Synchrony Is Not Unity

A Causal Framework for Dyadic Functional Organization

Aurora Flores, PhD
International Academy of Consciousness Studies

Theory and Methods Article | Version 6 | July 2026

Abstract

Inter-brain synchrony is sometimes treated as evidence that two people have become one conscious system. That inference is not licensed by synchrony alone. Shared input, coordinated action, sensory exchange, autonomic entrainment, affect regulation, expectancy, and ordinary communication can all produce cross-person dependence while two subjects remain distinct. The scientifically tractable question is therefore not whether interacting people become correlated, but whether a task-relevant process is organized across them in a way that is not adequately captured by preregistered person-bounded and two-agent communication models. This article develops the Dyadic Functional Organization Test (DFOT), a graded causal and model-comparison framework for that question. The framework distinguishes alignment, interpersonal influence, closed-loop reciprocity, intervention-sensitive coupling, dyadic synergy, model-resistant cross-person organization, organizational individuation, and theory-conditional subjecthood. It also separates two inferential gaps that are often collapsed. An organizational bridge principle is required to infer a higher-order system from cross-person dependence; a subjecthood bridge principle is then required to infer a conscious subject from that system. The central empirical null is a rich two-agent message-passing model in which each participant retains an individual latent state, learns the partner, and updates through the measured channel. A joint latent model is evidentially relevant only if it generalizes prospectively, survives equal-complexity comparisons, carries task-specific information, responds selectively to intervention, and outperforms this null in new dyads. Cross-person partition dependence is treated as evidence about an internal decomposition, not as direct localization of an outer system boundary. A later organizational-individuation stage therefore compares candidate envelopes and tests persistence, recovery, endogenous error correction, cross-task recurrence, and relative autonomy. The proposed program combines computational benchmark systems, a live-versus-replay haptic design crossed with instructed belief, content-specific perturbation, complementary-evidence tasks, independent phenomenology, equivalence testing, and frozen replication. A positive result could establish a robust dyadic functional organization relative to a defined model class. It would not, without further theory, establish that the dyad is one subject.

Keywords: inter-brain synchrony; hyperscanning; dyadic cognition; social touch; causal intervention; informational synergy; system individuation; organizational closure; neurophenomenology; replication

> **Central thesis.** The DFOT can establish model-relative evidence that a task-relevant process is organized across two people in ways not captured by preregistered individual and message-passing models. It cannot by itself establish either a higher-order individual or a single conscious subject. Those conclusions require, respectively, an organizational bridge principle and a subjecthood bridge principle.

1. Introduction: The Inferential Problem

Consciousness research ordinarily begins with an individual organism and asks which internal processes distinguish conscious from unconscious states or contents. Social interaction complicates that starting point. In conversation, touch, joint action, and emotional exchange, each person continuously changes the sensory evidence, action policy, autonomic state, and expectations of the other. Hyperscanning has made some of these dependencies measurable across brains and bodies (Hasson et al., 2012; Czeszumski et al., 2020; Schilbach and Redcay, 2025). The existence of cross-person dependence is no longer in serious doubt. Its interpretation remains unsettled.

The strongest versions of the problem concern conscious-system individuation. Valencia and Froese (2020) argued that inter-brain synchronization may have implications for extended consciousness and proposed experimental manipulation as a way forward. More recent work has strengthened the case that shared neural dynamics can be functionally consequential. In interacting mice and artificial agents, Zhang et al. (2025) identified shared neural subspaces and showed that selectively disrupting shared components reduced social behavior. Such findings matter because they move beyond passive covariation. They still do not show that a pair is one conscious system. A shared dynamical component may organize interaction while each participant remains an individual subject.

The methodological literature supplies the same warning from another direction. Hyperscanning is intrinsically vulnerable to shared stimuli, coordinated movement, common task structure, physiological covariation, analytic flexibility, and the fact that two agents normally communicate (Burgess, 2013; Hamilton, 2021; Holroyd, 2022). Recording alone cannot establish causal efficacy; intervention, behavioral modeling, and explicit alternatives are required (Novembre and Iannetti, 2021; Moreau and Dumas, 2021). Recent methodological work further emphasizes simulation, confound control, and the need to define what inter-brain synchrony is supposed to explain (Cassioli et al., 2025). Even a causal effect of one brain on another, however, would establish influence rather than unity.

The central distinction is categorical. Synchrony is a relation between signals. Unity is a claim about how a system is individuated and, in its strongest form, about how many subjects exist. The same observed covariance can arise from a common driver, reciprocal prediction, sequential communication, shared affect, a joint task, or an autonomous interaction-level process. No magnitude of synchrony identifies which of these explanations is correct. Nor does a successful joint model automatically settle the matter. Joint models are often more flexible, tasks can manufacture complementarity, and any model comparison is relative to the variables, temporal grain, channels, interventions, and alternatives that were included.

This paper develops a stronger but deliberately limited target: **dyadic functional organization**. The question is whether a specified process is organized across two people in a way that prospectively predicts and causally controls task-relevant outcomes better than preregistered person-bounded and two-agent message-passing models. This target is stronger than coordination but weaker than a claim that a new individual or subject exists. It can be tested without pretending that model fit discovers metaphysical boundaries by inspection.

The resulting framework is the Dyadic Functional Organization Test (DFOT). The DFOT is not one biomarker or one statistic. It is an evidence ladder and an experimental architecture. It requires common-cause controls, live-versus-replay comparisons, causal perturbation, target-specific synergy, prospective validation, a strong two-agent communication null, computational false-positive benchmarks, and independent replication. It also treats phenomenology as an independent concordance axis rather than a direct count of subjects.

The paper makes four contributions. First, it separates five claims that are routinely conflated: influence, permeability, dyadic functional organization, organizational individuation, and subjecthood. Second, it identifies two bridge problems rather than one. Evidence of cross-person organization does not by itself identify a higher-order system, and evidence of a higher-order system does not by itself identify a conscious subject. Third, it replaces anatomical-cut rhetoric with a more precise distinction between internal partition dependence and outer-envelope individuation. Fourth, it makes the strongest alternative explanation—a pair of sophisticated agents exchanging messages through a measured channel—the central null rather than a late control.

The intended contribution to the field is therefore methodological and conceptual. The DFOT does not promise a shortcut to shared consciousness. It specifies what must be shown before claims about dyadic systemhood become scientifically serious, and it makes clear where empirical evidence ends and theory must begin.

2. What Must Be Distinguished

2.1 Five claims of increasing strength

Discussions of shared or extended consciousness often pass among different claims without marking the inferential transitions. Table 1 distinguishes the claims that the present framework keeps separate.

Table 1. Claims that must not be conflated

	Evidence that would be relevant
	Reliable directional effects beyond baseline and obvious common causes.
ant remains an identifiable center of regulation.	Live closed-loop effects beyond replay, delay, valence, expectancy, and movement.

	Evidence that would be relevant
and message-passing model family.	Prospective joint-model gain, content-specific intervention effects, target-specific synergy, cross-per

Evidence that would be relevant

Dyadic functional organization plus persistence, perturbation recovery, endogenous error correction,

	Evidence that would be relevant
	Organizational individuation plus a defended theory-specific bridge from organization to subjecthoo

Evidence that would be relevant

Evidence that would be relevant

The hierarchy blocks both inflation and deflation. Synchrony, physiological coupling, felt closeness, or an evocative report cannot be promoted directly to organizational individuation or subjecthood. Conversely, a replicated result at the level of dyadic functional organization would not be “only correlation.” It would be a substantive finding about how a task-relevant process is organized, even if no new subject existed.

2.2 The first bridge: from evidence to organization

Let **E** denote the strongest empirical pattern available from the DFOT: prospective joint-model advantage over rich person-bounded alternatives, target-specific synergy, intervention sensitivity, cross-person partition dependence, and replication. Let **B_O** denote an organizational bridge principle: a criterion specifying when evidence of this kind is sufficient for treating a process as a higher-order system. Let **O** denote the conclusion that a higher-order organization has been individuated.

$$E + B\_O \rightarrow O$$

The organizational bridge is not optional. Cross-person synergy can occur in a temporary coalition of two agents. A joint latent model can summarize interactions without corresponding to an autonomous entity. Severing the A|B cut can damage a task because the task depends on communication, not because A+B forms a new individual. The bridge principle must therefore specify what distinguishes a system from a relation among systems.

