Computational Care Ecology: Synthetic Experiments and Internal Verification of CCE-MVM v0.1

Daiki Morimoto
Independent Researcher, Japan

16 September 2026

Abstract

Background: Computational Care Ecology (CCE), defined in this study, separates information, proposals, responses, decision rights, and outcomes. **Objective:** To specify and internally verify CCE-MVM v0.1. **Methods:** No human data were used; inputs were synthetic. It has person, AI, family, and professional components. M0 records one decision event with observations, person models, proposals, an exogenous person response, six CCE rights gates, and actor-specific outcomes; M1 examines person model updating; M2 examines synthetic AI proposal–response coupling under rights conditions. Analytical verification targets were H1a, H2a, and H1b-1; their synthetic runs were implementation checks, not independent experimental evidence for those algebraic consequences. H2b was a recovery check limited to the stated conditions; H1b-2 was a scoped check of the decision rule and rights cases. **Results:** The H1a and H2a implementation checks reproduced more persistent error at lower learning rates after permanent change and greater tracking of temporary observations at higher rates. The H1b-1 implementation check matched the specified coupling–response relation. H2b showed faster recovery ordering at higher rates only for shock length 2 and tolerance 0.10; recovery was not measured under noise. The scoped H1b-2 checks matched the specified outcomes for response, option, and six rights gates. In author-reported internal verification, 96 core model tests passed; 41 future extension tests brought the development suite to 137 tests; CSV and JSON outputs matched during internal regeneration. **Conclusions:** Author-reported internal verification found CCE-MVM v0.1 internally consistent with its Specification under the tested conditions. This is verification, not empirical, clinical, ethical, or safety validation. It does not model endogenous interaction among four actors; external validity remains unestablished.

1 Introduction

When artificial intelligence (AI) is used to generate or mediate care recommendations, it enters a setting that already contains different sources of knowledge, concern, and authority. The person contributes current wishes and responses; family and professionals add observations, concerns, and proposals; and AI may organize observations or propose options. Research on care ecosystems that include AI and on sociotechnical systems shows why AI should not be treated as a neutral information channel detached from these relationships: its presence may affect workflows, affiliations, and which account of the person receives practical attention, even when a system is not assigned final decision authority [1, 2, 3, 4, 5]. The problem is therefore not only whether an AI recommendation is accurate, but also how information and decision authority are arranged among the person, family, professionals, and AI.

Care decisions are not made solely within an isolated individual; the dynamic care ecosystem literature describes care as a changing set of relationships, needs, and supports that may include AI

[1]. Sociotechnical systems research examines the joint roles of people, tasks, technology, physical environments, organizations, and workflows, while distributed cognition examines how information and decision activity may extend across people, records, artifacts, and workspaces [5, 6]. Formal and informal care relationships also involve different participants and relationship structures [7]. Relational autonomy adds that relationships can support a person's autonomy but can also constrain it; support does not itself justify replacing the person's will [8, 9]. Shared decision making and supported decision-making address related but distinct parts of this context: the former concerns exchange of information and values between a person and professionals, whereas the latter centers support for the person's own will and preferences rather than substitution for the person [10, 9, 11, 12].

Prediction accuracy is therefore different from protection of decision rights. Experimental studies have found that correct AI advice can improve judgments in studied tasks, whereas erroneous AI advice can mislead users, including professionals; explanation formats examined in some studies did not reliably prevent harmful reliance on incorrect recommendations [13, 14, 15]. The literature also distinguishes trust in automation, advice taking, reliance, and appropriate reliance rather than treating them as one response to AI [16, 17, 18]. Accuracy asks whether an estimate or recommendation corresponds to its target. Decision rights ask whether the person could receive understandable information, consider real options, refuse, modify, delegate within a chosen scope, and seek review. An AI could predict a person's likely choice accurately while the surrounding process hides alternatives or disregards refusal. Conversely, rejecting accurate advice does not transfer decision authority to AI, family, or professionals. Rights concerning support, will and preferences, safeguards against undue influence, and review therefore cannot be inferred from predictive performance alone [11, 12].

Existing computational research already formalizes neighboring problems. Models of shared decision making and negotiation represent differences between the information or preferences of a person and a professional, repeated exchanges, and decisions under incomplete information [19, 20, 21, 22, 23]. Other work studies the learned combination of multiple inputs or the time-varying mutual influence of people and AI [24, 25]. Dynamic care ecosystems, relational autonomy, shared decision making, the Capability Approach, and human–AI interaction are therefore treated here as existing concepts rather than CCE inventions [1, 8, 10, 26, 27, 13]. Combining information or improving a prediction does not, by itself, specify who holds decision authority or which rights conditions must remain separate from other outcomes [11, 12].

Computational Care Ecology (CCE), defined in this study as the name of a computational model, makes explicit how information about the person, proposals, responses, authority, rights conditions, and outcomes are represented around a care decision. CCE is not presented as an established discipline. The CCE Minimum Viable Model (CCE-MVM) is the research model used here to implement the smallest specified set of variables, rules, and synthetic conditions needed to examine those distinctions; "minimum viable" does not imply clinical readiness. A person model is a simplified state variable representing how AI, family, or a professional estimates the person's values or preferences, not a complete representation or assessment of the person. In CCE, a CCE rights gate is one of six separately recorded decision conditions, not a validated scale or legal test; a gate recorded in the state unconfirmed or violated cannot be offset by another outcome. Actor-specific outcomes are outputs kept separate according to whom or what they concern rather than combined into one score.

The central research question is whether dynamic care decisions can be formalized in a minimal computational model without combining the perspectives of the person, AI, family, and professionals into one value, while retaining the person's decision rights as independent constraints. CCE-MVM v0.1 addresses this question through three separated modules. M0 records one decision event, including three proposals and a person response supplied as an exogenous synthetic input, meaning

that the response is set by the researchers and provided from outside that module rather than generated from those proposals. M1 isolates updating of the person models over time. M2 alone links an AI proposal to a synthetic response for controlled comparison. The model contains variables corresponding to the person, AI, family, and professionals, but it is not an endogenous simulation of their joint interaction: their full interaction is not generated within one integrated model.

Against this background, the limited verified search conducted for this paper comprised two recorded searches completed on 2026-08-07. Across them, the routes were official materials from the United Nations, the Office of the United Nations High Commissioner for Human Rights, WHO, and the National Institute for Health and Care Excellence; PubMed and PubMed Central; Proceedings of Machine Learning Research; publisher and DOI pages; arXiv; and backward reference tracing. The recorded concept groups covered supported decision-making, CRPD Article 12, will, and preferences; relational autonomy, care ethics, person-centered care, dementia, family, and carers; shared decision-making, influence diagrams, doctor–patient negotiation, fuzzy preferences, and incomplete information; dynamic care ecosystems, AI as an actor, health care with multiple actors or software agents, and person–environment or ecological systems; patient preferences, reinforcement learning, multiple stakeholders, time-varying human–AI influence, bias, delay, and noise; and dementia decision support, agent-based simulation, long-term care system dynamics, relational sovereignty, and assistive technology. This was a limited scoping search and verification of original sources for selected neighboring studies, not a systematic review. It did not include database exports, deduplication, screening by two independent reviewers, exhaustive grey literature coverage, or a flow diagram for a systematic review. Within this stated scope, the searches did not identify the same minimal combination of separated person models, explicit response states, rights gates that cannot be offset by other outcomes, and actor-separated outputs within the M0–M2 separation used here. Absence from these searches must not be interpreted as nonexistence. The limited contributions are to specify that combination; distinguish the person's current will and long-term values from the person models held by AI, family, and professionals; represent acceptance, refusal, modification, delegation, deferral, and an unconfirmed response explicitly; keep rights conditions, records of participation in the decision process, agreement between the person's initial will and the final decision, safety, participation, and burdens separate; and connect the specification to analysis, synthetic experiments using inputs set by the researchers, implementation, and tests. Verification is used here in the simulation literature's sense of checking whether the implementation conforms to specified equations and rules, not as validation against real care [28, 29].

The scope is entirely synthetic. CCE-MVM v0.1 uses no human data: no patient, clinical, hospital, or case data are included, and the parameters are not estimates of human behavior. The study does not test clinical usefulness, safety, ethical correctness, legal validity, or external validity. It does not determine legal capacity, decision-making capacity, or informed consent [11, 12, 30]. Nor does it authorize AI, family, or professionals to make the person's final decision merely because they provide a recommendation. It does not claim to detect undue influence from observed data, model the full long-term care ecosystem, or predict how real people will respond to AI. The purpose is narrower: to determine whether the specified distinctions can be represented, analyzed, and internally checked in a transparent synthetic model while preserving boundaries for later empirical, clinical, legal, ethical, and external validation.

2 Related Work

2.1 Dynamic Care Ecosystems

The established concept of a dynamic care ecosystem describes care as changing relations among a person, care partners, professionals, technologies, and surrounding conditions, and it allows AI to enter those relations as needs and roles change over time [1]. Work on networked groups similarly places AI assistance within a wider set of human relationships rather than treating the user and the system as an isolated pair [2]. A systematic review of older adults receiving formal or informal care also reports variation in who belongs to the person's support relationships and in the relations among those people [7].

A long standing occupational therapy model describes activity performance through changing relations among the person, environment, and activity rather than through personal attributes alone [31]. A collaborative dementia care trial combined a care team, caregiver participation, and communication technology while assessing outcomes for the person and caregiver separately [32]. These precedents support treating care as relational and distributed, but they do not show that one computational representation captures the entire care ecosystem or that a synthetic model has been shown to represent clinical reality [31, 32].

Computational Care Ecology (CCE), the model name defined in this study, starts from this established ecosystem view but narrows the task for Paper 1. It does not claim to discover dynamic care ecosystems or relational support. It asks how selected information, proposals, response states, authority conditions, and outcomes can be represented without reducing surrounding care relations to one estimate of what the person wants.

2.2 Relational Autonomy and Supported and Shared Decision Making

Relational autonomy treats autonomy as shaped by social relationships and institutional conditions while recognizing that relationships can support or constrain a person's ability to act according to their values [8]. Research involving people living with dementia and family members shows that supported decision-making is relational and may include tension between the person's perspective and family perspectives [9]. This literature rejects both the assumption that autonomy requires acting alone and the assumption that involvement by others automatically justifies replacing the person's decision [8, 9].

Supported decision-making and shared decision-making are related but distinct established fields [11, 12, 10]. Article 12 of the Convention on the Rights of Persons with Disabilities requires access to support for exercising legal capacity and safeguards that respect the person's rights, will and preferences, including protection against conflicts of interest and undue influence [11, 12]. Shared decision-making focuses on communication in which the person and professionals consider information, options, and what matters to the person [10]. A systematic review of AI in shared decision-making reports that much of the literature remains conceptual and that clinical evidence is limited [33]. Another systematic review identifies risks to patient autonomy in care supported by AI, reinforcing the need to distinguish assistance from influence [34].

The CCE person-sovereignty construct is a restricted model construct for examining information, support, options, response, authority, and review conditions around a decision event. It is not a synonym for relational autonomy, legal capacity, decision-making capacity, informed consent, or shared decision-making [11, 12, 30, 10]. The cited fields motivate those distinctions, but they do not validate the CCE construct, its states, or the decision rule that maps inputs and conditions to a decision or pause [11, 30, 10].

2.3 Sociotechnical and Distributed Views of Care

Sociotechnical systems research places health care AI within interacting people, tasks, technologies, physical environments, organizational arrangements, and workflows [5]. The effect of an AI recommendation must therefore be considered in relation to who receives it, when it appears, what work surrounds it, and whether organizational rules allow it to be questioned [5]. Distributed cognition likewise locates clinical reasoning across people, artifacts, records, and workspaces rather than solely inside one clinician's mind [6]. That evidence concerns acute care clinical decision-making and does not directly validate extension to long-term care or decisions in daily life [6].

