radically independent puppeteer. It also sharpens the causal question. If the emergent self acts back on the brain, what changes because of that action? If its state is fixed by the physical base and its effects are fully carried by that base, the causal exclusion problem returns. If it contributes an additional influence, the influence should eventually appear in a difference-making model. The present evidence does not refute emergence; it leaves the added causal contribution unconfirmed.

9.4 The subject is an agent cause, not another event cause

A deeper reply rejects the assumption that every action is caused only by prior events. On an agent-causal or substance-causal view, the person exercises a basic power to act for reasons; the agent, not an extra neural event or a pulse of energy, is the cause. Lowe develops such an account of personal agency and substance causation (Lowe 2008). If that is the proposal, demanding a controller-shaped energy residual would mistake the kind of cause being offered.

The reply is serious, but it does not remove every evidential burden. Difference-making need not mean direct manipulation of an immaterial substance. An agent cause could constrain which choices occur under specified reasons, support counterfactuals about what the agent would do, or explain an out-of-sample pattern not supplied by the neural model. If agent causation is instead offered only as a metaphysical account of who acts while making exactly the same public predictions, the neural findings remain neutral between it and the no-controller account. Its case must then rest on independent arguments for agency or substance, not on neural dependence as such.

9.5 Fine-grained brain prediction may be impossible

Swinburne argues that conscious and brain events differ deeply and that the total conscious and neural conditions relevant to difficult choices may never repeat in a way that supports a complete predictive law (Swinburne 2013). That warning blocks an unrealistic demand for exact prediction of an individual life.

The argument here does not require perfect prediction. Medicine and cognitive neuroscience routinely infer causes without reproducing a complete person. They use controlled contrasts, lesion patterns, stimulation, within-person changes, and probabilistic models. A controller need not yield a clockwork law. It must only support some stable difference that is not borrowed entirely from the neural account.

10 What Would Change the Verdict

A fair argument should state what could make it lose. Direct manipulation of a proposed controller would be powerful evidence, but interventionism does not require that every cause be physically adjustable. Historical events and latent causes can be supported by distinctive counterfactual patterns. At least four developments would change the comparison.

1. **An advance, out-of-sample prediction:** an interactionist theory predicts a stable class of neural or behavioral transitions that a well-supported no-controller model misses. A residual noticed only after a model fails is not enough.
2. **A controller-specific counterfactual pattern:** the theory explains how outcomes would vary with reasons, intentions, or agent-level conditions even if the proposed cause cannot be directly manipulated. The pattern must be independently motivated rather than relabeled neural variance.
3. **A principled psychophysical law:** a dualist account specifies why particular subjects pair with particular brains and yields risky expectations about when the pairing or influence succeeds and fails.
4. **An independent argument:** a sound argument from possibility or necessity, knowledge, or metaphysics establishes a nonphysical subject and its causal powers even if current experiments do not discriminate it.

These standards are demanding but not hostile. Brain-dependent accounts face parallel demands: they must improve mechanisms, avoid mistaking report for consciousness, connect levels of explanation, and explain why organized activity has a first-person character. The current evidential advantage is not a blank check for future neuroscience.

The important asymmetry is present performance. Brain research supplies interventions that change mental capacities and models that organize those changes. The interactionist accounts assessed here currently supply no validated controller-specific constraint or independent causal support in these domains. Future evidence or argument could alter that balance; current mystery does not.

11 Conclusion

The puppeteer mind is not ruled out merely because mental life depends on the brain. A dualist can coherently say that the brain is the subject's instrument. Mediation is therefore not the decisive issue. The question is whether the added subject gives us a controller-specific reason to expect, explain, or change any public outcome.

The evidence identifies capacity-specific neural dependence and usable causal levers. Memory systems decompose with different neural injuries. Focal damage changes selected forms of judgment. Stimulation manipulates circuits, though therapeutic efficacy must be tested separately. Perturbation distinguishes conscious and unconscious conditions. Connectivity shapes coordinated access. Neural dynamics help explain the timing of voluntary movement. Each line has limits, but together they give the brain-dependent research program explanatory depth. Because an instrument dualist can also predict neural dependence, these findings are shared evidence for mediation rather than direct evidence for a brain-only ontology.