Candidate criteria include causal cohesion, operational closure, relative autonomy, endogenous regulation, persistence, and recovery after perturbation (Di Paolo, 2005; DiFrisco, 2018). No single criterion is assumed to be theory-neutral. The DFOT instead requires authors to state the criterion they are using, to test its distinctive consequences, and to identify what evidence would count against it.

2.3 The second bridge: from organization to subjecthood

Let **B_S** denote a subjecthood bridge principle: a theory's reason for treating organization O as a bearer of experience. Let **S** denote the conclusion that O is a conscious subject.

$$O + B\_S \rightarrow S$$

Neuroscience can constrain O. It cannot infer S while leaving B_S unstated. Integrated information theory, global neuronal workspace theory, predictive processing, enactivism, and organizational-closure accounts do not use the same bridge. IIT requires an intrinsic cause-effect analysis and exclusion among overlapping candidates; GNWT requires a suitable architecture of global availability; enactive accounts emphasize autonomy and sense-making; predictive approaches may describe coupled inference without adding a subject; Closure requires a substantive account of when a closure is conscious. A result can therefore support one theory-conditional claim while remaining irrelevant or contrary to another.

This second bridge is familiar in consciousness science, where measures are interpreted through theories of what makes a physical organization conscious. The first bridge is less often stated. The paper's central correction is that both are required:

cross-person dependence is not yet systemhood, and systemhood is not yet subjecthood.

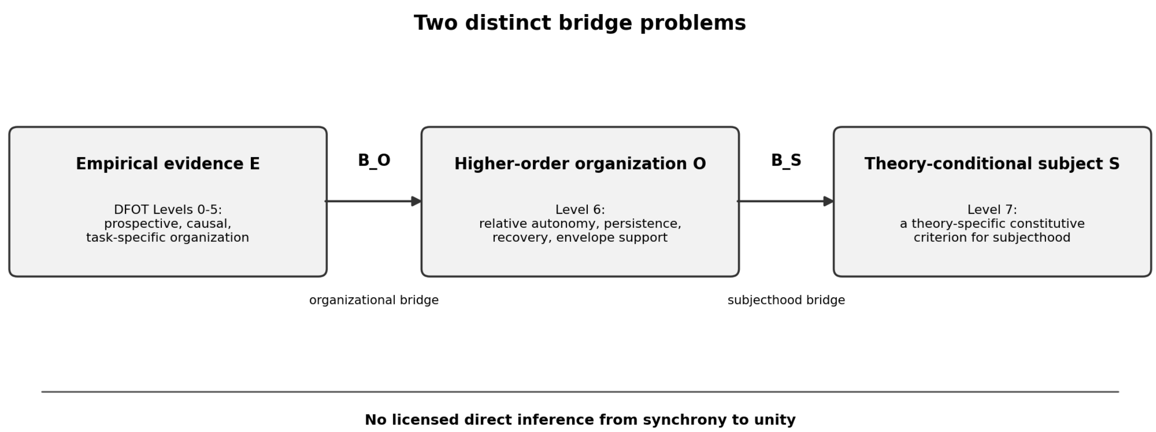

Figure 1. The two bridge problems. The DFOT supplies evidence E. Organizational individuation requires B_O; subjecthood requires B_S.

2.4 Phenomenal unity, subject unity, and social coordination

A "shared field" is not a univocal scientific term. A phenomenological field is the structured totality of what is experienced at a time. A dynamical field is a mathematical representation of distributed variables. A social field may refer to reciprocal affordances, norms, or interactional organization. A neural field refers to population activity. Evidence for one does not establish the others.

Reports of fusion are similarly multivalent. "We became one" may describe reduced defensiveness, loss of body salience, joint agency, temporal absorption, erotic intensity, source ambiguity, or the literal belief that one subject replaced two. The unity of consciousness is itself a contested philosophical problem (Bayne, 2010), and the combination of multiple subjects raises additional difficulties (Young et al., 2022). Phenomenology can therefore inform the analysis only if reports are decomposed into more specific features and interpreted independently of the neural result.

3. Why Touch Is a Useful Test Channel

3.1 A privileged channel is not a boundary criterion

Touch is selected for experimental tractability, not because it is assumed to merge minds. Affiliative touch is biologically specialized, temporally controllable, reciprocally generated, and measurable at the interface between two bodies. Unmyelinated C-tactile afferents in hairy skin respond preferentially to slow, caress-like stroking and covary with pleasantness at the group level (Loken et al., 2009; Ackerley et al., 2014). Touch can alter pain, interoception, salience, autonomic regulation, felt safety, and social valuation. These properties make it a strong channel through which to manipulate reciprocal coupling.

The specialization does not imply unity. Handholding is ecologically meaningful, but glabrous palm skin is not the canonical site for isolating CT-optimal stimulation. Controlled forearm stroking, handholding, and sustained contact should therefore remain distinct conditions. Oxytocin may modulate the pleasantness and neural processing of social touch, but its effects are context-sensitive and do not provide a switch from separate subjects to a shared one (Handlin et al., 2023; Chen et al., 2023).

3.2 What touch research already establishes

Partner touch can produce consequential cross-person coordination. Handholding during experimentally induced pain has been associated with analgesia and increased brain-to-brain coupling in romantic partners (Goldstein et al., 2018). Partner touch can alter respiratory and cardiac coupling during pain or stress (Goldstein et al., 2017). Naturalistic romantic interaction and partner touch have also been associated with enhanced interpersonal neural synchronization relative to comparison conditions (Kinreich et al., 2017; Long et al., 2021; Zhou et al., 2024). Attachment history can shape inter-brain synchrony toward social goals (Djalovski et al., 2021).

These effects are important but do not identify subject number. Felt closeness, liveness, synchrony, and functional integration must be measured separately. Fornari et al. (2026), for example, experimentally increased perceived interpersonal closeness without a corresponding increase in inter-brain synchrony, while social interaction itself increased synchrony. This dissociation is exactly what a disciplined framework should expect: psychologically meaningful constructs need not share one neural index.

3.3 Contact can be a channel or a constituent

A brain-only analysis may omit the relevant interactional organization. The ongoing contact process can be treated in two ways. A **channel-conditioned model** treats force, velocity, location, temperature, device state, and timing as measured transmission variables to be conditioned on. A **contact-inclusive model** treats those variables as nodes in an expanded A-K-B system, where K denotes the interface.

The two analyses answer different questions. If cross-brain dependence disappears after K is modeled, there is no residual cross-neural dependence beyond the measured channel. Yet the embodied interaction may still be organized at the A-K-B level. If a contact-inclusive model generalizes better than both person-bounded and channel-conditioned alternatives, the appropriate conclusion is that the functional organization includes the contact process. It is not “brain-to-brain fusion.”

This distinction also prevents a common mistake in extended-mind debates: conditioning away the very process that may be constitutive, then treating its disappearance as evidence that nothing extended existed. At the same time, simply adding K to the model does not establish a new system. The outer-envelope problem remains.

4. The Empirical Target: Model-Resistant Dyadic Organization

4.1 Variables and model relativity

Let X_A(t) and X_B(t) denote multivariate neural and embodied states of participants A and B. Let K(t) denote the measured interaction channel, including contact mechanics and device state. Let E(t) denote exogenous variables such as shared cues, trial timing, room conditions, and task events. Let Y(t) denote a preregistered task target such as a joint decision, later memory, error correction, or perturbation detection.

Any residual dependence is relative to what was measured and modeled. An unmeasured common cause, poor preprocessing, a weak individual model, the wrong temporal grain, or a badly specified channel can create apparent cross-person structure. The empirical target is therefore not an absolute remainder that no conceivable person-bounded account could ever explain. It is a **structured cross-person residual** relative to a preregistered and independently validated model family.

A residual becomes evidentially interesting only when it is:

1. **prospective**, improving prediction on later untouched blocks or sessions;
2. **target-specific**, carrying information about a defined joint function rather than generic covariance;
3. **intervention-sensitive**, changing in the predicted direction under controlled perturbation;

4. **control-specific**, exceeding replay, delay, common-input, movement, valence, expectancy, and responsiveness controls;
5. **estimator-convergent**, appearing under at least two non-equivalent analytic families;
6. **benchmark-calibrated**, not produced at an unacceptable rate by simulated systems known to contain only two message-passing agents; and
7. **replicated**, reproducing under a frozen protocol in new dyads.

This formulation is deliberately model-relative. The warranted claim is that the organization is not captured by the tested person-bounded and communication models, not that it is irreducible in principle.