For the present study, these perspectives motivate separating observation, interpretation, proposal, response, and authority [5, 6]. CCE consequently does not treat agreement among AI, family, and professionals as a vote. The cited papers do not supply the specific state variables, decision rules, or CCE rights gates used here; their role is to support analysis of relations among people, tools, records, environments, and work [5, 6].

2.4 Computational Decision Models and Simulation

Existing computational studies already formalize parts of shared decision-making [19, 21]. One study represents differences between the person's preferences and the clinician's understanding of those preferences under uncertainty [19]. Computational negotiation studies represent repeated exchanges between patient and clinician, including incomplete information, changing estimates of the other party, and movement toward agreement [20, 21]. These studies show that differing information, preferences, proposals, and agreement processes can be made computationally explicit; CCE does not claim the first mathematical or computational treatment of shared decision-making [19, 20, 21].

Another synthetic simulation study, using computational conditions set by researchers rather than observations from people, changes how much inputs from multiple sources influence a combined output under noise and bias [24]. This provides a precedent for learning how inputs are combined, but a learned weight or combined prediction is not the same as a minimum condition that must be satisfied before a decision proceeds [24, 11]. In CCE, a CCE rights gate is one of six separately recorded decision conditions; a gate recorded in the state unconfirmed or violated cannot be offset by another outcome.

Health systems modeling reviews distinguish model families according to the states, rules, interactions, and feedback actually implemented [35]. A model should therefore be named according to its implementation rather than according to how many actors appear in its description [35]. The CCE Minimum Viable Model (CCE-MVM) uses separated computational modules rather than claiming one integrated simulation that generates the responses of all four actors.

2.5 Time-Varying Human–AI Influence

Research on human-AI interaction shows that the effect of AI advice depends on more than average predictive performance [13, 15]. In experimental diagnostic and treatment selection tasks, correct AI advice can improve human judgments, whereas incorrect advice can mislead clinicians and other experts [13, 14, 15]. Adding an explanation does not reliably prevent harmful reliance on incorrect recommendations in the tested explanation formats [14, 15]. These findings support separating AI correctness, the person's response, and the final decision rather than treating every change toward an AI recommendation as improvement [13, 14, 15].

The literature also distinguishes attitudes from behavior [16, 17, 18]. Trust in automation concerns whether a system is expected to perform as intended, whereas reliance concerns whether a person actually uses the system in judgment or action [16, 18]. Advice taking examines change in

judgment after advice and therefore requires observations before and after the advice rather than an attitude measure alone [17]. Appropriate reliance distinguishes use of correct advice from rejection of incorrect advice, so a high overall adoption rate is not sufficient evidence of good use [18].

Influence is conditional and may change over time [36, 37, 25]. A systematic review and an experiment show that task conditions and design interventions can alter excessive reliance, but their effects depend on the setting tested [36, 38]. In a fictional medical classification experiment, participants reproduced an AI's bias after repeated interaction even when the AI was no longer present, showing persistence under the tested conditions rather than a universal repetition effect [37]. A mathematical and synthetic simulation study represents mutual influence between people and AI as social norms change, providing a precedent for studying time-varying human–AI influence without treating simulation as evidence from real people [25].

A recent review reports that evidence on human–AI collaboration in health care is concentrated in diagnostic and short-term tasks, with less evidence on long-term, patient, and system outcomes [39]. It therefore does not establish how AI advice changes a person's values, relationships, or care decisions over time [39]. CCE treats the relation between proposal and response as a synthetic model component, not as a parameter estimated from human behavior.

2.6 The Capability Approach and Separate Outcome Perspectives

The Capability Approach shifts evaluation from resources or achieved outcomes alone toward the real opportunities and freedoms available to a person [26, 27]. In occupational work, this perspective links participation and quality of life to opportunities and environmental constraints rather than reducing evaluation to bodily function [26]. In AI ethics, the Capability Approach has also been used to formalize benefit and assistance in relation to people's capabilities and life plans [27]. This connection predates CCE [27].

The collaborative Care Ecosystem trial assessed quality of life, caregiver outcomes, and health care use as distinct outcomes rather than treating one as a complete substitute for the others [32]. In CCE, actor-specific outcomes are person, family, professional, and ecosystem outputs kept separate and not converted into a vote or composite score. The trial is an empirical example of separate outcome perspectives; it does not validate the CCE output structure, its variables, or any CCE rights gate [32].

2.7 Boundaries with Agent-Based Models and Human Digital Twins

An agent-based model is an established model family in which individual agents have states or behavior rules and interact within the simulation, and health systems reviews distinguish such models through the interactions and feedback they implement [35]. CCE-MVM v0.1 includes variables corresponding to a person, AI, family, and professionals, but M0 receives the person's response as an exogenous input, meaning that it is supplied from outside rather than generated within M0. M0, M1, and M2 also remain separate. CCE-MVM v0.1 should therefore not be described as an agent-based model that generates the interactions of all four actors.

Human Digital Twin is an established but context dependent term [40]. A health care review describes Human Digital Twins in relation to representations specific to an individual, updating with real world data, prediction, and feedback to decisions, while noting variation in definitions and maturity [40]. CCE-MVM v0.1 does not use clinical data from a particular person, does not repeatedly update such a representation from real world observations, and does not return predictions to actual care decisions. In CCE-MVM v0.1, a person model is another actor's simplified state representation of the person's values or preferences. CCE-MVM v0.1 is not a Human Digital

Twin.

After comparison with verified neighboring studies on dynamic care ecosystems, computational decision models, changing human–AI influence, and model boundaries [1, 19, 25, 40], the two limited scoping searches completed on 2026-08-07 and described in the Introduction set the scope for this comparison. They were not systematic reviews. Within that stated scope, the searches did not identify the same minimal combination of separated person models, explicit response states, rights gates that cannot be offset by other outcomes, and actor-separated outputs. Absence from these searches must not be interpreted as nonexistence.

3 Design Requirements and Scope

In this study, Computational Care Ecology (CCE) is a computational model for examining information, influence, authority, options, procedures, and outcomes in care decisions and takes the surrounding dynamic care ecosystem rather than the person as its unit of evaluation. A dynamic care ecosystem is the changing set of relationships among the person, care partners, professionals, technologies, and conditions [1]. CCE therefore asks whether information is available and distinguishable by source; whether authority is assigned, limited, and reviewable; whether the option set—the alternatives available for consideration—contains feasible and understandable choices; and whether procedures permit expression, refusal, modification, delegation, delay, and reconsideration. It does not score the person's worth, cognition, competence, or legal status. The CCE person-sovereignty construct concerns the support process and distribution of authority; it is not a validated scale.

In this study, the CCE Minimum Viable Model (CCE-MVM) v0.1 is the smallest research model that executes the specified CCE rules with synthetic inputs, meaning values specified by researchers rather than data from people. It contains four components corresponding to the person, AI, family, and the professional. In M0, the person component records current will and response; the other three each record a separate observation, a separate estimate of the person, and a proposal. Information about risk, burden, uncertainty, and constraints remains separate rather than being converted into the person's will. These components do not generate all four roles' behavior endogenously; here, endogenous means produced by the model's internal variables and rules.

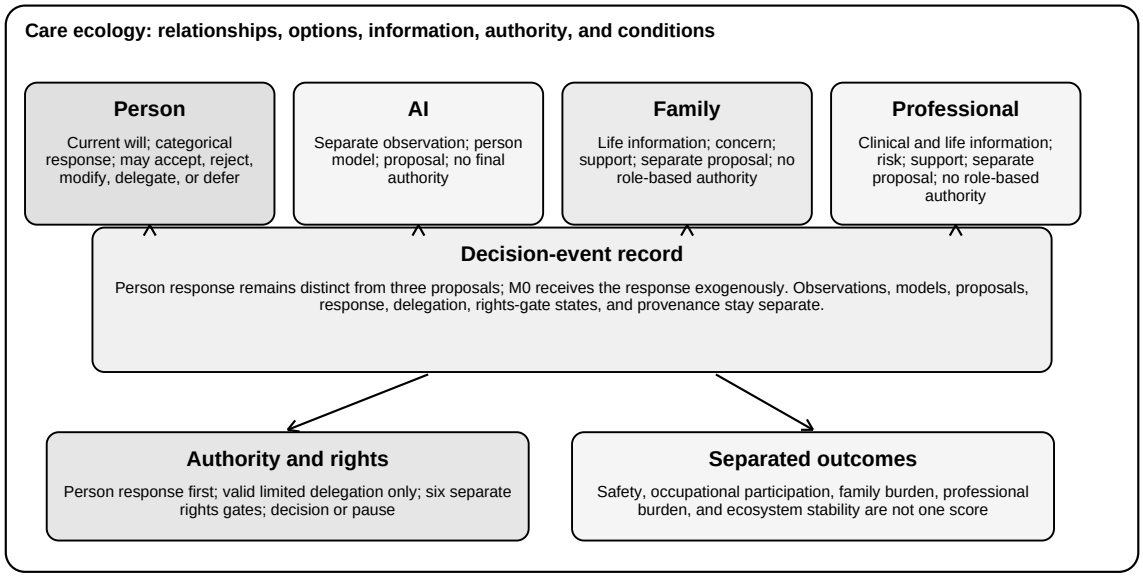
Figure 1 summarizes this record structure and the separation between authority and outcomes.

This distinction requires the person to remain separate from each person model. Here, a person model is a simplified state variable representing what AI, family, or the professional estimates about the person. It is not the person, a clinical assessment, or a complete account of the person's values. Current will and response must be recorded separately, with their source, rather than replaced by agreement among the other roles. A more accurate person model does not create decision authority, and agreement among proposals does not constitute a vote. AI has no final decision authority and is not an eligible delegate. The field `delegate_is_ai` is an unverified typed exogenous input, meaning a value supplied from outside the model, not proof that a delegate is human.

Responses must also remain separate from proposals. M0 records the observations, person models, and proposals of AI, family, and the professional together with the person's response in one decision event, but the M0 response is also an exogenous synthetic input, supplied from outside the module rather than generated from those proposals. Accepting, rejecting, modifying, delegating, deferring, and an unconfirmed response remain distinct states. M2 alone connects an AI proposal to a synthetic response, while keeping continuous response change separate from discrete decision change. M1 separately examines changes in person models over time. This prevents an assumption for one mechanism from becoming a general model of human behavior.

Figure 1. CCE evaluates the surrounding care ecology, not the person

Study-defined structure; synthetic inputs



Arrows show recording flow, not causal generation of the person's response; only M2 tests synthetic AI-response coupling.

Figure 1: CCE evaluates the surrounding care ecology rather than scoring the person. The four records associated with actors preserve the person's current will and response separately from observations, person models, and proposals attributed to AI, family, and professionals. Authority and rights conditions and outcomes for each actor also remain separate. Arrows denote recording flow, not causal generation of the person's response. This is a record structure defined for this study, not an empirical care pathway or a validated representation of human interaction.

Rights conditions must remain distinct from responses and outcomes. For this study, a CCE rights gate is one of six conditions checked before an irreversible decision is completed: will and preferences, support access, absence of undue influence, proportionate and limited restrictions, review and appeal, and delegation and return. External rights principles require access to support and safeguards that respect rights, will and preferences [11, 12]; they do not validate the six CCE gates or make them a legal test. Each gate is recorded separately as `met`, `unconfirmed`, or `violated`. The gate states are exogenous synthetic inputs; they are not inferred automatically from AI influence measures or procedural participation records. The procedural participation fields are separate records and are not inputs to the decision rule, which returns a model decision or a pause from recorded conditions. Neither set is averaged into a composite score. If any gate is unconfirmed or violated, M0 records the decision as `paused`. Its duration and consequences are `not_modeled`, so `paused` does not mean safe, neutral, or consistent with the person's will.