The controller nevertheless faces an evidential fork. If it adds a distinct influence, the theory should identify a successful prediction, counterfactual constraint, psychophysical law, or other independently supported difference and accept empirical risk. If it makes the same public predictions as the no-controller account, the experiments are neutral between the ontologies and give the added controller no incremental causal-explanatory support. Economy, testability, and research productivity then favor the no-controller account provisionally; they do not prove it complete.

The bounded conclusion matters. Global physical sufficiency has not been assumed or established. Consciousness remains philosophically difficult. Higher-level mental causation can remain real. Epiphenomenal, property-dualist, emergent, agent-causal, and independently supported soul theories require separate assessment. What fails is the simple promise that neural dependence itself confirms an immaterial controller or that adding one automatically improves the public causal explanation. A controller earns that explanatory place by constraining something.

A Formal Reconstruction of the Central Thesis

1. **Premise 1:** Neural findings give an added controller incremental causal-explanatory support over a no-controller account only if they are more expected on the controller view, the controller supplies a successful distinctive prediction or counterfactual constraint, or independent evidence establishes its relevant causal power.
2. **Premise 2:** Selective lesions, stimulation, state perturbation, connectivity changes, and temporal models establish structured neural dependence and causal relevance across memory, judgment, mood, conscious access, coordinated control, and action.
3. **Premise 3:** Both the no-controller account and an instrument-mediated controller view can predict that these public capacities depend on neural organization, and the assessed controller accounts do not make the cited patterns more expected by specifying their detailed structure.
4. **Premise 4:** The interactionist accounts assessed here have not yet supplied a validated controller-specific prediction or counterfactual constraint in these domains, nor independent support for the controller's causal role in these outcomes.
5. **Conclusion:** Therefore the cited neural findings currently give the added controller no incremental public causal-explanatory support over the no-controller brain-dependent account in these domains.
6. **Corollary:** If a controller theory supplies a distinctive difference, it becomes empirically risky and can gain or lose support. If the two views remain observationally equivalent, the evidence is neutral between their ontologies; any provisional preference for the no-controller account comes from economy, testability, and research productivity.
7. **Scope limit:** This conclusion does not solve phenomenal consciousness, establish global physical sufficiency, eliminate higher-level mental causation, or refute dualist theories supported by independent evidence or argument.

The inference is valid at its stated, evidential strength. Premise 1 specifies the routes by which the neural record could favor the added cause. Premises 3 and 4 deny those routes for the findings and accounts assessed here. The conclusion follows: these findings do not give the controller incremental support. This does not entail that the controller makes no actual difference or does not exist. Premise 2 is supported inductively by methodologically diverse interventions rather than one decisive experiment. Premise 3 is a comparative claim about what the two views can expect. Premise 4 is a current literature judgment and can be overturned by a controller-specific success or independent argument.

A.1 Premise Support and Limits

Premise 1 does not require that every cause be directly visible or independently adjustable. Historical and latent causes can be inferred from distinctive patterns and counterfactual relations. The requirement is that the proposed cause constrain the world somehow or have independent support. A controller that changes no expectation and has no independent case may still be part of a possible metaphysical interpretation, but the neural findings have not improved its evidential position.

Premise 2 is deliberately cumulative. Lesions probe necessity, stimulation probes manipulability, state perturbation probes dynamical capacity, connectivity changes probe integration, and temporal models probe evolving control. Each method has confounds and limits. Their methodologically diverse convergence supports organized neural causal dependence. It does not convert shared mediation evidence into several independent tests of ontology.

Premise 3 captures the pianist point. A controller can act through an instrument while every measured effect appears in the instrument. Mediation is therefore compatible with both views. Physical sufficiency—the stronger claim that total physical conditions fix the outcome—would rule out an additional difference, but the formal argument does not assume that global claim.