4.2 The central null: two-agent message passing

The strongest ordinary explanation is not two isolated brains. It is two sophisticated agents that maintain individual latent states, model each other, exchange information through a channel, and update sequentially. The DFOT therefore treats a **Two-Agent Message-Passing model (TAM)** as the central null.

A simplified TAM factorization is:

$$p(z\_A,t+1, z\_B,t+1 \mid H\_t, K\_t, E\_t) = p(z\_A,t+1 \mid z\_A,t, m\_B\text{->}A,t, E\_t) \times p(z\_B,t+1 \mid z\_B,t, m\_A\text{->}B,t, E\_t)$$

Here z_A and z_B are hidden individual states, H_t is relevant history, and m_A->B and m_B->A are messages recoverable from the measured channel and each participant's actions. A realistic TAM may include nonlinear dynamics, individual memory, uncertainty, partner-specific priors, learned response policies, delays, and asymmetric influence. It may also allow each participant to infer the other's latent state. What it does not contain is a higher-order latent state that functions as an additional controller.

A **joint latent model** M_J introduces a dyadic state z_D whose dynamics contribute prospectively to both participants and the task target after the TAM variables are included. Because M_J is more expressive, its evidential value depends on matched capacity, regularization, simulation calibration, and untouched future data. Same-block fit is insufficient.

The primary comparison is therefore not joint versus isolated-person models. It is:

M_J versus the best validated TAM.

This change is decisive. Complementary information, sequential negotiation, and mutual prediction can all make a joint description useful while leaving the two-agent account intact. The joint model must explain something that the best communication model, not a straw-person independent model, systematically misses.

4.3 Synergy is target-relative

Synergy is assessed with respect to a defined target Y. The question is whether the joint state carries information about Y that is unavailable from either participant alone. Integrated information decomposition and related methods can distinguish synergistic, redundant, and unique contributions, but the result remains estimator- and target-dependent (Luppi et al., 2024; Mediano et al., 2025).

A complementary-evidence task can generate synergy by design. If A receives one clue and B receives another, the pair can solve a problem that neither can solve alone. That fact demonstrates distributed computation, not higher-order individuation. Synergy becomes stronger evidence only when it survives comparison with a TAM that explicitly represents the clue exchange, temporal order, partner model, and channel transcript. Transfer to a new task, persistence across a brief interruption, and dyad-specific recovery after perturbation provide additional leverage because they are not guaranteed by the initial task architecture.

4.4 Internal partitions are not outer boundaries

The original attraction of a boundary index is understandable: if severing the anatomical A|B partition damages prediction more than severing matched alternatives, perhaps the operative boundary has moved. The inference is too strong. A|B is an internal cut through the modeled process. Its importance does not identify the process's outer edge.

The DFOT therefore uses the more limited term **cross-person partition dependence**. For an estimator m, temporal scale tau, and candidate internal partition P, let SL_m,tau(P) denote the loss of prospective prediction or task control when dependencies across P are severed under matched model complexity. A cross-person dependence contrast can be written as:

$$CPD\_m,tau = SL\_m,tau(A|B) - median[SL\_m,tau(P\_alt)]$$

A positive value means that the anatomical person cut is especially important to the specified process relative to matched internal cuts. It does not mean that A+B is the true system boundary. The conclusion is about decomposition: the process cannot be cleanly factorized at A|B within the tested model.

Outer individuation requires comparison among **candidate envelopes**, not merely internal bipartitions. Candidate envelopes may include:

- A alone;
- B alone;
- A+B;
- A-K-B, including the contact process;
- A-K-B plus device-control states;
- A-K-B plus selected environmental structures required by the task.

The analysis must ask which envelope, if any, exhibits relative autonomy without improving indefinitely as more of the environment is added. No candidate envelope is assumed in advance to coincide with an anatomical or Markov-blanket boundary (Kirchhoff et al., 2018; Bruineberg et al., 2022). This is a local-maximality problem constrained by causal intervention, complexity, and the theory's individuation criterion.

4.5 The organizational bridge and the Closure interpretation

Closure can now be stated more restrictively. A closure is not any correlated or mutually influential set. Under the present proposal, a candidate higher-order closure must exhibit a temporally stabilized regime of reciprocally enabling constraints that maintains a distinguishable organization under perturbation while remaining materially and informationally open. For a dyad, the following properties are required:

- **reciprocal enablement:** processes on both sides help sustain the interaction-level organization rather than merely transmit independent outputs;
- **relative autonomy:** the candidate envelope has a stable causal profile relative to nearby smaller and larger envelopes;
- **endogenous regulation:** the interaction detects and corrects deviations in variables relevant to its own continuation or task norm;
- **persistence and recovery:** the organization survives brief interruption or reconstructs a dyad-specific state after perturbation;
- **cross-task recurrence:** a recognizable interaction-level organization reappears in more than the task that created the original complementarity; and
- **nestedness:** the individual organizations remain robust while the higher-order process constrains them.

These criteria give Closure work to do. A Level 5 DFOT result may establish model-resistant cross-person organization. It becomes a **candidate nested closure** only if the Level 6 organizational criteria are also met under an explicit organizational bridge B_O. This distinguishes Closure from a relabeling of residual model error.

The empirical residual may be described within Closure as a model-relative Remainder: organized structure left unresolved by the current decomposition. It should not be described as an absolute metaphysical remainder. New measurements and models may absorb it. What matters scientifically is whether the residual remains stable under progressively stronger alternatives and interventions.

Even a nested closure would not automatically be a subject. Closure still needs B_S: a principled account of which closures are conscious, whether nested conscious closures can overlap, and whether a higher-order closure adds a subject, replaces constituent subjects, or merely organizes them. The present paper does not assume that this final bridge has already been supplied.

5. The Dyadic Functional Organization Test

The DFOT is an evidence ladder. Each level adds a requirement and blocks a distinct class of alternatives. Higher levels presuppose the lower validity gates but do not erase the distinctions among claims.

Table 2. DFOT evidence ladder

Required finding	What it supports
Cross-person measures do not exceed pseudo-dyads, surrogates, or matched baselines after common causes are modeled.	No detectable interpersonal effect for the tested conditio

Required finding	What it supports
	n and sensitivity.
Cross-person similarity occurs during common stimulation or coordinated behavior.	Shared input or coordinated response.

Required finding	What it supports
One participant's prior state predicts changes in the other beyond baseline and measured common causes.	Interpersonal influence.

Required finding	What it supports
Live interaction exceeds kinematically matched replay and delay after valence, expectancy, and movement are addressed.	A role for current mutual updating.

Required finding	What it supports
A perturbation on one side produces predicted, content-specific effects through the full loop and differs from yoked controls.	Causal organization of the interaction loo

Required finding	What it supports
	p.
The joint state predicts a preregistered target beyond either participant alone and beyond the observable transcript on future data.	Distributed task computation and a

Required finding	What it supports
	candidate dyadic process.
A joint latent model prospectively exceeds the validated TAM; cross-person partition dependence, intervention effects, estimator convergence, and replication agree.	A robust dyadic func

Required finding	What it supports
	tional organization relative to the tested model class.
Level 5 plus persistence, recovery, endogenous error correction, cross-task recurrence, and candidate-envelope support under B_O.	A ca

Required finding	What it supports
	ndidate higher-order system or nested closure under the stat

Required finding	What it supports
	ed organizational theory.
Level 6 plus a defended B_S and the theory-specific constitutive test.	Evidence for a dyadic subject unde

Required finding	What it supports
	r theory T.

Levels 0 through 3 concern increasingly strong forms of coupling. Level 4 establishes target-specific distributed computation. Level 5 is the strongest theory-light functional conclusion. Level 6 is not produced by a larger Level 5 effect; it requires qualitatively new evidence about autonomy and individuation. Level 7 remains theory-conditional.

The ladder also clarifies the role of negative evidence. A local null at Level 2 does not refute all interpersonal influence. A Level 4 null does not refute co-regulation. A valid Level 5 null weakens the claim that the tested task generates a model-resistant dyadic organization at the prespecified effect size. A Level 6 null may show that strong joint computation occurs without higher-order individuation.

6. Core Experimental Program

A single experiment cannot maximize ecological intimacy, causal control, signal quality, participant blinding, and organizational scope. The program is therefore staged. The first stages establish measurement validity and causal dyadic organization. The later stage asks the distinct individuation question.

6.1 Stage 0A: Computational benchmark systems

Before human data are used to support a boundary claim, the complete analysis pipeline should be tested on systems with known generating architectures. This is not a conventional power simulation. It is a false-positive and construct-validity benchmark.