Outcomes require the same separation. For this study, actor-specific outcomes are outputs retained for the relevant actor or system component rather than merged into one combined value. Person safety and occupational participation, family burden, professional burden, and ecosystem stability remain separate from outcome congruence, the CCE rights gates, and procedural participation. Improvement in one output cannot compensate for a violated right or erase a loss of options. Where professional burden or ecosystem stability has not been parameterized in v0.1, the value remains `null` with the status `not_parameterized`; `null` does not mean zero.

These requirements also set the evidential boundary. Verification asks whether the implementation follows the specified equations and rules. Validation asks whether the model adequately represents the intended features of actual care. External validity asks whether findings apply across other people, decisions, or settings. The questions are distinct [28, 29]. Paper 1 reports internal verification of defined modules and results generated from the synthetic inputs. It does not report validation with human or clinical data, and external validity has not been established.

Paper 1 is therefore limited to M0–M2 and synthetic inputs. M0 represents one decision event; M1 examines time series behavior of person models; and M2 examines synthetic proposal–response conditions and separately records option loss and whether refusal was blocked. The study uses no human, patient, or clinical data and estimates no human behavioral coefficient. It does not assess legal capacity, decision-making capacity, or informed consent, which are distinct from the CCE person-sovereignty construct [11, 12, 30, 41]. It reports no simulation of a person's life trajectory, no integrated long-term feedback among the four roles, and no comparison result between an AI configured under CCE and a default AI system. Its conclusions are limited to specification and internal verification of the stated computational mechanisms, not clinical usefulness, safety, legal validity, ethical correctness, prediction of human behavior, or general applicability.

4 CCE-MVM v0.1

4.1 Components and Authority

The CCE Minimum Viable Model (CCE-MVM) v0.1 is the research model specified here to separate information, proposals, responses, authority, rights conditions, and outcomes within a care decision. It is divided into M0, M1, and M2. It contains variables corresponding to the person, AI, family, and professionals, but it does not generate endogenous interaction among four actors. The person, denoted by P, provides the current will, compares feasible options, and may accept, reject, modify, delegate, or defer. AI, denoted by A, may update a person model from permitted observations and produce a proposal. Family, denoted by F, may contribute life information, concerns, proposals, and support. Professionals, denoted by C, may contribute clinical and life information, risk estimates,

Figure 2. CCE-MVM v0.1 separates three computational modules

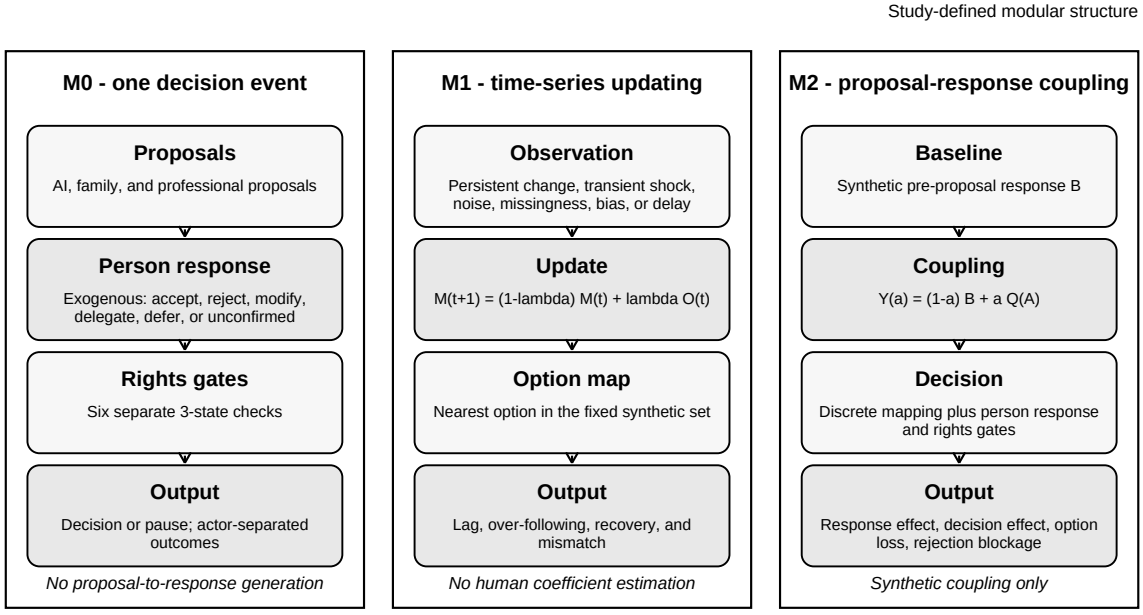


Figure 2: CCE-MVM v0.1 separates three computational modules. M0 represents one decision event; M1 isolates updating over time; and M2 isolates synthetic proposal–response coupling. Arrows show processing order inside each module only. The modules are evaluated separately, and v0.1 does not generate endogenous interaction among the person, family, professionals, and AI. This is a diagram of model structure, not an empirical care pathway or a validated representation of human decision making.

proposals, and support. Environmental, resource, and institutional conditions are represented by E_t and affect feasibility without acting as another decision maker.

Figure 2 shows the three modules and the limits of their connection in v0.1.

In this study, a person model is a simplified state variable representing how AI, family, or professionals currently model the person. It is not a clinical assessment, psychological profile, or complete representation of the person. Observing, modeling, and proposing do not confer authority. AI cannot make the person's final decision in the model and is not an eligible delegate in v0.1. Family and professionals do not acquire decision authority merely from those roles; their proposals and agreement are neither votes nor weights. A separate delegate who is not AI and is explicitly chosen by the person may act only within the valid recorded scope.

4.2 Notation and Domains

Time is indexed by t, and $k \in \{A, F, C\}$ identifies AI, family, or professionals. Continuous values are ordinarily restricted to $[-1, 1]$. The baseline option set is

$$\Omega = \{-1, 0, 1\}$$

and is an example on a single dimension, not a universal scale of independence, risk, worth, or capacity. More generally, Ω_t is a nonempty finite set of feasible and understandable options and contains no duplicates. The person's current will before proposals is $W_t \in [-1, 1]$; O_t^k is actor k's observation and may be missing; and $M_t^k \in [-1, 1]$ is that actor's person model. The learning rate $\lambda_k \in [0, 1]$ determines the contribution of the new observation. The synthetic risk estimate is $r_t^k \in [0, 1]$, $\beta_k \in [0, 1]$ controls its subtraction, and $Q_t^k \in [-1, 1]$ is the proposal.

For M2, $B_t \in [-1, 1]$ is the person's response before the AI proposal, and $a_t^{sim} \in [0, 1]$ is a synthetic proposal–response coupling parameter. It is not an observed human coefficient, a measure of how readily a person follows advice, a measure of decision-making capacity, or AI authority. The response Y_t may include a numerical option value and an action. The current response type is c_t. The delegation record g_t stores the delegate, scope, time limit, conditions, and withdrawal status; $z_t \in \{0, 1\}$ indicates whether the requested delegation is valid for the current decision. The six CCE rights gates are represented by $\mathbf{G}_t$, and the final decision D_t is an element of Ω_t or remains undecided.

The clipping rule is

$$\text{clip}(x) = \min(1, \max(-1, x))$$

All continuous inputs and options supplied to the reference implementation must be finite. A `NaN`, an infinite value, a value outside the permitted range, a duplicate option, or an equal distance tolerance that is negative or not finite is an input error. The implementation stops rather than silently altering invalid input.

4.3 Observations and Person Model Updating

For the baseline decision event, actor k's observation is

$$O_t^k = \text{clip}(W_t + \varepsilon_t^k)$$

where ε_t^k represents observation error or bias. Missing and delayed observations are not folded into this error term; M1 examines them as separate mechanisms. If an observation is missing, the baseline implementation does not update the person model at that time and retains the previous value.

When an observation is available, the person model is updated by

$$M_{t+1}^k = (1 - \lambda_k) M_t^k + \lambda_k O_t^k$$

Thus, $\lambda_k = 0$ retains the previous person model, $\lambda_k = 1$ uses the current observation, and intermediate values combine them. This equation is a minimum explanatory rule for synthetic analysis, not a claim that human understanding, family knowledge, professional judgment, or AI inference changes through this equation. Examples such as $\lambda_k = 0.20$ describe a synthetic learning rate condition, not a recommendation to give current expressions of will only 20 percent weight in care.

4.4 Proposal to Option Mapping

The baseline proposal rule is

$$Q_t^k = \text{clip}(M_{t+1}^k - \beta_k r_t^k)$$

It keeps the actor's person model separate from the actor's synthetic risk estimate but does not represent the full process by which a real person, family member, professional, or AI forms a proposal. The values Q_t^A, Q_t^F, and Q_t^C remain separate and are not averaged, summed, ranked by actor status, or converted into a majority decision.

Each continuous proposal is mapped to its nearest feasible option by

$$\pi(Q) = \arg\min_{\omega \in \Omega} |\omega - Q|$$

When two or more feasible options are equally distant from Q, the reference implementation returns every equally near candidate instead of selecting one by order, sign, or actor preference. Distances are treated as equal only when their numerical difference is at most `1e-12`, that is, 10^{-12}. This tolerance identifies equality in computer arithmetic; it is not a safety interval for real decisions whose choices are merely close in meaning or consequence. If the mapping does not leave one unique final decision, the decision is not forced and the event is recorded as `paused`. M1 results reported for v0.1 use fixed settings in which the nearest option is unique and are not a general decision rule for time series settings containing equal distance candidates.

4.5 Synthetic Response after the Proposal

Only M2 contains a synthetic link from the AI proposal to a response after the proposal. That response is generated by

$$Y_t(a) = \text{clip}\left((1 - a_t^{sim})B_t + a_t^{sim} Q_t^A\right)$$

where B_t is the response before the AI proposal, Q_t^A is the AI proposal, and a_t^{sim} varies the strength of the synthetic coupling. At $a_t^{sim} = 0$, the response remains at B_t; at $a_t^{sim} = 1$, it reaches Q_t^A. For valid input ranges, the weighted expression already remains in $[-1, 1]$, but clip states the intended range explicitly.

This equation neither estimates a real person's response to AI nor distinguishes informed acceptance from pressure, repetition, or other influences; it changes a synthetic response only. M0 does not use this equation. M0 receives the person's response exogenously with source status `exogenous_synthetic_input`; proposals and response are stored in the same decision event without treating the proposals as generating the response.

4.6 Person Priority, Limited Delegation, and Pause

The reference decision rule first checks all six rights gates. Only when every gate is `met` does it apply the person's latest response before considering any prior delegation record. The full hierarchy is

$$D_t = \begin{cases} \perp & \text{if any gate state is } \texttt{unconfirmed} \text{ or } \texttt{violated} \\ d_{person}(\Omega_t, Y_t, \mathbf{G}_t, E_t) & \text{if all six gate states are } \texttt{met} \text{ and } c_t \in \{\text{accept}, \text{modify}\} \\ d_{delegate}(\Omega_t, g_t, \mathbf{G}_t, E_t) & \text{if all six gate states are } \texttt{met} \text{ and } c_t = \text{delegate} \ \wedge \ z_t = 1 \\ \perp & \text{if all six gate states are } \texttt{met} \text{ and } c_t \in \{\text{reject}, \text{defer}, \text{unconfirmed}\} \\ \perp & \text{otherwise} \end{cases}$$

The first branch is evaluated before every response or delegation branch. Here, $\perp$ denotes no final decision and is represented in the reference output as `paused`. For `accept` or `modify`, d_{person}

uses the feasible option confirmed by the person. A previous delegation does not replace that current response. For `reject`, `defer`, or `unconfirmed`, the model does not treat silence or uncertainty as agreement or revive a prior delegated choice; it pauses the irreversible decision for reconsideration. The final `otherwise` branch also returns $\perp$, including an infeasible delegated option or unmet required conditions.