Premise 4 is not an argument from ignorance. It does not say that because no controller has been detected, none exists. It says that the assessed accounts have not supplied the distinctive work needed for these findings to favor the controller. A future theory can defeat the premise through an advance prediction, an independently motivated counterfactual pattern, a successful psychophysical law, or an independent argument for the relevant causal power.

The argument also distinguishes an *added substance* from a *higher-level cause*. A person's reason for leaving a room may explain the action even when neural and muscular processes realize that reason-guided behavior. The two descriptions need not compete for one slot, any more than a software instruction and the electrical activity implementing it must be rival pushes. The redundancy problem arises here because the puppeteer is proposed as a further subject that can act on the realizing process. If the proposal instead says only that one embodied process can be truly described in mental and neural terms, it has moved away from additive substance control.

Nor does a controller-specific difference have to look like a burst of mysterious energy. It could appear as a stable change in transition probabilities, a reason-sensitive counterfactual pattern, a systematic and predicted failure of an otherwise successful neural dependency, or a psychophysical rule—a rule connecting mental and physical states—that predicts when influence occurs. Pitts is therefore right that investigators should not assume in advance that interaction must violate a simple conservation equation. A raw residual is insufficient because incomplete models and measurement error also leave residuals. The present difficulty is evidential: the assessed theories do not yet motivate a distinctive pattern and then succeed where the no-controller model fails.

Finally, explanatory economy is not a proof of nonexistence. When two accounts make the same predictions, the account with fewer independent causes is ordinarily easier to test and use, but reality is not obligated to be simple. Instrument-mediated dualism may be true even if it is empirically tied with a brain-dependent rival. The formal conclusion is correspondingly modest: the shared evidence does not give the extra controller incremental causal-explanatory support. A separate argument may still establish it.