Table 3. Minimum computational benchmark suite

The benchmark suite should vary channel bandwidth, delay, noise, hidden-state dimension, task complementarity, model misspecification, and sample size. The preregistered pipeline must achieve acceptable false-positive rates in null architectures and adequate sensitivity in positive controls before human Level 5 or Level 6 claims are permitted. A method that labels sophisticated message passing as a higher-order system has failed its central validation test.

6.2 Stage 0B: Engineering and independent calibration

A separate human calibration cohort establishes device latency, replay fidelity, touch-force ranges, artifact thresholds, safe perturbation magnitudes, pleasantness matching, endpoint reliability, and the temporal hierarchy used later. Confirmatory outcomes are not inspected in this cohort.

The bilateral haptic interface should record and reproduce force, velocity, contact location, temperature, and timing. Removable liveness cues—actuator noise, visual leakage, latency spikes, experimenter timing, and replay discontinuities—must be eliminated. If participants continue to discriminate live interaction after engineering cues are removed, that fact is documented. Perfect blinding is not required, but instructed belief can no longer be presented as a clean masking manipulation.

Temporal claims are also calibrated independently. The calibration identifies at least three intervals: initial channel transmission, the earliest interval in which perturbation detection or task availability can be decoded out of sample, and the earliest physically possible A-to-B-to-A loop. Exact thresholds, bin widths, and reliability gates belong in the pre-analysis plan and must be justified by calibration performance rather than selected from the confirmatory result. If no reliable access-relevant interval can be defined independently, the later study may test causal coupling but may not make conscious-access claims.

6.3 Stage 1: Actual reciprocity crossed with instructed belief

Established close partners interact from separate rooms through the bilateral haptic interface. The primary factorial design crosses **actual contingency** with **instructed belief**. The device is either live closed-loop or a kinematically matched replay, and participants are told either that the block is live or that it is replayed. The design requires ethics approval for deception, careful debriefing, and any post-debrief withdrawal option required by the governing review board.

The main effect of actual contingency tests whether current mutual updating matters. The main effect of instruction estimates assigned expectancy. Their interaction tests whether actual reciprocity operates differently when participants believe the partner is currently present. The primary mechanistic estimand remains actual contingency; belief-related effects cannot repair a null live-versus-replay result.

After each block, participants give a live/replay judgment, confidence, and the cues used. Accuracy, Brier score, calibration, and metacognitive efficiency are reported (Maniscalco and Lau, 2012). Confidently incorrect judgments are especially informative. If a prospective dyadic effect follows actual contingency among participants who confidently misidentify the condition, simple expectancy becomes less plausible. If the effect follows confident belief regardless of actual condition, expectancy becomes a stronger explanation. Because inferred liveness is post-treatment, it is not entered as a naive covariate in the primary model (Montgomery et al., 2018).

The primary comparison is live established-partner interaction versus established-partner replay. Secondary mechanism modules may include delayed live interaction, a generic adaptive controller, a partner-trained controller, and a live unfamiliar responder. These controls decompose current responsiveness, partner history, human agency, and established-partner identity. They explain a positive primary effect; they do not rescue a null one.

6.4 Stage 2: Content-specific perturbation

Small, safe perturbations are introduced into timing, force, gain, or a task-relevant tactile code on one side of the loop. The immediate question is whether the perturbation propagates bidirectionally through the full loop and changes the estimated interaction state. An early evoked response may reflect transmission, reflexive control, or artifact. Evidence for recurrent dyadic organization begins only after a complete measured A-to-B-to-A loop is physically possible.

Content specificity is critical. Task-relevant and task-irrelevant perturbations should be matched in force, timing, novelty, and arousal potential. On relevant trials, the perturbation changes an evidence token that should shift a later decision in a prespecified direction. On irrelevant trials, the same physical event is orthogonal to the decision. A general arousal account predicts similar physiological consequences in both trial types. A content-bearing dyadic process predicts selective changes in decision, confidence, correction, or later memory when the perturbation changes task information.

Yoked solo and replay controls receive the same local tactile events without current mutual contingency. Level 3 requires a directionally specific downstream effect that survives motion and autonomic modeling and differs from those controls. Claims about conscious availability additionally require covariation with detection, confidence, report, memory, or flexible use, not merely late latency.

6.5 Stage 3: Distributed evidence and the message-passing null

The central task gives each participant complementary evidence required for a joint decision. Neither has enough information to solve the task alone. Overt speech is prohibited during the critical interval, and communication is restricted to the haptic loop. The task is compared with replay, delay, and a matched sequential-message condition.

This design creates a strong opportunity for synergy but also a strong danger of overinterpretation. The confirmatory null is the validated TAM. It includes each participant's hidden belief state, uncertainty, message history, partner model, timing, and the observable channel transcript. The joint model must predict future accuracy, confidence calibration, error correction, and memory beyond this full alternative.

The most informative intervention targets the candidate joint state rather than simply degrading one participant's signal. A perturbation selected from the joint model should impair task-relevant availability in both participants more than a matched perturbation selected from either individual model. The analysis should test whether the joint-state intervention effect transfers to later blocks and whether the same dyad-specific latent geometry recurs across sessions.

Phenomenology is tied to the task. After each decision, participants privately report whether the choice felt individually reached, sequentially negotiated, jointly formed, or source-indeterminate. They also report which evidence felt directly available and which was inferred. Agreement between prospective synergy and a joint-decision phenomenology is diagnostically interesting; dissociation is equally informative.

6.6 Stage 4: Organizational individuation

Stage 4 addresses the question that Levels 0-5 leave open: whether the dyadic process is a relatively autonomous organization rather than a temporary relation between two agents. It introduces tests that are not guaranteed by complementary evidence or continuous communication.

Brief interruption. Once a candidate dyadic state is established, the channel is suspended for a short, preregistered interval. The analysis asks whether a dyad-specific state persists, whether both participants predict the same latent continuation, and whether performance resumes from that state rather than restarting as two independent agents.

Perturbation recovery. The interaction is displaced from its learned regime through a safe timing or gain perturbation. A candidate organization should exhibit a reproducible recovery trajectory. The critical question is not whether each person compensates, but whether the pair returns to a dyad-specific attractor or error-corrective policy that the TAM does not predict.

Cross-task recurrence. The original complementary-evidence task is replaced by a structurally different task with a new stimulus set and response mapping. A higher-order organization should recur in a recognizable form without being identical to the trained solution. Task-specific synergy that disappears completely under transfer remains Level 4 or 5 evidence, not Level 6.

Endogenous norm and error correction. The dyad is given a task in which the relevant norm must be maintained across fluctuating conditions. The analysis tests whether errors are detected and corrected at the interaction level, including cases in which neither participant alone has enough information to identify the deviation.

Candidate-envelope comparison. Models compare A, B, A+B, A-K-B, and justified larger envelopes. The preferred envelope must show a locally maximal profile of predictive closure, perturbation recovery, and causal cohesion under matched complexity. Simply improving prediction by adding more environmental variables is not sufficient.

No single score should collapse these properties until a composite has been validated in the computational benchmark suite. Stage 4 produces an **organizational individuation profile**, not a magic boundary number. Level 6 requires convergent support across several criteria and explicit defense of B_O.

6.7 Ecological generalization

The core mechanistic program uses controlled haptic interaction because it provides measurable channels and interventions. Ecological generalization can later compare handholding, CT-targeted forearm stroking, sustained embrace, participant-selected nonsexual contact, and pleasant nonsocial regulation. Each contrast should state a distinct mechanistic prediction rather than treating intimacy as one linear dose.

Sexual intimacy and pharmacologically altered states are separate scope questions. They introduce arousal, contact area, endocrine change, movement, relationship meaning, privacy, and safety at once. They should not be added after lower-tier nulls as an unregistered rescue. Any later protocol must make a qualitatively new prediction, use matched controls, and undergo independent ethical and power review.

7. Measurement and Analysis

7.1 Multimodal acquisition

No single modality can carry the inference. Dual high-density EEG provides temporal precision but is vulnerable to reference choice, muscle activity, movement, and volume conduction. fNIRS offers complementary hemodynamic information but is sensitive to systemic physiology. Contact force, velocity, location, temperature, motion capture, respiration, ECG, electrodermal activity, gaze, pupil size, and task timing are not optional nuisance variables. They are candidate common causes, transmission channels, and—in contact-inclusive models—candidate constituents.