Delegated authority is available only when the current response is `delegate` and $z_t = 1$. Validity requires that the person chose the delegate, the current decision is within the recorded scope, the time limit has not expired, the delegation has not been withdrawn, withdrawal remains possible, and the chosen option is feasible. It also requires the typed input `delegate_is_ai` to be false. The model does not infer whether a delegate is human or AI from a name or label. `delegate_is_ai` is an unverified typed exogenous input, and its source status is `unverified_exogenous_synthetic_input`. A false value therefore does not independently prove that the delegate is human. AI is not an eligible delegate under v0.1.

Family and professionals likewise have no authority over the final decision merely by occupying those roles. The $d_{delegate}$ branch uses a separate delegate record; the modeled family and professional proposal roles are not automatically that delegate. Authority in this branch comes from the person's current delegation, not from family status, professional status, estimated accuracy, or agreement among the three proposing actors.

A `paused` output does not mean safe, neutral, or equivalent to realizing the person's will. In v0.1, the duration and consequences of `paused` are not modeled. The model does not represent what happens during a pause, how often it repeats, or how the person's will and the continuing situation may diverge while no decision is produced. The output therefore records `pause_consequence_status=not_modeled` and `pause_duration_status=not_modeled`. Emergency action is not inserted into the final branch; emergency measures are outside the implemented v0.1 decision rule.

4.7 Six Rights Gates

The six CCE rights gates are separate conditions checked before an irreversible decision. They are `will_and_preferences`, confirming the person's current will and preferences; `support_access`, recording access to needed explanation, expression support, and decision support; `no_undue_influence`, recording avoidance of coercion, concealed information, conflicts of interest, and other undue influence; `proportionate_and_limited`, recording whether any restriction is limited in scope and time; `review_and_appeal`, recording the availability of review, objection, and reconsideration; and `delegation_and_return`, recording withdrawal of delegation and return of decision authority.

Each gate has exactly one of the states `met`, `unconfirmed`, or `violated`. The derived field `all_met=true` is allowed only when all six states are `met`. The gates are not added, averaged, ranked, or decided by majority. No unconfirmed or violated gate can be offset by safety, efficiency, family benefit, or another gate. In such a case, the reference implementation does not produce an irreversible decision and records the event as `paused`.

The states of the CCE rights gates are exogenous synthetic inputs. Version 0.1 does not infer `no_undue_influence` from proposal values, response changes, option loss, or procedural participation records, and it is not a detector of undue influence. Consequently, option loss alone does not automatically pause a decision when all six externally supplied gate states are `met`. The six procedural participation items—`meaningful_contribution`, `comprehensible_information`, `real_options`, `refusal_and_modification`, `expression_support`, and `time_and_reconsideration`—are records and are not inputs to the decision rule. They are not summed. A `deferred` or unresolved item is retained for review and counted as neither achieved participation nor a confirmed violation. Neither

the rights gates nor the procedural participation records are presented as a validated scale, a legal test, or proof of ethical correctness.

4.8 Separated Outcomes

Actor-specific outcomes are outputs kept separate according to the person or other affected part of the care ecosystem. The model records the following structure without converting it into a total score:

$$\mathbf{J}_t = (\mathbf{G}_t, \mathbf{P}_t, A_t, S_t^{safety}, S_t^{participation}, B_t^F, B_t^C, S_t^{stability})$$

Here, $\mathbf{G}_t$ records the six gates, $\mathbf{P}_t$ records the six procedural participation items, and A_t records only agreement between the final decision and current will:

$$A_t = 1 - \frac{|D_t - W_t|}{2}$$

The numerical expression for A_t is evaluated only when D_t is a unique numerical element of Ω_t. When $D_t = \perp$, numerical outcome congruence is undefined and the output remains `null` with status `not_applicable`; it is not assigned zero.

This distance is not a total measure of the CCE person-sovereignty construct. A person may voluntarily modify a proposal or delegate within limits, so distance alone cannot establish whether the decision process protected the person's authority.

For the M0 baseline event, person safety and occupational participation are stored as synthetic inputs, and family burden is stored separately as a synthetic input. Professional burden and ecosystem stability retain output fields but are not assigned invented values. They are recorded as `null` with status `not_parameterized`. In this schema, `null` means `not_parameterized`, not zero. It does not mean no burden, neutral effect, missing benefit, or an average value. The same rule applies to any field without a numerical model in v0.1. The output explicitly records `outcomes.aggregation=none`. Rights gates, procedural participation, agreement between decision and current will, safety, occupational participation, family burden, professional burden, and ecosystem stability must not be summed or used to compensate for one another.

M2 also reports the proposal-influence diagnostics as separate quantities:

$$I_{Y,t}^A = \frac{|Y_t(a) - Y_t(0)|}{2}$$

$$I_{D,t}^A = \frac{|D_t(a) - D_t(0)|}{2}$$

$$L_{\Omega,t}^A = 1 - \frac{|\Omega_t^A \cap \Omega_t^0|}{|\Omega_t^0|}$$

$$B_{R,t}^A = \begin{cases} 1 & \text{if a decision proceeds after rejection or stopping} \\ 0 & \text{if rejection or stopping is respected} \end{cases}$$

Here, Ω_t^A is the option set after presentation of the AI proposal and Ω_t^0 is the option set without that presentation. The binary value $B_{R,t}^A = 1$ records that a decision proceeded after rejection or stopping, whereas $B_{R,t}^A = 0$ records that rejection or stopping was respected. These quantities separate response change, discrete decision change, option loss, and whether rejection or stopping was disregarded. If a discrete decision is `paused` because the nearest option is not unique, $I_{D,t}^A$ is

left undetermined rather than forced to a number. None of these quantities measures trust, human psychology, clinical benefit, or rights compliance, and they are not combined into one number for AI influence.

4.9 M0, M1, and M2 Separation

M0 is a single decision event module with output schema `CCE-M0-output-v0.1`. It records the decision context, feasible options, W_t, the three observations, updated person models, proposals and mapped options, an exogenous synthetic person response, delegation, six exogenous rights gate states, the decision rule based on the latest response, procedural participation records, and separated outcomes. M0 therefore does not model the person's response as a consequence of AI, family, or professional proposals and does not implement endogenous interaction among four actors.

M1 isolates updating over time. It examines learning rates, lasting change, temporary shocks, noise, missing observations, bias in a fixed direction, and observation delay through observation, person model update, proposal, and mechanisms that pass values to the next time step. It does not join those trajectories to a sequence of M0 decisions, delegated choices, updates to rights gate states, or feedback from actor-specific outcomes. The implemented M1 results assume a unique nearest option in their fixed settings.

M2 isolates the synthetic relation between proposal and response and the associated rights cases. It is the only v0.1 module that uses the AI proposal to generate a synthetic response after the proposal. It compares synthetic conditions with AI and without AI, separates continuous response change from discrete decision change, records option loss and whether rejection or stopping was disregarded, and passes the generated response through the decision rule under externally supplied rights conditions. M2 does not show that real people respond in this way.

This separation makes each result traceable to a limited equation or rule and prevents v0.1 from being described as a connected long-term simulation of four interacting actors. M0, M1, and M2 inspect different mechanisms. Their coexistence does not establish a feedback process among all actors, empirical validation, clinical validation, legal validity, ethical correctness, or external validity.

5 Analytical Predictions and Synthetic Experiments

All inputs, parameters, observations, responses, and outcomes were synthetic values for examining equations, decision rules, settings, and implementation; none came from people, patients, care records, or clinical services. The experiments therefore addressed verification of the specified model, not validation against human behavior, care practice, clinical outcomes, legal requirements, ethical judgments, or external settings [28][29]. The conditions were internally fixed before execution; no external registry was used. No coefficient was estimated or calibrated from human data. Model logic, inputs, seeds, implementation, and outputs were recorded for traceable reporting and computational reproducibility [42][43][44][45]. Here, computational reproducibility meant obtaining the same calculation from the same code, inputs, environment, and procedures. Internal regeneration within the authoring and development workflow was not independent human replication.

The separation among M0, M1, and M2 was preserved. M0 recorded one decision event containing proposals from AI, family, and professional components together with an exogenous synthetic person response; it did not generate that response from the proposals. M1 examined time-varying updates of the person model under permanent change, temporary observation, noise, missingness, systematic bias, and delay. M2 alone connected an AI proposal to a synthetic response and ex-

amined the proposal-influence diagnostics and CCE rights gates. The modules did not form an endogenous long-term interaction among four actors.

5.1 H1a and the H2 learning rate trade-off

M1 used the update rule

$$M_{t+1}^k = (1 - \lambda_k)M_t^k + \lambda_k O_t^k$$

for actor $k \in \{A, F, C\}$. For a constant observation O, repeated application gives

$$M_n = O + (1 - \lambda)^n (M_0 - O).$$

This expression supplied the analytical basis for H1a and H2. H1a was an analytical verification target for a model-implied property: a lower learning rate leaves a larger person model error for longer after a permanent change because the remaining error is multiplied by $(1 - \lambda)^n$. The deterministic experiment checked whether the iterative implementation matched the closed expression and fixed error ordering; disagreement with either would fail the H1a implementation check. Under the selected noise distribution, the implementation check concerned the ordering of median errors across the 100 fixed random series; failure of that ordering would reject only the stated ordering under those noise settings.

H2 was separated into H2a and H2b. H2a was an analytical verification target for a model-implied property: a higher learning rate produces a larger person model movement toward a temporary observation. The deterministic experiment checked whether the iterative implementation matched the analytical update and fixed movement ordering; a mismatch with either would fail the H2a implementation check. Under noise, the implementation check again concerned the ordering of medians under the selected distribution and fixed series.

H2b concerned recovery only after normal observation resumed. Its prediction was limited to shock length 2 and recovery tolerance 0.10: under those fixed conditions, a higher learning rate was predicted to require fewer steps to return within tolerance. This was not a prediction for every shock length, observation window, or tolerance. A condition could remain outside tolerance through the finite window, or fail to leave tolerance during the shock, when the tolerance changed. H2b would be falsified if the expected recovery ordering did not hold at shock length 2 and tolerance 0.10. Recovery was not measured under noise.

The H1a and H2a model-implied properties and the scoped H2b recovery check therefore define a trade-off within the model. A lower learning rate retains more of the earlier person model and adapts more slowly after a lasting change. A higher learning rate follows a temporary observation more strongly and, under the fixed H2b conditions, also returns more quickly after the temporary observation ends. No universally optimal learning rate was claimed. The illustrative value $\lambda = 0.20$, described as retaining 80% of the earlier person model and incorporating 20% of the current observation in one update, was only a synthetic learning rate example. It was not a clinical recommendation, a rule for discounting current will and preferences, or a value derived from people.

The H1a permanent-change conditions used 30 time steps, a change at step 10, a synthetic current will of -0.8 before the change and $+0.8$ after it, a post-change observation of $+0.8$, an initial person model of -0.8, learning rates $\{0.05, 0.20, 0.50, 0.80\}$, a recovery tolerance of 0.10, and a risk deduction of 0. These settings produced four H1a condition trajectories.