References

- Casali, Adenauer G., Olivia Gosseries, Mario Rosanova, Mélanie Boly, Simone Sarasso, Karina R. Casali, Silvia Casarotto, Marie-Aurélié Bruno, Steven Laureys, Giulio Tononi, and Marcello Massimini (2013). “A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior.” In: *Science Translational Medicine* 5.198, 198ra105. DOI: [10.1126/scitranslmed.3006294](https://doi.org/10.1126/scitranslmed.3006294).
- Chalmers, David J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York, NY: Oxford University Press.
- Fodor, Jerry A. (1974). “Special Sciences (or: The Disunity of Science as a Working Hypothesis).” In: *Synthese* 28.2, pp. 97–115. DOI: [10.1007/BF00485230](https://doi.org/10.1007/BF00485230).
- Haan, Edward H. F. de, Paul M. Corballis, Steven A. Hillyard, Carlo A. Marzi, Anil Seth, Victor A. F. Lamme, Lukas Volz, Mara Fabri, Elizabeth Schechter, Tim Bayne, Michael Corballis, and Yair Pinto (2020). “Split-Brain: What We Know Now and Why This Is Important for Understanding Consciousness.” In: *Neuropsychology Review* 30.2, pp. 224–233. DOI: [10.1007/s11065-020-09439-3](https://doi.org/10.1007/s11065-020-09439-3).
- Hasker, William (1999). *The Emergent Self*. Ithaca, NY: Cornell University Press.
- Holtzheimer, Paul E. et al. (2017). “Subcallosal Cingulate Deep Brain Stimulation for Treatment-Resistant Depression: A Multisite, Randomised, Sham-Controlled Trial.” In: *The Lancet Psychiatry* 4.11, pp. 839–849. DOI: [10.1016/S2215-0366\(17\)30371-1](https://doi.org/10.1016/S2215-0366(17)30371-1).
- Kim, Jaegwon (1998). *Mind in a Physical World: An Essay on the Mind–Body Problem and Mental Causation*. Cambridge, MA: MIT Press.
- Koenigs, Michael, Liane Young, Ralph Adolphs, Daniel Tranel, Fiery Cushman, Marc Hauser, and Antonio Damasio (2007). “Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgements.” In: *Nature* 446.7138, pp. 908–911. DOI: [10.1038/nature05631](https://doi.org/10.1038/nature05631).
- Lowe, E. J. (2008). *Personal Agency: The Metaphysics of Mind and Action*. Oxford: Oxford University Press. DOI: [10.1093/acprof:oso/9780199217144.001.0001](https://doi.org/10.1093/acprof:oso/9780199217144.001.0001).
- Mayberg, Helen S., Andres M. Lozano, Valerie Voon, Heather E. McNeely, David Seminowicz, Clement Hamani, Jason M. Schwalb, and Sidney H. Kennedy (2005). “Deep Brain Stimulation for Treatment-Resistant Depression.” In: *Neuron* 45.5, pp. 651–660. DOI: [10.1016/j.neuron.2005.02.014](https://doi.org/10.1016/j.neuron.2005.02.014).
- Moreland, J. P. and Scott B. Rae (2000). *Body and Soul: Human Nature and the Crisis in Ethics*. Downers Grove, IL: InterVarsity Press.
- Pinto, Yair, David A. Neville, Marte Otten, Paul M. Corballis, Victor A. F. Lamme, Edward H. F. de Haan, Nicoletta Foschi, and Mara Fabri (2017). “Split Brain: Divided Perception but Undivided Consciousness.” In: *Brain* 140.5, pp. 1231–1237. DOI: [10.1093/brain/aww358](https://doi.org/10.1093/brain/aww358).
- Pitts, J. Brian (2020). “Conservation Laws and the Philosophy of Mind: Opening the Black Box, Finding a Mirror.” In: *Philosophia* 48.2, pp. 673–707. DOI: [10.1007/s11406-019-00102-7](https://doi.org/10.1007/s11406-019-00102-7).
- Reddy, Sandesh, Katherine E. Kabotyanski, Samad Hirani, Tommy Liu, Zain Naqvi, Nisha Giridharan, Mohammed Hasen, Nicole R. Provenza, Garrett P. Banks, Sanjay J. Mathew, Wayne K. Goodman, and Sameer A. Sheth (2024). “Efficacy of Deep Brain Stimulation for Treatment-Resistant Depression: Systematic Review and Meta-Analysis.” In: *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 9.12, pp. 1239–1248. DOI: [10.1016/j.bpsc.2024.08.013](https://doi.org/10.1016/j.bpsc.2024.08.013).
- Schurger, Aaron, Pengbo Ben Hu, Joanna Pak, and Adina L. Roskies (2021). “What Is the Readiness Potential?” In: *Trends in Cognitive Sciences* 25.7, pp. 558–570. DOI: [10.1016/j.tics.2021.04.001](https://doi.org/10.1016/j.tics.2021.04.001).
- Schurger, Aaron, Jacobo D. Sitt, and Stanislas Dehaene (2012). “An Accumulator Model for Spontaneous Neural Activity Prior to Self-Initiated Movement.” In: *Proceedings of the National Academy of Sciences* 109.42, E2904–E2913. DOI: [10.1073/pnas.1210467109](https://doi.org/10.1073/pnas.1210467109).
- Searle, John R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Squire, Larry R. (2004). “Memory Systems of the Brain: A Brief History and Current Perspective.” In: *Neurobiology of Learning and Memory* 82.3, pp. 171–177. DOI: [10.1016/j.nlm.2004.06.005](https://doi.org/10.1016/j.nlm.2004.06.005).
- Swinburne, Richard (2013). *Mind, Brain, and Free Will*. Oxford: Oxford University Press. DOI: [10.1093/acprof:oso/9780199662562.001.0001](https://doi.org/10.1093/acprof:oso/9780199662562.001.0001).
- Woodward, James (2004). *Making Things Happen: A Theory of Causal Explanation*. New York, NY: Oxford University Press. DOI: [10.1093/0195155270.001.0001](https://doi.org/10.1093/0195155270.001.0001).