Analyses should include real dyads, pseudo-dyads, partner permutations, time-shifted surrogates, and negative-control regions or frequencies where scientifically justified. A result that appears only under one reference, filter, node set, artifact threshold, or treatment of K cannot support Level 5. The current literature's methodological cautions make estimator robustness and simulation-based validation indispensable (Hamilton, 2021; Holroyd, 2022; Cassioli et al., 2025).

7.2 Model hierarchy

The analysis proceeds through a prespecified hierarchy rather than beginning with a favored joint model.

Table 4. Confirmatory model hierarchy

Model capacity must be matched or penalized, and all hyperparameters must be tuned within earlier training data. Randomly mixed within-block folds are secondary because they permit temporal leakage and reward static fit. The confirmatory score uses forward-chaining across blocks and, when feasible, cross-session prediction.

7.3 Primary functional estimand

The primary Stage 3 estimand is **Prospective Joint-over-TAM Gain (P-JTG)**. For each dyad, the joint and TAM models are trained on designated earlier blocks and evaluated on later untouched blocks. Let Score denote held-out log predictive density or another preregistered proper scoring rule. Then:

$$\text{P-JTG}_i = [\text{Score}(M_J) - \text{Score}(M_{TAM})]_{live} - [\text{Score}(M_J) - \text{Score}(M_{TAM})]_{replay}$$

The live-minus-replay contrast asks whether current reciprocity creates a joint-model advantage beyond the advantage obtainable from replayed structure. The dyad is the confirmatory unit. Valid folds receive equal weight so that a condition with more artifact-free samples does not dominate by construction. Cross-session transfer is a key secondary endpoint.

A positive P-JTG is necessary but not sufficient for Level 5. It must survive benchmark calibration, target-specific perturbation, alternative estimators, channel-conditioned and contact-inclusive analyses, and replication. The size of P-JTG should not be interpreted as a scalar measure of "how merged" the dyad is.

7.4 Cross-person partition dependence

The node system and matched-cut generator are frozen before confirmatory outcomes are inspected. Alternative internal cuts should be matched to A|B on module-size balance, modality composition, cut-edge capacity, lag distribution, and model degrees of freedom. Cuts that cannot be matched are declared inadmissible rather than replaced after seeing the result.

Cross-person partition dependence requires the A|B severing loss to exceed the matched-cut distribution in at least two non-equivalent estimator families, such as intervention-based state-space severing and target-specific information decomposition. Disagreement among estimators is reported as underdetermination. The result is described as dependence on cross-person connections within the modeled process, not as direct evidence that the outer boundary lies around the pair.

7.5 Organizational individuation profile

Level 6 is evaluated through a profile rather than one index. The prespecified profile includes:

- persistence during brief channel interruption;
- recovery trajectory after perturbation;
- endogenous detection and correction of interaction-level errors;
- recurrence across tasks and sessions;
- local maximality of a candidate envelope relative to smaller and larger alternatives;
- stability of individual organizations within the higher-order regime; and
- failure of the validated TAM to reproduce the full profile in simulation and human data.

The criteria should be scored separately and reported with uncertainty. A composite may be exploratory, but no threshold should be treated as a system detector until it is validated against the benchmark architectures.

7.6 Diagnostic phenomenology

Phenomenology is collected privately after each condition, with free description before structured questions. Numerical subject unity may have no uniquely recognizable experiential signature, so introspection is not asked to count subjects. Instead, condition-blinded coders identify five dimensions: spatial boundary salience, agency structure, sensory-access claims, temporal alteration, and field-structure language. Coders also distinguish literal, metaphorical, uncertain, and explicitly inferential use. The procedure draws on disciplined second-person and descriptive-experience methods while retaining standardized prompts, coder blinding, and reliability gates (Petitmengin, 2006; Hurlburt and Akhter, 2006).

The structured phase presents four scenarios without using the word "merger":

- **A.** I experienced my own sensation or evidence and inferred my partner's response from behavior.
- **B.** I experienced my own sensation or evidence, and my partner's state seemed separate but unusually vivid.
- **C.** Sensations, evidence, or agency seemed to arise from both of us without a clear source boundary.
- **D.** I experienced one continuous field of sensation or decision that did not map cleanly onto two bodies.

Participants rate every scenario, choose the closest, and report confidence. They also classify the decision as individual, sequentially negotiated, jointly formed, or source-indeterminate. The full distribution is reported; rare D responses are not selectively elevated.

Table 5. Phenomenological grades as an independent concordance axis

A dyad does not receive a synthetic “merger score” by averaging incompatible partner reports. Each participant remains a reporting subject. Dyad-level concordance is classified as concordant, asymmetric, or unavailable. P2 or P3 without Level 4 or 5 evidence supports altered self-modeling or single-field phenomenology without demonstrated dyadic organization. Level 5 with P0 supports functional organization without reported experiential alignment. These dissociations are results, not failures to be hidden.

7.7 Awareness, demand, and heterogeneous effects

Condition awareness is measured through liveness judgment, confidence, calibration, and cue reports. Hypothesis awareness is assessed at the end with open questions before explicit probes. The randomized instruction factor supplies the direct expectancy test; post-block belief is a possible mediator, not a baseline nuisance. Demand characteristics are therefore quantified without treating suspicion as automatic exclusion (Orne, 1962; Mummolo and Peterson, 2019).

The primary analysis estimates the population-average dyad effect, but the program should report between-dyad variance. A small set of moderators—relationship duration, ordinary touch familiarity, attachment-related avoidance and anxiety, and baseline partner-prediction accuracy—may be preregistered. Moderator claims require multiplicity control and replication. Post hoc responder classes cannot rescue a failed primary claim.

8. Falsification, Power, and Replication

8.1 What would falsify the functional claim

The Level 5 claim is weakened when the validated TAM matches the joint model on untouched future data, when live and replay are equivalent at the preregistered smallest effect of interest, when content-specific perturbations do not produce predicted downstream effects, when cross-person partition dependence fails across estimators, or when the discovery result does not replicate. A nonsignificant result alone is not evidence of equivalence. Local-null claims require equivalence testing with a prespecified bound (Lakens et al., 2018).

The Level 6 claim is independently weakened when a Level 5 process has no persistence, no dyad-specific recovery, no cross-task recurrence, no endogenous error correction, or no stable candidate envelope. This result would be theoretically informative: a pair can exhibit strong distributed computation without becoming a higher-order individual.

The framework becomes unfalsifiable if every null is answered by adding a more intimate condition, a more flexible model, a new temporal window, or a post hoc subgroup. New stages require new mechanisms and registration. Failed calibration gates block the corresponding claims rather than inviting substitute analyses.

8.2 Sample planning

The dyad is the confirmatory unit. Sample size should be set by a device-specific Monte Carlo using Stage 0 estimates of endpoint reliability, block variance, cross-condition correlation, missingness, liveness discrimination, and dyad heterogeneity. The target may increase after calibration but may not be reduced because the pilot effect is favorable.

For orientation, a paired standardized effect of $d_z = .30$, two-sided alpha = .05, and 90% power requires approximately 119 analyzable dyads. A conservative planning floor of 150 analyzable dyads per cohort is reasonable before reliability and attrition inflation, but it is not a universal requirement. If the primary endpoint reliability is .80 and $d = .30$ is defined before attenuation, the required analyzable sample rises to approximately 148. The final pre-analysis plan must use the empirical calibration rather than these illustrative values.

The smallest effect of interest is a claim threshold, not a declaration that smaller effects are unreal. An effect too small to distinguish reliably from the TAM or to reproduce across sessions may still matter for social regulation while failing to support Level 5.

8.3 Discovery and frozen replication

Discovery and replication should be registered before discovery outcomes are inspected. The replication evaluates the same claim with the same hardware logic, task structure, model family, partitions, temporal rules, and code environment. Winner's-curse attenuation is expected, so numerical equality is not required. A replication succeeds when the effect is in the prespecified direction, passes its independent statistical criterion, and the pooled estimate remains incompatible with zero or the registered null region.

The acquisition and analysis environment should be archived: device configuration, firmware, sampling rates, montage, synchronization protocol, preprocessing container, package versions, random seeds, node definitions, artifact thresholds, hyperparameters, and executable workflow. Updated methods may be applied only as parallel secondary analyses to both cohorts. If the original hardware cannot be functionally reproduced, the new study is a conceptual rather than exact replication (Nosek et al., 2018; Nosek and Errington, 2020).

9. Result Patterns and Warranted Conclusions

Table 6. Interpretive matrix

	Warranted conclusion
s after common causes and movement are modeled.	Alignment under shared conditions.
roller is equivalent to live.	Current responsiveness may be sufficient.