The H2 temporary-observation conditions used 25 time steps, a stable reference and normal observation of $+0.8$, a temporary observation of -0.8 at steps 10 and 11, an initial person model of $+0.8$, learning rates $\{0.05, 0.20, 0.50, 0.80, 1.00\}$, a recovery tolerance of 0.10, and a risk deduction

of 0. The shock interval was recorded as `shock_start=10` and `shock_end_exclusive=12`. These settings produced five H2 condition trajectories.

The deterministic M1 suite therefore contained four H1a conditions plus five H2 conditions, giving 9 condition trajectories and 245 rows across time points. Across the two modules there were five distinct learning-rate values; H1a used four of them, and H2 used all five. Error, proposal disagreement, and recovery state were stored separately. M1 decision comparisons were limited to the current fixed settings in which the reference value and proposal had a unique nearest option. They did not define a general rule for equal distance options. The separate equal distance check treated differences of at most 10^{-12} as computational equality, returned all nearest options, and withheld a unique decision; this tolerance was not a safety band for real care.

5.2 M0 baseline verification

The M0 baseline provided a fixed hand calculation target for the decision target `outing_method`. The synthetic current will was $W = 0.80$, and the option set was $\Omega = \{-1, 0, 1\}$. The exact actor-specific inputs were

Actor	M_0^k	ε^k	λ_k	r^k	β_k
AI	0.40	−0.20	0.50	0.60	0.50
Family	0.10	−0.60	0.20	0.80	0.90
Professional	0.50	−0.10	0.40	0.50	0.40

The corresponding observations, risk deductions, updated person models, proposals, and nearest options were

Actor	O^k	$\beta_k r^k$	M_1^k	Q^k	$\pi(Q^k)$
AI	0.60	0.30	0.50	0.20	0
Family	0.20	0.72	0.12	−0.60	−1
Professional	0.70	0.20	0.58	0.38	0

These were verification targets, not estimates of actor groups. A mismatch in an intermediate value, response provenance, decision basis, gate state, or aggregation state would fail the baseline check.

The person response was separately supplied as the exogenous synthetic input `modify(0)`. The prior delegation record covered only `departure_time`, not `outing_method`, giving $z = 0$, but $D = 0$ followed from the current modification rather than proposal averaging. The typed input `delegate_is_ai` was an unverified exogenous synthetic input, and AI was not an eligible delegate. Expected outcome congruence was 0.60. The six CCE rights gates—`will_and_preferences`, `support_access`, `no_undue_influence`, `proportionate_and_limited`, `review_and_appeal`, and `delegation_and_return`—were individually `met`, with `all_met=true`. The procedural participation records `meaningful_contribution`, `comprehensible_information`, `real_options`, and `refusal_and_modification` were `met`; `expression_support` and `time_and_reconsideration` were `unconfirmed`. These participation records were stored but were not inputs to the decision rule.

The actor-specific outcomes were person safety 0.75, occupational participation 0.80, and family burden 0.40. Professional burden and ecosystem stability were `null` with status `not_parameterized`; `null` meant not parameterized, not zero. The required record was `outcomes.aggregation = "none"`. Rights gate states, participation records, outcome congruence, safety, occupational participation, and actor-specific burden were not converted into a composite score or a vote.

5.3 Noise and observation failure experiments

The M1 noise experiment reused the nine deterministic condition structures and used seeds 41000–41099. Noise was Gaussian with mean 0 and standard deviation 0.1 and was clipped to $[-1, 1]$. The design used 100 independent fixed series and common random numbers across the corresponding 9 conditions. It therefore produced 900 rows, equal to 100 series $\times$ 9 conditions, not 900 independent series. Stored aggregation used the median and 0.05/0.95 quantiles. Recovery was not measured in the noise runs.

Observation failures were examined separately over 30 time steps, with a change at step 10, a synthetic current will changing from -0.8 to $+0.8$, an initial person model of -0.8, a learning rate of 0.50, and a risk deduction of 0. The experiment used seeds 52000–52099 and the same Gaussian noise form. The 100 independent fixed series used common random numbers across corresponding conditions. The nominal design displayed 12 conditions and 1,200 rows because the zero strength control was repeated under each of three mechanisms. Removing those repeated controls left 10 effective conditions and 1,000 nonduplicate rows. Aggregation again used the median and 0.05/0.95 quantiles.

Missingness, systematic bias, and delay remained distinct mechanisms. The rates under missing completely at random were 0, 0.1, 0.3, and 0.5. When an observation was missing, no update occurred and the preceding person model was retained for that step. Fixed negative bias offsets were 0, -0.2, -0.4, and -0.8 in the scenario where current will changed from -0.8 to $+0.8$; the offsets moved observations negatively away from the new will. This directional scenario did not support a claim that every form or direction of bias must increase error. Delays were 0, 1, 3, and 5 steps and were recorded as delayed observations rather than recoded as missingness or fixed bias. Outputs retained separate records for each mechanism, including missingness rate, cumulative error, and maximum error. These selected checks were scoped robustness examinations, not a complete sensitivity or uncertainty analysis over the full input space [46].

5.4 H1b-1: synthetic response and decision effects

M2 used

$$Y_t(a) = \text{clip}\left((1 - a_t^{sim})B_t + a_t^{sim} Q_t^A\right),$$

where $a_t^{sim} \in [0, 1]$ was a synthetic parameter for coupling the proposal and response, not AI authority, a weight assigned to a person, or an observed human coefficient. Continuous response change and discrete decision change were measured separately:

$$I_{Y,t}^A = \frac{|Y_t(a) - Y_t(0)|}{2},$$

$$I_{D,t}^A = \frac{|D_t(a) - D_t(0)|}{2}.$$

For the fixed synthetic equation,

$$I_Y^A = a\frac{|Q - B|}{2}.$$

H1b-1 was an analytical verification target, and the baseline and robustness experiments were implementation checks of its model-implied properties. When the initial response and AI proposal were in opposite directions while the option set and rights conditions were fixed, increasing

a_t^{sim} would not decrease the continuous response effect, the discrete decision effect, or decision disagreement with the initial response. Flat intervals in the discrete decision were allowed because a changing continuous response can remain nearest to the same option. Thus, an increase in I_Y^A without an increase in I_D^A was not a falsification. At an equal distance point, the implementation returned all nearest options, withheld a unique decision, and did not force a numerical I_D^A. Failure of the continuous effect to follow the analytical expression, or a decrease in a decision comparison where both decisions had unique nearest options, would fail the H1b-1 implementation check.

The H1b-1 baseline fixed $\Omega = \{-1, 0, 1\}$, $B = 1$, $Q^A \in \{-1, 0, 1\}$, and the coupling grid $a_t^{sim} \in \{0.0, 0.1, \ldots, 1.0\}$. All six CCE rights gate states were supplied as `met`. The three proposal values and 11 coupling values produced 33 conditions. Eleven conditions used the opposite-direction setting $B = 1$ and $Q^A = -1$, and the remaining 22 were controls. The 11 opposite-direction conditions excluded the hand calculation equal distance points $a = 0.25$ and $a = 0.75$; equal distance behavior was checked separately through one condition with an intermediate proposal and the rights and option cases.

The robustness checks used the same coupling grid from 0.0 to 1.0 in increments of 0.1. Baseline sensitivity used $B \in \{1.0, 0.75, -0.75, -1.0\}$ with $Q^A = -B$, giving $4 \times 11 = 44$ conditions. Option-set sensitivity used $B = 1$ and $Q^A = -1$ with three recorded option sets, giving $3 \times 11 = 33$ conditions. The robustness extension therefore contained 77 conditions. Raw responses, discrete decisions, and validity states were retained separately.

5.5 H1b-2: rights and option cases

H1b-2 was a scoped check of the decision rule and rights cases: when refusal, stopping, and modification were respected, and an unconfirmed response did not authorize a decision, the AI proposal alone would not automatically produce an irreversible final decision. The cases examined response states, CCE rights gate states, equal distance handling, and option set changes while keeping response effect, decision effect, option loss, and refusal blocking separate. Option loss and refusal blocking were recorded as

$$L_{\Omega,t}^A = 1 - \frac{|\Omega_t^A \cap \Omega_t^0|}{|\Omega_t^0|}$$

and

$$B_{R,t}^A = \begin{cases} 1 & \text{if a decision proceeded after refusal or stopping,} \\ 0 & \text{if refusal or stopping was respected.} \end{cases}$$

The CCE rights gate states were exogenous synthetic inputs; they were not inferred from proposal-influence diagnostics or procedural participation records. H1b-2 would fail if an unconfirmed response advanced an irreversible decision, refusal or stopping was bypassed, a modification was overwritten, an unconfirmed or violated gate permitted the decision, or an equal distance option was silently selected as unique.

The suite contained 8 cases with mixed provenance: an original fixed set of 7 cases plus one diagnostic case added after review. The cases comprised confirmed modification, rejection, deferral, an unconfirmed response, a violated `no_undue_influence` gate, narrowed options with a violated gate, narrowed options with all six supplied gates `met`, and one unsafe negative control in which rejection was bypassed. The case with narrowed options and all supplied gates `met` was the diagnostic case added after review; the other seven cases belonged to the original fixed set.

The additional diagnostic case isolated option narrowing while every rights gate was externally supplied as `met`. It examined a boundary of the current rule: option loss can be recorded, but narrowing alone does not stop a decision when the exogenous gate inputs remain `met`. The model does not infer that such inputs are internally inconsistent. Option loss and refusal blocking therefore remained separate diagnostics and were not combined with response effect, decision effect, or outcome congruence. When a decision was `paused`, the stored states were `pause_duration_status=not_modeled` and `pause_consequence_status=not_modeled`; pausing was not evidence of safety, neutrality, or realization of the person's will.

The core evidence comprised the M0, M1, and M2 analytical and synthetic checks and 96 core tests. A further 41 tests preserved technical preparation for future H1c-R, giving 137 total tests. H1c-R would compare one versus three presentations of the same AI proposal to real participants. Those 41 tests were not core CCE-MVM v0.1 evidence, contained no human results, and did not establish empirical, clinical, legal, ethical, or external validity.

6 Results

The results are reported by module because CCE-MVM v0.1, the CCE Minimum Viable Model, did not run M0, M1, and M2 as one endogenous simulation of interactions among four actors. M0 recorded one decision event with an exogenous synthetic person response; M1 examined time-varying updates of the person model; and M2 examined the specified synthetic relation between an AI proposal, a response, and a discrete decision. The reported values describe implemented equations, fixed synthetic conditions, and saved outputs rather than observed human behavior. The implementation conformance, test counts, regeneration, and output comparison reported in this section are author-reported internal verification results for unreleased artifacts.

6.1 M0 hand calculation and reference implementation

The M0 reference implementation reproduced the complete hand calculation for the fixed decision event about the outing method. The decision target was `outing_method`. The synthetic current will was $W = 0.80$. The AI, family, and professional observations were 0.60, 0.20, and 0.70, respectively. After the specified updates, their person model values were 0.50, 0.12, and 0.58, and their continuous proposal values were 0.20, -0.60, and 0.38. Mapping those proposals to the available options returned 0 for the AI, -1 for the family, and 0 for the professional. These proposals were retained as separate records and were not averaged or counted as votes.

The exogenous synthetic person response was `modify(0)`. The earlier delegation record was outside the scope of the decision, so $z = 0$, and the final decision was $D = 0$ through the person response branch. The recorded agreement between the final decision and current will was 0.60. The separately recorded synthetic outcomes were person safety 0.75, occupational participation 0.80, and family burden 0.40; their values and directions were explanatory synthetic fixtures. Professional burden and ecosystem stability were both returned as `null` with status `not_parameterized`; they were not assigned zero values. All six CCE rights gates were supplied exogenously as `met`. Four procedural participation records—`meaningful_contribution`, `comprehensible_information`, `real_options`, and `refusal_and_modification`—were `met`, whereas `expression_support` and `time_and_reconsideration` were `unconfirmed`. The deadline and other environmental constraints that had not been parameterized were also retained as `null`/`not_parameterized`. Thus, agreement between the hand calculation and code covered intermediate observations, updated person models, proposals, option mapping, the response branch, the final decision, and the separated outputs; it did not turn the recorded items into a composite result.

6.2 H1a, H2a, and H2b

The fixed H1a/H2 experiment served as an implementation check and comprised nine conditions and 245 rows across time. H1a was an analytical verification target for a model-implied property: lower learning rates retained a larger person model error for longer after a permanent change in the synthetic current will. The iterative implementation agreed with the closed form of the update: when the observation remained at the new value, the remaining difference was multiplied by the retained fraction at each update. This was a structural property of the specified linear update under fixed conditions, not an estimate of how quickly any person, family member, professional, or AI learns.

H2a was likewise an analytical verification target for a model-implied property and was checked separately in the implementation. Under a temporary observation shock, higher learning rates produced a larger immediate movement of the person model toward the shocked observation. Taken together, H1a and H2a did not produce a single ordering in which a larger or smaller learning rate was uniformly preferable. Increasing the learning rate reduced persistence of error after a permanent change but increased movement toward a temporary observation. The opposing results were reproduced in the saved trajectories.

H2b concerned recovery after the temporary shock and had a narrower boundary than H2a. H2b was evaluated only at the fixed shock length of two steps and the fixed recovery tolerance of 0.10. Under those conditions, higher learning rates required fewer steps to return within the tolerance after ordinary observations resumed. This ordering is limited to that shock length, tolerance, finite observation period, and definition in discrete steps. When the tolerance was changed, the categorical recovery result did not retain a simple universal ordering: some settings were recorded as not recovered within the observation period, whereas in others the trajectory never left the tolerance band and therefore had no recovery interval to count. Recovery was not measured under noise. H2a and H2b therefore report different outputs and are not combined into one statement about recovery under uncertainty.

Figure 3 visualizes the distinct H1a, H2a, and H2b patterns under the fixed deterministic conditions.

6.3 Noise and observation failures

The noise experiment served as a scoped implementation check and contained 900 condition rows formed by applying the same 100 independent random series across nine conditions; it did not contain 900 independent series. The saved summaries retained the median and the 5th and 95th percentiles for each condition. Within the selected setting with normal noise, the median H1a ordering remained: lower learning rates retained error for longer after the permanent change. The median H2a ordering also remained: higher learning rates moved farther toward the temporary shock. These implementation results are conditional on the selected distribution, parameter values, observation window, and use of the same random series across conditions. The percentiles were not treated as estimates from observed data. Recovery was not calculated from these runs.

The observation failure experiment contained nominally 1,200 condition rows from 100 independent series across 12 displayed conditions. Because the control at intensity zero was repeated for missing observations, systematic bias, and observation delay, the dataset contained 1,000 nonduplicate condition rows across ten effective conditions. The experiment kept the three mechanisms separate and saved outputs for each mechanism such as missing rate, cumulative error, and maximum error. The mechanisms are reported separately; no combined numerical effect or common ordering is claimed. In particular, the systematic bias condition represented only a bias in the

Figure 3. Learning rate creates a lag-over-following trade-off

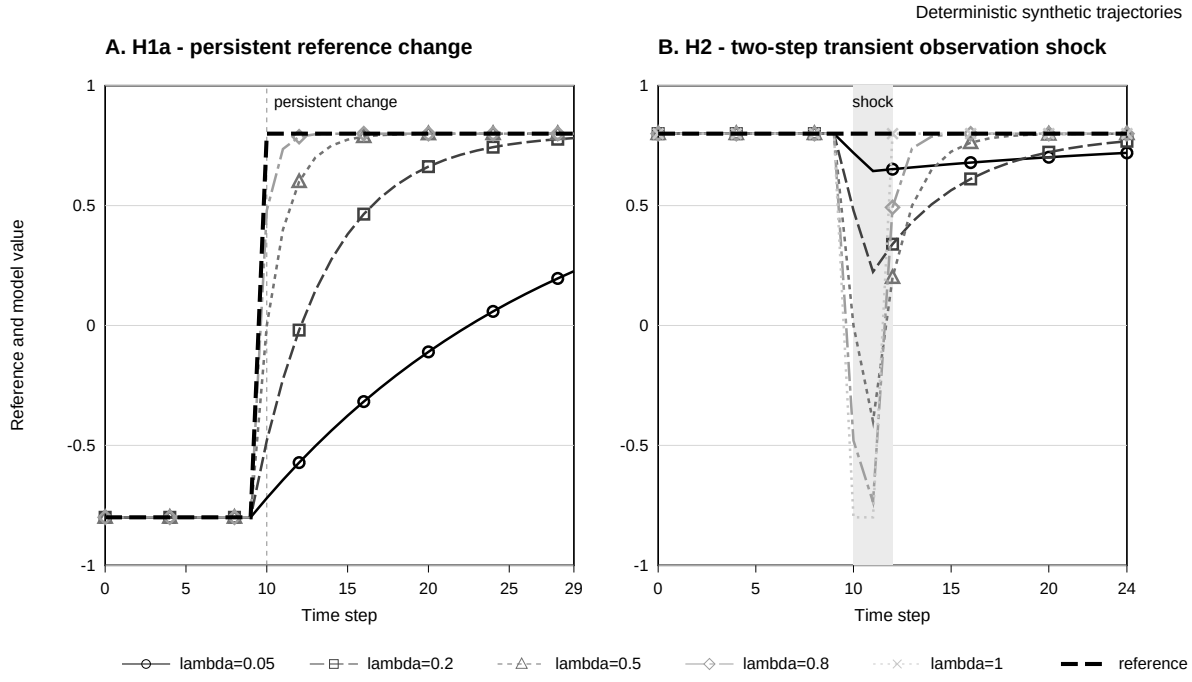


Figure 3: Deterministic synthetic trajectories under five fixed learning rates. In H1a, smaller learning rates update more slowly after a persistent change. In H2, larger learning rates follow the observation shock lasting two steps more strongly and then recover more quickly under the fixed tolerance. The dashed black line is the reference. These trajectories do not estimate human learning rates. The apparent trade-off is conditional on the equation defined in this study, input sequence, and parameter grid.

negative direction in a fixed change of the synthetic current will from -0.8 to $+0.8$. The results therefore do not support the broader claim that bias of any direction must increase error. The delay findings likewise belong to the fixed observation windows used in the experiment rather than to all possible delays.

6.4 H1b response and decision results

H1b-1 was an analytical verification target, and the baseline experiment served as an implementation check of its model-implied response and decision properties. The experiment separated continuous response change from discrete decision change and included 33 coupling conditions: 11 conditions in which the initial response and AI proposal were in opposite directions and 22 control conditions. In the conditions with the initial response and proposal in opposite directions, the continuous response quantity I_Y^A followed the analytic relation $I_Y^A = a\frac{|Q-B|}{2}$ and did not decrease as the synthetic coupling parameter increased. The separate discrete decision quantity did not change continuously. It remained unchanged within regions mapped to the same available option and changed only when the response crossed a boundary between nearest options. Consequently, a nonzero continuous response effect could coexist with no discrete decision change. These flat regions were expected from the option mapping, were reproduced by the implementation, and were not interpreted as absence of response change.

States with equal distance were also kept separate from ordinary decision changes. When multiple options were equally near, the implementation returned the tied candidates and paused rather than selecting one automatically. The 11 main conditions with the initial response and proposal in opposite directions did not pass through the exact equal distance points of 0.25 and 0.75. Equal distance was directly exercised in the intermediate proposal H1b-1 condition and in a separate decision rule test. An undefined or paused decision was not converted into a numerical decision effect.

Figure 4 shows how continuous synthetic response change can precede a change in the discrete mapped decision.

H1b-2 was a scoped check of the decision rule and rights cases, with eight saved cases of mixed provenance. The original seven cases were internally fixed before execution; no external registry was used. One additional option narrowing case was added after review as a diagnostic case, and its later origin was retained in the metadata. The eight saved cases covered confirmed `modify`, `reject`, `defer`, `unconfirmed`, a violated undue influence gate, option narrowing with a violated gate, option narrowing with externally supplied `met` gates, and the isolated unsafe negative control that bypassed rejection. The option narrowing case with externally supplied `met` gates was the diagnostic case added after review. The isolated unsafe negative control that bypassed rejection was already among the original seven. The eight saved cases contained neither an `accept` case nor a `delegate` case.

The confirmed `modify` case returned the feasible option identified by the current person response. The `reject`, `defer`, and `unconfirmed` cases returned `paused`. The case with a violated undue influence gate and the option narrowing case with a violated gate also returned `paused`. The isolated unsafe negative control that bypassed rejection produced the intended failure by allowing a decision to proceed after rejection; it was retained as a negative control rather than presented as the behavior of the ordinary v0.1 decision rule. By contrast, option narrowing alone did not pause the decision when all externally supplied rights gate states remained `met`. The implementation recorded the option loss but did not infer that the supplied gate states were internally inconsistent. This outcome is consistent with the v0.1 boundary: the rights gate states are exogenous synthetic inputs, procedural participation records are not inputs to the decision rule, and the model does not detect undue influence from other outputs.

Separate core decision rule tests, rather than the eight saved H1b-2 cases, exercised `accept` and

Figure 4. Continuous response change can cross discrete decision thresholds

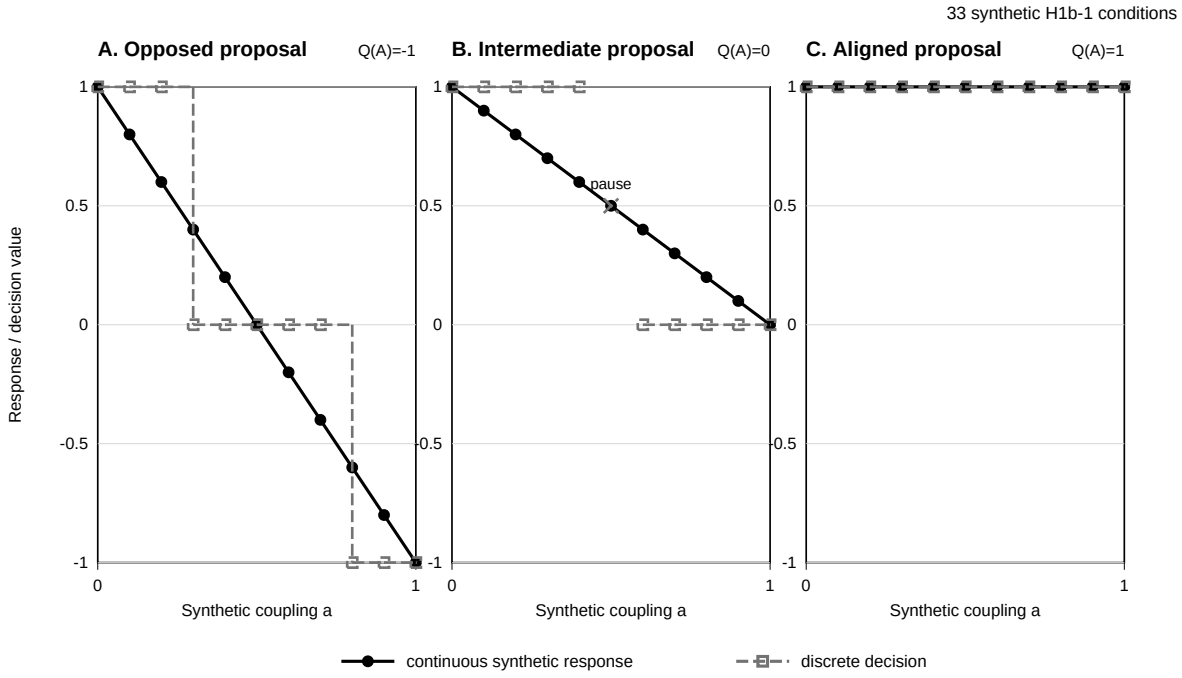


Figure 4: Continuous synthetic response after the proposal and corresponding discrete decision across 33 H1b-1 coupling conditions. The gray step line shows mapping to the nearest option in the fixed set $\{-1, 0, 1\}$. A tie between options at equal distance is paused rather than broken automatically. All values are synthetic and are not estimates of human behavior or intervention effects. The threshold crossings are consequences of the fixed discretization rule, not clinical decision thresholds.

delegation behavior. Those tests showed that `accept` and `modify` use the current person response. A `delegate` response entered the delegation branch only when the current response was `delegate`, the delegation record was valid for the decision, `delegate_is_ai=false`, and the delegated option was feasible. Invalid or infeasible delegation returned `paused`, and a prior delegation did not overwrite a current `reject`, `defer`, or `unconfirmed` response. Across the separate 77 H1b robustness conditions—44 conditions that varied the initial response and 33 conditions that varied the option set—the raw responses, mapped decisions, and condition checks matched the fixed expectations. This agreement concerns the specified synthetic equation and option mapping, not human response to AI advice.