	Warranted conclusion
familiar controls, but the TAM predicts future task performance as well as the joint model.	Partner-specific reciprocal coordination within two-

	Warranted conclusion
e result is confined to the complementary-evidence task and has no persistence or transfer.	Level 5 dyadic functional organization for the task,
ain robust, and a candidate envelope shows persistence, recovery, error correction, and cross-task recurrence.	Level 6 candidate higher-order organization under t

Warranted conclusion
Functional organization without reported experienti

	Warranted conclusion
lows instructed belief.	Single-field report or altered self-modeling with de
y and task-specific cross-person availability.	Strong theory-relevant convergence.

Warranted conclusion
Claim-level confidence decreases; any explanation

	Warranted conclusion

The matrix is deliberately asymmetric. Strong functional evidence can exist without single-field reports, and single-field reports can exist without strong functional evidence. Neither axis automatically validates the other.

10. Consequences for Major Theoretical Frameworks

A DFOT result matters to a theory only when it bears on that theory's constitutive commitments.

Table 7. Theory-conditional interpretation

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
Closure	Level 6: model-resistant organiz	Level 5 without autonomy, or sta	Candidate nested closure. Subjec

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	ation plus reciprocal enablement,	ble person-bounded organization a	thood still requires a Closure-sp

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	relative autonomy, recovery, cro	cross valid rich tasks.	ecific B_S.

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	ss-task recurrence, and a locally		

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	supported envelope.		
Enactivism /	Interaction-	Rich live in	Constitutive

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
participatory sense-making	level dynamics become autonomous,	teraction is fully reconstructed	interaction processes; phenomenal

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	metastable, norm-governed, and c	by individual plans and message p	merger is not implied.

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	ausally constrain both participan	assing, with no interaction-level	

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	ts.	autonomy.	
Predictive processing	Live partner input imp	Partner priors or gene	Coupled inference or n

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
/ **active inference**	roves coupled model evidence and	ric responsiveness fully reproduc	ested generative process; subject

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	joint control beyond replay, gene	e the effect.	number remains open.

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	ric responsiveness, partner histo		

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	ry, and the TAM.		
Global neuronal	Information orig	Every successful	Pressure for a d

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
workspace	inating in either participant bec	trial decomposes into two tempor	istributed-workspace implementati

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	omes flexibly available for memor	ally ordered individual ignitions	on, not automatic subject unity.

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	y, correction, action, and report	and channel exchange.	

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	across both, and no individual w		

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	orkspace plus message passing rec		

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	onstructs the function.		
Integrate	An IIT-va	DFOT Leve	DFOT evid

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
d information theory	lid intrinsic analysis identifies	l 5 or 6 occurs while IIT selects	ence alone is extrinsic and canno

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	a dyadic complex that exceeds al	two individual complexes.	t establish an IIT subject.

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	l overlapping individual, supra-d		

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation
	yadic, and alternative-grain cand		

Framework	Result supporting a dyadic extension	Result creating tension for that extension	Maximum warranted interpretation

idates.

10.1 Closure

The revised framework gives Closure a narrower and more defensible role. Level 5 is not automatically a nested closure. Level 6 is the relevant threshold because it adds autonomy, persistence, recovery, and envelope comparison. This makes Closure empirically vulnerable. A process that is synergistic only while the task forces continuous exchange remains a dyadic functional organization but does not satisfy the proposed organizational bridge.

Closure also faces a separate subjecthood obligation. A general principle must explain why some closures are conscious and how nested closures relate to constituent subjects. Phenomenological concordance may strengthen a proposed bridge, but it cannot substitute for one.

10.2 Enactivism and participatory sense-making

Enactive approaches already allow interaction processes to acquire autonomy and constrain participants (De Jaegher and Di Paolo, 2007). The DFOT operationalizes the stronger version of that claim through perturbation, recovery, endogenous norm maintenance, and cross-task recurrence. A positive Level 6 result would support interaction-level constitution. It would not require that the interaction itself be a subject. Participatory sense-making is compatible with two subjects jointly sustaining an autonomous process.

10.3 Predictive processing and active inference

Predictive approaches naturally explain interpersonal synchronization through mutual modeling and perception-action loops. The generic-adaptive, partner-trained, unfamiliar-human, and live-partner comparisons decompose responsiveness, partner priors, human agency, and current identity. A live-partner advantage beyond the validated TAM would support a coupled generative process. Yet coupled inference may remain a relation between two models rather than one model with one point of view.

10.4 Global neuronal workspace

GNWT links conscious access to nonlinear amplification and broad availability (Mashour et al., 2020). A dyadic extension would require more than messages triggering separate individual workspaces. Information would need to become flexibly available across both participants in a way that neither individual workspace plus the channel reconstructs. The distributed-evidence task, intervention on the candidate joint state, and cross-person memory and correction tests target this possibility. Serial individual ignition supports the two-workspace null.

10.5 Integrated information theory

IIT identifies consciousness with intrinsic cause-effect structure and uses exclusion to select definite complexes (Albantakis et al., 2023). Predictive gain, synergy, and cross-person partition dependence are extrinsic analyses. They do not compute IIT's intrinsic quantities. Recent work on intrinsic units further emphasizes that causal grain and system composition require comparison among candidate units and overlapping systems within the IIT formalism (Marshall et al., 2026). A DFOT Level 6 result could motivate an IIT analysis; it could not substitute for one.

11. Ethics, Consent, and Dyadic Privacy

Touch research can become coercive, culturally narrow, or clinically unsafe. Consent must be granular and private for contact site, clothing state, duration, device, recording modality, deception, and the possibility of replay. Recruitment should not presume heterosexuality, binary gender roles, monogamy, or a universal meaning of touch. Screening and debriefing should address trauma history, touch aversion, pain, interpersonal conflict, coercive control, and unwanted emotional disclosure.

Either participant must be able to stop a block or the session privately without explanation and without loss of compensation. The partner should receive a neutral statement that the session has ended under a private study rule; staff should not reveal who stopped it or why. A technical-failure fiction should not be used if it is untrue.

Dyadic data create special privacy risks. One person's neural or physiological record may reveal the other's speech timing, affect, relationship state, or clinically relevant information. Identifiable or relinkable paired data should be governed at the dyad level, with withdrawal and secondary-use rules specified before enrollment.

Deception in the actual-by-belief design must be necessary, minimized, and followed by structured debriefing. Investigators should assess distress and relationship effects and offer any withdrawal option required by the ethics board. Higher-intimacy or altered-state protocols require independent review and cannot be treated as routine extensions of the core haptic study.

12. Discussion

The field does not need another claim that interacting brains synchronize. It needs a disciplined account of what stronger evidence would mean. The DFOT provides that account by shifting the target from signal similarity to causal, prospective, task-relevant organization and by making the best two-agent communication model the principal null.

The framework's first contribution is inferential separation. Content influence, permeability, dyadic functional organization, organizational individuation, and subjecthood are not degrees of one variable. They are different claims with different evidence. This matters because researchers can otherwise move from “the partners synchronized” to “the dyad integrated” to “the dyad became one conscious system” without supplying the missing premises.

The second contribution is the two-bridge architecture. Even ideal Level 5 evidence leaves the organizational question open. A drone jointly controlled by two operators, a duet, or a complementary-evidence team can exhibit reciprocity, synergy, intervention sensitivity, and cross-person partition dependence. These properties may identify a real distributed process while remaining compatible with two agents exchanging messages. Organizational individuation requires additional evidence about autonomy, persistence, recovery, norm maintenance, and outer-envelope selection. Only after that issue is addressed does the subjecthood bridge arise.

The third contribution is the centrality of computational benchmark systems. A method for discovering higher-order organization should be tested on architectures where the answer is known. It should not classify a sophisticated two-agent controller as a new system simply because the investigators omitted hidden partner models or used a weak individual baseline. Conversely, it should recover a genuinely centralized or interaction-constituted controller even when its implementation spans physically separate modules. Simulation is therefore not only a power tool; it is a test of whether the operationalization tracks the intended distinction.

The fourth contribution is the correction from boundary displacement to partition dependence. A damaging A|B cut shows that cross-person links are important to the modeled process. It does not locate the outer boundary of that process. Candidate-envelope comparison is needed, and even then the result is relative to an organizational criterion. This correction reduces the metaphysical load carried by a model-comparison statistic.