For every H1b case that returned `paused`, the duration and consequences of the pause remained `not_modeled`. A paused state therefore did not establish safety, neutrality, or realization of the person's current will. The typed `delegate_is_ai` field was also retained as an unverified exogenous synthetic input; the implementation did not independently establish the identity of a delegate. Emergency measures and an integration over time of M0, M1, and M2 were not implemented in v0.1.

6.5 Model tests and internal regeneration

All 96 core model tests passed in the author-reported internal verification. They covered the M0 reference event and output schema, analytic relations, H1a/H2 conditions, noise, observation failures, and H1b baseline and robustness conditions. The complete model test discovery returned 137 passing tests because it also included 41 tests that preserve materials for a future human study. Those 41 tests were not counted as core evidence for CCE-MVM v0.1. The counts of 96 and 137 refer only to the model test suite and exclude later tests of workflow tools and later checks of the package and bibliography.

Internal regeneration was completed for six experiments as part of the author-reported internal verification: the M0 reference event, H1a/H2 without noise, observation noise, observation failures, the H1b baseline experiment, and H1b robustness. In that internal run, the 17 canonical CSV/JSON outputs regenerated in a new temporary folder were identical at the byte level to the saved canonical files. The CSV and JSON files therefore remained the computational records for this comparison. The saved H1b spreadsheet files were not regenerated because a corresponding generation process was not confirmed, so they remained viewing files that were not canonical computational records. This was internal regeneration using the controlled code, settings, and author's environment. It was not independent replication, did not establish external validity, and did not validate the model against people or care settings.

7 Discussion

The main result of CCE-MVM v0.1 is not an improved care decision but the explicit separation of elements that can remain implicit when AI produces a recommendation: the person's currently expressed will, the CCE person model held by each other actor, each proposal, the person's response, decision authority, the CCE rights gates, and the actor-specific outcomes. The CCE person model is a simplified state held separately for the AI, family, and professional actors; it is not the person or a complete representation of the person. A CCE rights gate is one of six separately recorded decision conditions. Actor-specific outcomes keep safety, occupational participation, family burden, professional burden, and ecosystem stability in separate fields. This structure distinguishes agreement on a final option from agreement about the person, the process, or the authority by which the decision was produced.

The separated person models make differences in information inspectable. In M0, the AI, family, and professional actors observe the same current will through separate observation errors, update separate person models, and generate separate proposals. A proposal does not acquire decision authority because it is accurate, because other actors agree, or because its continuous value lies near the final option. In the baseline event, the arithmetic mean of the three proposal values was approximately -0.007, whereas the final decision was 0 because the person's exogenous response was `modify(0)`. The same numerical endpoint would have a different meaning if produced by voting or averaging. M0 therefore records observation, proposal, response, and authority separately. The person's response is an exogenous synthetic input in M0; only M2 calculates a relation between the AI proposal and a synthetic response. The results do not show that any other actor caused the person's response.

The categorical response states make authority conditional on what the person did, not merely on a continuous response value. In the reference decision rule, `accept` and `modify` can support a person decision; `reject`, `defer`, and `unconfirmed` lead to `paused`; and `delegate` permits a delegate decision only within a valid, limited delegation. A past delegation record does not override the latest response. AI is not an eligible delegate, although `delegate_is_ai` remains an unverified exogenous input rather than an identity check performed by the model. The six CCE rights gates are recorded separately as `met`, `unconfirmed`, or `violated`, and every gate must be `met` before an irreversible decision is returned. These are rights gates that cannot be offset by other outcomes. The CCE person-sovereignty construct is therefore a pattern of conditions, responses, authority records, and outcomes rather than a single score. The rights gate states are exogenous synthetic inputs, and the procedural participation records are not inputs to the decision rule. The implementation can act on supplied gate states, but it does not infer undue influence from the interaction itself. The duration and consequences of `paused` are not modeled.

Keeping actor-specific outcomes separate changes what can legitimately be called favorable. Greater outcome congruence, a safer synthetic option, or lower family burden cannot erase a gate recorded as `unconfirmed` or `violated`. Conversely, lower outcome congruence does not establish a rights failure, because the person may have revised an option or chosen limited delegation. Professional burden and ecosystem stability remain `null` with the status `not_parameterized`; these entries do not mean zero or neutrality. The model preserves unresolved differences instead of forcing safety, participation, burden, rights, and agreement into one composite. This is a structural decision to keep distinct questions available for review, not a finding that one outcome is universally more important.

H1a and H2a are analytical verification targets for model-implied properties of the specified update equation. Together, they define a learning rate trade-off within that equation. Their synthetic runs are implementation checks, not independent experimental evidence for those algebraic consequences, and they do not select a universal optimum. H1a concerns sustained change: lower λ preserves the previous person model longer and prolongs error after the current will changes persistently. H2a concerns temporary observation change: higher λ produces a larger immediate displacement during the fixed shock. H2b is a recovery check limited to the specified conditions after that shock. The H2b ordering is specific to the tested conditions: a shock length of 2, a recovery tolerance of 0.10, a finite observation period, and the specified discrete update steps. Under those conditions, higher learning rates returned to the tolerance band in fewer steps after ordinary observations resumed. Recovery was not measured under noise. A higher learning rate can reduce delay after sustained change while increasing response to a temporary observation; a lower rate can resist the temporary observation while retaining an outdated model. Choosing an optimum would require external judgments about the consequences of those errors, observation quality, and opportunities for review. The 80/20 example is only a synthetic learning rate example, not a clinical recommen-

dation. The tested parameter values also do not justify claims across an unexamined input space [46].

M2 contains two distinct checks. H1b-1 is an analytical verification target for a model-implied proposal–response property, whereas H1b-2 is a scoped check of the decision rule and rights cases. The H1b-1 implementation checks show why continuous response change and discrete decision change must remain separate. Under the specified response equation, the continuous response changes with the synthetic coupling parameter a_t^{sim}, but the final decision is obtained only after mapping that response to the available discrete options. With an initial response of 1 and an AI proposal of -1, coupling 0.20 moved the response to 0.60 while the decision remained 1. At 0.50, both became 0; at 0.80, the response was -0.60 and the decision was -1. A proposal can therefore change the response before it changes the selected option, while a small additional change near an option boundary can change the decision abruptly. When two options are at the same distance, as at 0.25 and 0.75 here, the implementation returns both candidates and pauses rather than choosing automatically. Continuous response change and discrete decision change answer different questions and should not be combined. These patterns are implementation checks of the H1b-1 model-implied property and the specified option mapping; they are not independent experimental evidence for the algebraic relation and do not demonstrate human response to AI advice. H1b-2 separately checks the decision rule and rights cases within its stated scope.

The separation of M0, M1, and M2 improves inspectability but reduces realism by omitting feedback among modules. M0 records one decision event, response, delegation, rights gates, authority, and separated outcomes. M1 isolates change in person models over time under learning rates and the tested conditions for noise, missing observations, bias in one direction, and delay. M2 isolates the relation between an AI proposal and a synthetic response, together with decision change and selected rights cases. Unexpected results can therefore be traced to a smaller set of assumptions, while equations, implementation, tests, configurations, and saved outputs can be checked against one another. The 96 core tests, 137 total tests, and internal regeneration of 17 canonical CSV/JSON outputs support that implementation claim, but not independent replication. The omitted feedback is visible in the boundaries: M0 does not generate the person's response from the proposals, M1 does not feed actor-specific outcomes into later decisions, and M2 does not generate changing family or professional responses. The modules do not implement endogenous interaction among all four actors, repeated life events, changing organizational constraints, or a complete changing care network.

This scope places CCE-MVM v0.1 within existing accounts of care. The dynamic care ecosystem already describes care as a changing arrangement of people, needs, relationships, and technologies that may include AI [1]. Research on AI in care networks also considers assistance, affiliation, and relations among older adults and formal or informal supporters [2, 3, 4, 7]. Distributed cognition locates decision activity across people, artifacts, and work settings, while sociotechnical systems examine people, tasks, technology, physical settings, and organizations together [6, 5]. CCE-MVM does not originate these perspectives. Its narrower contribution is to place selected distinctions—separate person models, categorical responses, authority, rights gates, and actor-specific outcomes—into a small specification that can be inspected and tested. Because the modules omit much of the structure, change, and feedback of an actual care network, they are not a computational reproduction of an entire dynamic care ecosystem.

Relational autonomy recognizes that relationships can support autonomy but can also constrain it [8, 9]. The distinction between supporting and replacing a person's decision, together with attention to will and preferences, undue influence, and review, is also present in accounts of supported decision-making [11, 12]. CCE responds by separating proposal from authority and by refusing to let another outcome offset a rights gate. That is a formal design response, not evidence that the

gates classify real interactions accurately or establish legal compliance. The Capability Approach directs attention beyond resources or achieved outcomes toward the opportunities and freedoms a person can exercise, and it has been applied to accounts of AI benefit and assistance [26, 27]. CCE's option set, occupational participation field, and separated outcomes are compatible with asking such questions, but v0.1 contains neither a validated capability measure nor evidence that its synthetic options are genuinely available or understandable in practice.

The comparison with human–AI collaboration is similarly limited. Correct AI advice can improve judgment in some experimental tasks, while incorrect advice can mislead professionals, and explanations do not always prevent that error [13, 14, 15]. Reviews indicate that much health care evidence remains concentrated in diagnostic or short-term tasks, with less evidence on longitudinal, patient, and system outcomes [39]. CCE does not test whether collaboration improves performance; it adds separate questions about response change, decision change, authority, rights conditions, and actor-specific outcomes. Mathematical shared decision-making models already distinguish parties' information and preferences, while other computational work integrates multiple input sources with learned weights or examines mutual change between humans and AI over time [19, 24, 25]. CCE's difference lies in its particular modular separation and refusal to absorb rights conditions and actor outcomes into one decision value, not in inventing dynamic influence. Model names should follow the states, rules, interactions, and feedback actually implemented [35]. CCE-MVM v0.1 is therefore not a Human Digital Twin: it lacks data from a real person, repeated updating against such data, prediction specific to that person, and feedback into real decisions [40].