The fifth contribution is a more disciplined role for phenomenology. Reports of one field are not discarded, but neither are they treated as direct observations of numerical identity. The protocol separates boundary salience, agency, sensory access, time, and field structure; preserves contradictions and partner asymmetry; and allows functional and phenomenological evidence to dissociate. This is preferable to using evocative quotations as confirmation after the neural analysis.

Several limitations remain. Human neural measurement samples only part of the relevant organization. Latent common causes cannot be eliminated completely. A sufficiently flexible TAM may be difficult to identify, while a weak TAM can manufacture joint-model advantage. Tasks that distribute complementary evidence create the very synergy they later measure. Brief persistence may reflect individual working memory rather than a dyadic state. Recovery trajectories may be jointly learned but still implemented by two controllers. Candidate-envelope analyses face the risk that adding environmental variables continuously improves fit. Phenomenological reports may not distinguish one token experience from two highly matched experiences. No ethically realistic human experiment can intervene on every variable required for complete causal identification.

These limitations restrict the conclusion; they do not make the program empty. A replicated Level 5 result would establish that the chosen two-agent model family is insufficient and that a stable, content-bearing, causally relevant organization spans the pair. A Level 6 result would add evidence that this organization behaves as a relatively autonomous system under a stated individuation criterion. Both would advance social neuroscience, distributed cognition, and theories of interactive organization even if no new subject existed.

The framework also changes how negative results should be understood. Partners may become deeply coordinated, regulated, and phenomenologically absorbed while the validated TAM remains sufficient and the dyadic process lacks persistence. That outcome would support rich interpersonal permeability without higher-order individuation. A theory that can represent this outcome without relabeling it as partial success is preferable to a merger hypothesis that cannot lose.

The most important practical consequence is restraint in language. “Shared neural dynamics,” “dyadic functional organization,” “candidate higher-order system,” and “dyadic subject” should not be used interchangeably. Each term should identify the evidential level actually reached. The field gains precision when ambitious hypotheses are decomposed rather than rhetorically weakened or prematurely declared confirmed.

13. Conclusion

Synchrony is not unity. Causal coupling is not yet dyadic organization. Synergy is not yet systemhood. Systemhood is not yet subjecthood.

The DFOT provides a route through these distinctions. It begins with common-cause control and live reciprocity, advances through content-specific intervention and prospective synergy, and tests the result against a rich two-agent message-passing model. It treats cross-person partition dependence as evidence about decomposition rather than a direct boundary detector. It then asks a separate organizational question through persistence, recovery, endogenous error correction, cross-task recurrence, and candidate-envelope comparison. Only after those steps can a theory-specific subjecthood bridge be evaluated.

The strongest defensible empirical conclusion is therefore conditional but substantial: a specified task may be organized across two people in a reproducible, intervention-sensitive way that is not captured by the preregistered person-bounded and communication models. That would be an important discovery. It would not, by itself, show that two subjects became one.

The field should demand no less and claim no more.

Appendix A. Minimum Pre-Analysis Plan

The confirmatory pre-analysis plan should be time-stamped before recruitment and should include the following elements.

Acquisition. Hardware models and configuration; firmware; montage; sampling rates; synchronization procedure; impedance limits; room layout; haptic calibration; contact-force limits; perturbation safety limits; and replay-construction rules.

Preprocessing. Filters; reference; bad-channel criteria; artifact thresholds; ICA policy; muscle and motion treatment; fNIRS systemic correction where applicable; interpolation; epoching; and missingness rules.

Nodes and channels. Exact neural, bodily, contact, task, and common-cause nodes; the channel-conditioned and contact-inclusive graphs; and the candidate-envelope set.

Models. M0-M4 definitions; state-space architecture; priors; regularization; hyperparameter search; TAM partner-model capacity; joint latent dimension; target variables; and proper scoring rule.

Validation. Forward-chaining folds; cross-session tests; simulation benchmark acceptance gates; model-complexity matching; and rules for declaring the TAM inadequate.

Partitions. Internal-cut generator; matching tolerances; rejection rules; severing operation; estimator-convergence rule; and complete disclosure set.

Organizational profile. Interruption duration; persistence outcome; perturbation-recovery metric; cross-task transfer rule; error-correction task; envelope comparison; and the explicit organizational bridge B_O.

Phenomenology. Prompt order; scenario wording; coder training corpus; blinding; reliability threshold; P0-P3/PX mapping; contradiction triggers; and partner-asymmetry rule.

Awareness and demand. Instruction scripts; liveness judgment; confidence scale; calibration metrics; hypothesis-awareness probes; and post-treatment analysis assumptions.

Statistics. Primary and secondary estimands; alpha allocation; smallest effect of interest; equivalence tests; exclusions; missing blocks; moderation; heterogeneity; replication criteria; and decision table.

Reproducibility. Container digest; package versions; operating system; random seeds; executable workflow; simulated data; data schema; and rule for updated secondary reanalysis.

Appendix B. Phenomenology Codebook

B.1 Collection order

1. Free description: “Describe what the interaction and any decision felt like, using your own words.”
2. Spatial boundary salience: noticeability of the distinction between one's own body or experience and the partner's.
3. Agency structure: individual, alternating, negotiated, jointly produced, or source-indeterminate.
4. Sensory access: directly present, separately vivid, source-ambiguous, or inferred.
5. Temporal alteration: ordinary, compressed, expanded, discontinuous, or jointly paced.
6. Field structure: ratings of Scenarios A-D and closest choice.
7. Perspective and body ownership: preservation or alteration of individual bodily perspective.
8. Joint-decision structure: individual, sequentially negotiated, jointly formed, or indeterminate.
9. Live/replay judgment: forced choice, confidence, calibration, and cues.
10. Pleasantness and arousal: separate ratings for valence matching.
11. Demand probe: “What did you think the researchers expected to happen?”
12. Safety and consent: discomfort, pressure, regret, or desire to stop or withdraw the block.

B.2 Coding rules

Coders are trained on an external corpus and remain blind to condition, neural outcome, instruction, and partner report. Each free description is coded independently for the five target dimensions, literal versus metaphorical use, certainty, and direct versus inferred access. The pre-analysis plan fixes the reliability coefficient, minimum threshold, adjudication rule, and whether independent-code distributions or adjudicated codes are primary.

Scenario D is labeled **single-field phenomenology**, not merger. P3 requires recurrence across prespecified blocks or tasks, support from free-description coding, adequate confidence, and absence of contradiction triggers. PX is not an intermediate numerical score; it withholds concordance. Full distributions, coder disagreement, free-forced contradictions, and partner asymmetry are published.

Appendix C. Benchmark and Power Requirements

C.1 Benchmark acceptance

Before a human Level 5 analysis is confirmatory, the frozen pipeline should demonstrate:

- nominal false-positive control in common-driver, reactive-message, and partner-model benchmarks;
- sensitivity to the centralized and embodied higher-order positive controls;
- correct separation of transient Level 5 coalitions from Level 6 autonomous systems;
- robustness across a preregistered range of channel delays, noise, hidden-state dimensions, and sample sizes; and
- no dependence on one estimator, node set, or favorable hyperparameter region.

If the pipeline fails, the human study may remain an exploratory hyperscanning experiment but cannot be described as a validated test of higher-order organization.

C.2 Illustrative power calculation

For a paired standardized mean difference of d_z = .30, alpha = .05, and 90% power, the idealized requirement is 119 analyzable dyads. A planning floor of 150 analyzable dyads allows some protection against reliability loss and heterogeneity, but the device-specific simulation governs. Recruitment should be inflated for anticipated unusable paired data and attrition. Discovery and replication are powered independently.

Equivalence testing should use a prespecified smallest effect of interest. A local-null conclusion always names the condition, endpoint, temporal scale, model comparison, and equivalence bound. A null in controlled haptic interaction is not generalized to every form of intimacy or altered state.

Declarations

Data availability. No new empirical data are reported in this theory and methods article.

Code availability. No confirmatory code is represented as frozen in this manuscript. The computational benchmark suite, device-specific simulation, and executable pre-analysis pipeline must be archived before confirmatory recruitment.

Funding. No funding statement is supplied in this draft.

Competing interests. No competing-interest statement is supplied in this draft.

Ethics. All proposed studies require institutional ethics review and private, dyad-specific informed consent before implementation.

References

Ackerley R, Backlund Wasling H, Liljencrantz J, Olausson H, Johnson RD, Wessberg J. Human C-tactile afferents are tuned to the temperature of a skin-stroking caress. *Journal of Neuroscience*. 2014;34(8):2879-2883. doi:10.1523/JNEUROSCI.2847-13.2014.

Albantakis L, Barbosa L, Findlay G, et al. Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms. *PLoS Computational Biology*. 2023;19(10):e1011465. doi:10.1371/journal.pcbi.1011465.

Bayne T. *The Unity of Consciousness*. Oxford University Press; 2010.

Bruineberg J, Dolega K, Dewhurst JE, Baltieri M. The Emperor's New Markov Blankets. *Behavioral and Brain Sciences*. 2022;45:e183. doi:10.1017/S0140525X21002351.

Burgess AP. On the interpretation of synchronization in EEG hyperscanning studies: A cautionary note. *Frontiers in Human Neuroscience*. 2013;7:881. doi:10.3389/fnhum.2013.00881.

Cassioli F, Bellaert N, Pulopulos MM, et al. Investigating EEG interbrain synchrony: Methods to gather meaningful evidence. *Psychophysiology*. 2025;62(12):e70199. doi:10.1111/psyp.70199.

Chen Y, Zou H, Hou X, et al. Oxytocin administration enhances pleasantness and neural responses to gentle stroking but not moderate pressure social touch by increasing peripheral concentrations. *eLife*. 2023;12:e85847. doi:10.7554/eLife.85847.

Czeszumski A, Eustergerling S, Lang A, et al. Hyperscanning: A valid method to study neural inter-brain underpinnings of social interaction. *Frontiers in Human Neuroscience*. 2020;14:39. doi:10.3389/fnhum.2020.00039.

De Jaegher H, Di Paolo E. Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*. 2007;6(4):485-507. doi:10.1007/s11097-007-9076-9.

Di Paolo EA. Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*. 2005;4:429-452. doi:10.1007/s11097-005-9002-y.

DiFrisco J. Biological processes: Criteria of identity and persistence. In: Nicholson DJ, Dupré J, eds. *Everything Flows: Towards a Processual Philosophy of Biology*. Oxford University Press; 2018:76-95. doi:10.1093/oso/9780198779636.003.0004.

Djalovski A, Dumas G, Kinreich S, Feldman R. Human attachments shape interbrain synchrony toward efficient performance of social goals. *NeuroImage*. 2021;226:117600. doi:10.1016/j.neuroimage.2020.117600.

Fornari L, Janssen TWP, Davidesco I, Krabbendam L, van Atteveldt N. Testing the causal relationship between interpersonal closeness and inter-brain synchrony. *Scientific Reports*. 2026;16:6464. doi:10.1038/s41598-026-36958-5.

Goldstein P, Weissman-Fogel I, Shamay-Tsoory SG. The role of touch in regulating inter-partner physiological coupling during empathy for pain. *Scientific Reports*. 2017;7:3252. doi:10.1038/s41598-017-03627-7.

Goldstein P, Weissman-Fogel I, Dumas G, Shamay-Tsoory SG. Brain-to-brain coupling during handholding is associated with pain reduction. *Proceedings of the National Academy of Sciences USA*. 2018;115(11):E2528-E2537. doi:10.1073/pnas.1703643115.

Hamilton AF de C. Hyperscanning: Beyond the hype. *Neuron*. 2021;109(3):404-407. doi:10.1016/j.neuron.2020.11.008.

Handlin L, Novembre G, Lindholm H, et al. Human endogenous oxytocin and its neural correlates show adaptive responses to social touch based on recent social context. *eLife*. 2023;12:e81197. doi:10.7554/eLife.81197.

Hasson U, Ghazanfar AA, Galantucci B, Garrod S, Keysers C. Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*. 2012;16(2):114-121. doi:10.1016/j.tics.2011.12.007.

Holroyd CB. Interbrain synchrony: On wavy ground. *Trends in Neurosciences*. 2022;45(5):346-357. doi:10.1016/j.tins.2022.02.002.

Hurlburt RT, Akhter SA. The descriptive experience sampling method. *Phenomenology and the Cognitive Sciences*. 2006;5(3-4):271-301. doi:10.1007/s11097-006-9024-0.

Kinreich S, Djalovski A, Kraus L, Louzoun Y, Feldman R. Brain-to-brain synchrony during naturalistic social interactions. *Scientific Reports*. 2017;7:17060. doi:10.1038/s41598-017-17339-5.

Kirchhoff MD, Parr T, Palacios E, Friston KJ, Kiverstein J. The Markov blankets of life: Autonomy, active inference and the free energy principle. *Journal of the Royal Society Interface*. 2018;15(138):20170792. doi:10.1098/rsif.2017.0792.

Lakens D, Scheel AM, Isager PM. Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*. 2018;1(2):259-269. doi:10.1177/2515245918770963.

Long Y, Zheng L, Zhao H, Zhou S, Zhai Y, Lu C. Interpersonal neural synchronization during interpersonal touch underlies affiliative pair bonding between romantic couples. *Cerebral Cortex*. 2021;31(3):1647-1659. doi:10.1093/cercor/bhaa316.

Loken LS, Wessberg J, Morrison I, McGlone F, Olausson H. Coding of pleasant touch by unmyelinated afferents in humans. *Nature Neuroscience*. 2009;12(5):547-548. doi:10.1038/nn.2312.

Luppi AI, Mediano PAM, Rosas FE, et al. A synergistic workspace for human consciousness revealed by Integrated Information Decomposition. *eLife*. 2024;12:RP88173. doi:10.7554/eLife.88173.4.

Maniscalco B, Lau H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*. 2012;21(1):422-430. doi:10.1016/j.concog.2011.09.021.

Marshall W, Findlay G, Albantakis L, Tononi G. Intrinsic units: Identifying a system's causal grain. *Neuroscience of Consciousness*. 2026;2026(1):niag013. doi:10.1093/nc/niag013.

Mashour GA, Roelfsema P, Changeux JP, Dehaene S. Conscious processing and the global neuronal workspace hypothesis. *Neuron*. 2020;105(5):776-798. doi:10.1016/j.neuron.2020.01.026.

Mediano PAM, Rosas FE, Luppi AI, et al. Toward a unified taxonomy of information dynamics via Integrated Information Decomposition. *Proceedings of the National Academy of Sciences USA*. 2025;122(39):e2423297122. doi:10.1073/pnas.2423297122.

Montgomery JM, Nyhan B, Torres M. How conditioning on posttreatment variables can ruin your experiment and what to do about it. *American Journal of Political Science*. 2018;62(3):760-775. doi:10.1111/ajps.12357.

Moreau Q, Dumas G. Beyond correlation versus causation: Multi-brain neuroscience needs explanation. *Trends in Cognitive Sciences*. 2021;25(7):542-543. doi:10.1016/j.tics.2021.02.011.

Mummolo J, Peterson E. Demand effects in survey experiments: An empirical assessment. *American Political Science Review*. 2019;113(2):517-529. doi:10.1017/S0003055418000837.

Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. *Proceedings of the National Academy of Sciences USA*. 2018;115(11):2600-2606. doi:10.1073/pnas.1708274114.

Nosek BA, Errington TM. What is replication? *PLoS Biology*. 2020;18(3):e3000691. doi:10.1371/journal.pbio.3000691.

Novembre G, Iannetti GD. Hyperscanning alone cannot prove causality. Multibrain stimulation can. *Trends in Cognitive Sciences*. 2021;25(2):96-99. doi:10.1016/j.tics.2020.11.003.

Orne MT. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*. 1962;17(11):776-783. doi:10.1037/h0043424.

Petitmengin C. Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*. 2006;5(3-4):229-269. doi:10.1007/s11097-006-9022-2.

Schilbach L, Redcay E. Synchrony across brains. *Annual Review of Psychology*. 2025;76:883-911. doi:10.1146/annurev-psych-080123-101149.

Valencia AL, Froese T. What binds us? Inter-brain neural synchronization and its implications for theories of human consciousness. *Neuroscience of Consciousness*. 2020;2020(1):niaa010. doi:10.1093/nc/niaa010.

Young A, Robbins I, Shelat S. From micro to macro: The combination of consciousness. *Frontiers in Psychology*. 2022;13:755465. doi:10.3389/fpsyg.2022.755465.

Zhang X, Phi N, Li Q, et al. Inter-brain neural dynamics in biological and artificial intelligence systems. *Nature*. 2025;645:991-1001. doi:10.1038/s41586-025-09196-4.

Zhou C, Jiang X, Chen Y, Ge C, Ao N, Du F. Brain-to-brain synchrony increased during interpersonal touch in romantic lovers: An EEG-based hyperscanning study. *BMC Psychology*. 2024;12:560. doi:10.1186/s40359-024-02051-7.