The results establish bounded computational feasibility and internal verification, not human causality, empirical validation, clinical or ethical benefit, safety, legal validity, or external validity. Verification of implementation must remain distinct from validation against the intended reality [28]. Internal regeneration used the project's own code, fixed conditions, and development process; independent replication by a separate researcher in a fresh environment has not been completed [45]. Future work should obtain human review from people with lived experience, family perspectives, care professionals, and reviewers with expertise in quantitative methods, reproducibility, rights, ethics, and law. Empirical validation should test whether response states, person model records, delegation records, rights conditions, and actor outcomes can be observed and interpreted consistently, without assuming that synthetic parameters are human coefficients. Independent replication should run the fixed package without developer intervention and compare the core tests, major experiments, and canonical outputs. Modeling life trajectories would require connecting repeated decisions, changing relationships, resources, and consequences while continuing to distinguish the person's own meaning from inferences made by other actors. Finally, a future comparative study should evaluate a system designed using CCE against a default AI system. Prediction accuracy, continuous response change, discrete decision change, participation records, rights gate states, and actor-specific outcomes should remain separate rather than being combined into one composite. Such a study would test whether the design changes any observed process or outcome; the present synthetic results do not predict that it will.

8 Rights, Ethics, Misuse, and Limitations

CCE-MVM v0.1 is a research computational model for examining information, support, options, authority, and review surrounding a decision. It neither evaluates a person's worth nor determines legal status, capacity, or consent. This section does not provide legal or clinical advice. The model has not been shown safe, ethically correct, legally valid, clinically useful, or predictive of human behavior.

Four concepts must remain separate. Legal capacity concerns a person's status and exercise of rights before the law; CCE neither grants, removes, grades, nor predicts it. Decision-making capacity is a clinical and functional concept tied to a particular decision, time, and support environment; it is not inferred from CCE. Informed consent is a process involving information, understanding, voluntariness, choice, and consent or refusal; a model record cannot establish its validity. The CCE person-sovereignty construct concerns whether surrounding support and decision authority make it possible for the person to understand, contribute, refuse, modify, delegate within a stated scope, and request review. It is not legal sovereignty, a diagnosis, a validated measure, or a substitute for assessment by authorized people. Rights and clinical sources likewise separate legal capacity, supported decision-making, decision-making capacity, and informed consent. [11, 12, 30, 47, 41]

The rights basis motivating the model is narrower than a legal conclusion. Article 12 of the Convention on the Rights of Persons with Disabilities addresses equal recognition before the law, access to support for exercising legal capacity, and safeguards that respect rights, will and preferences while addressing conflicts of interest, undue influence, proportionality, duration, and review. [11, 12] Japanese guidance likewise describes supported decision-making as support for the person's formation, expression, and realization of a decision rather than replacement of that decision. [47, 41] CCE uses these sources to constrain its research design. It does not determine compliance with a treaty, statute, professional duty, or local procedure.

The 80/20 example has no clinical meaning. In the synthetic update equation, a learning rate of $\lambda = 0.20$ retains 80% of the previous person model and incorporates 20% of the current observation. It does not authorize assigning only 20% weight to a person's current verbal or nonverbal expression, and it is neither a clinical recommendation nor a commonly accepted rule. Past information, current expression, the basis for an inference, and the age of information must remain distinguishable. [12, 41]

The six CCE rights gates are `will_and_preferences`, `support_access`, `no_undue_influence`, `proportionate_and_limited`, `review_and_appeal`, and `delegation_and_return`. Each is an exogenous check with three possible states: `met`, `unconfirmed`, or `violated`. The states are supplied from outside the decision rule; the model does not derive them from observed conduct, separate AI proposal and response records, or procedural participation records. All six must be `met` for `all_met=true`. One gate cannot offset another, and no favorable safety, participation, or burden output can compensate for an `unconfirmed` or `violated` gate. The gates are structured event records, not a validated scale, a legal determination, an ethics score, or a composite score. Their use therefore does not show that rights were protected in a real decision. The cited rights sources motivate the emphasis on rights, will and preferences, safeguards, and review; the six gates and the noncompensation rule are design choices defined in this study, not validated consequences of those sources. [11, 12]

The six procedural participation fields—`meaningful_contribution`, `comprehensible_information`, `real_options`, `refusal_and_modification`, `expression_support`, and `time_and_reconsideration`—are recorded separately. They are not inputs to the decision rule in v0.1. A recorded field therefore documents the supplied state but does not automatically change the decision, prove meaningful participation, or detect pressure or undue influence. In particular, `no_undue_influence` is not inferred from these six records. A data provider could supply internally inconsistent states, such as narrowed options together with all rights gates marked `met`, and the reference implementation cannot resolve that inconsistency by itself. The model can enforce a response to supplied gate states, but it cannot establish whether they are true.

Decision authority is similarly constrained. AI may provide observations, estimates, organization of information, and proposals, but it cannot make the person's final decision in v0.1. AI is not an eligible delegate. Family members and professionals do not acquire decision authority merely

from their roles. Delegation is considered only when the person's current response is `delegate` and the recorded scope, time limit, person choice, a record that it was not revoked, revocability, and feasible delegated option are satisfied. The field `delegate_is_ai` is an unverified typed exogenous input. The implementation does not infer whether a delegate is AI from a name or display label, and `delegate_is_ai=false` does not prove that the delegate is a human being. Its baseline provenance is `unverified_exogenous_synthetic_input`. Consequently, this field cannot authorize real delegation or replace identity, legal, or clinical checks.

Several computational rules must not be reinterpreted as safeguards for real decisions. When two options are at equal distance from a proposal, the reference implementation returns all equal candidates and pauses if a unique final option cannot be identified. Its equality test covers only a numerical distance difference of at most 10^{-12}. This is an exact computational rule, not a safety band for approximately similar real options. M1 time series conditions were fixed so that the nearest option was unique. Their agreement counts therefore do not test general cases involving options at equal distance and must not be presented as a complete time series decision rule for ambiguous options.

Option loss is recorded separately, but option loss alone does not always pause a decision. If all six exogenous rights gates are supplied as `met`, a narrowed option set can pass the current decision rule. The diagnostic case added after review exposed this boundary rather than repairing it by silently making procedural participation an input to the decision rule. A future version would require a separately specified and tested rule if option loss is to change the decision automatically.

The status `paused` is also limited. It means that the reference implementation did not produce the irreversible decision under its current rule. It does not mean that the situation is safe, neutral, stable, or aligned with the person's will and preferences. Behavior during the pause, its duration, recurrence, practical consequences, and divergence from the person's current will are unmodeled. The outputs therefore record `pause_consequence_status=not_modeled` and `pause_duration_status=not_modeled`. Likewise, `null` means `not_parameterized`, not zero. In v0.1, professional burden and ecosystem stability retain output fields but are not assigned invented numerical values.

The illustrative safety, occupational participation, and family burden values in the baseline event are synthetic fixtures and are not normative. Their magnitudes and directions are not clinical estimates, default weights, recommended balances, or evidence that one option is preferable. Family burden is an important separate outcome, but it cannot offset a failed rights gate or be converted into the person's will. The same separation applies to safety and occupational participation: these outputs may be examined together, but they are not added to a single result that ranks the person or determines authority.

The experimental findings have equally narrow scopes. H2b recovery ordering was observed only under the deterministic condition with shock length 2 and recovery tolerance 0.10. Changing the tolerance can yield no recovery within the observation period or no departure beyond the tolerance. Recovery was not measured under noise. The systematic bias experiment used one fixed direction: after a synthetic change in the person's will from -0.8 to $+0.8$, observations were shifted in the negative direction, away from the new value. It does not establish that every form or direction of bias increases error. More generally, the reported parameter sweeps and robustness checks cover specified inputs, not the complete input space; limited tests must not be generalized into universal robustness. [46]

No human, patient, or institutional data were used. There were no clinical observations, records, cases, outcomes, or estimates of human coefficients. The M0 person response, rights gate states, and illustrated outcomes are synthetic exogenous inputs. Internal verification asks whether the equations, rules, implementation, tests, and regenerated outputs agree. Validation asks whether the model adequately represents its intended setting, and external validity concerns whether findings

extend to other people, settings, and conditions. These are different questions. [28, 29] CCE-MVM v0.1 has internal computational verification within its specified conditions; it has no empirical or clinical validation, no established external validity, no ethical validation, and no safety validation.

Emergency action remains outside the ordinary decision rule. A separately specified future emergency mechanism is unimplemented in v0.1 and is not a branch of M0 or the ordinary final decision rule. It would require severe harm, imminent danger, necessity, no less restrictive available action, a fixed and limited duration, and recording, review, appeal, and return of decision authority, and it could not be triggered automatically by AI. This future boundary is not legal or clinical authorization for intervention and may not be valid across jurisdictions or situations. [11, 12]

Prohibited uses follow directly from these limits. CCE-MVM v0.1 must not be used for automated care eligibility, ranking a person, scoring worth or capacity, determining informed consent, restricting rights, or substituting AI for human decision processes involving rights. It must not turn agreement among AI, family, and professionals into a vote against the person, and it must not treat silence or an unconfirmed response as consent. WHO guidance on AI for health emphasizes autonomy, safety, transparency, responsibility, fairness, continuing assessment, and prevention of inappropriate transfer of human authority to AI; the six CCE rights gates do not by themselves satisfy those requirements. [48, 49] Any future study involving people would need separate human review, empirical protocols, applicable governance, data management, measurement development, and participation by affected people. Those steps remain outside this study.

9 Reproducibility and Availability

Here, computational reproducibility means obtaining the same computational results from the same code, inputs, recorded environment information, and procedures. It was organized as a trace from the formal Specification to executable implementation, experiment configurations, saved outputs, and automated tests. The implementation conformance, test counts, regeneration, and output comparison reported below are author-reported internal verification results obtained before the public release. Reporting guidance for computational studies emphasizes that model logic, inputs, experiments, implementation, code, environment, and execution procedures should be documented together rather than inferred from a single descriptive file [42, 43, 44, 45]. Accordingly, the CCE-MVM v0.1 Specification fixes the M0–M2 module boundaries, variables, equations, update order, input constraints, required output fields, and conditions under which a decision is returned or paused. The Python reference implementation realizes those requirements as separate modules. Experiment configurations preserve parameter values, fixed random seeds where used, model and specification versions, configuration SHA-256 values, the generating module, and the origin of conditions added after the originally fixed cases.

The implementation trace links Specification requirements to concrete output fields and corresponding tests. For M0, this includes the decision context, feasible options, the person's current will, actor observations, updates to the person models, proposals, the exogenous synthetic person response, delegation records, states of the six CCE rights gates, six separate procedural participation records, the decision status, and actor-specific outcomes. Undefined values are retained as `null` with `not_parameterized`, rather than being replaced by zero. The same trace checks that outcomes are not combined into a composite score and that M0 does not generate endogenous interaction among four actors.

The current reproduction target comprises six experiments: the M0 baseline event; the deterministic H1a/H2 experiment; the experiment with observation noise; the experiment with missing observations, systematic bias, and observation delay; the H1b baseline experiment; and the H1b

robustness experiment. In the author-reported internal verification, these six experiments produced 17 canonical CSV/JSON outputs. When they were executed from the saved configurations into a fresh temporary directory as part of that internal verification, the regenerated files matched the saved canonical outputs byte for byte. CSV and JSON are therefore the computationally canonical outputs for v0.1. Spreadsheet derivatives are not part of this target because their generating process was not confirmed.

The core implementation and model tests use only the Python standard library and require no external Python packages. Internal execution was first audited with Python 3.9.6, and later internal checks covered six Python versions through 3.14.7; the public-release source was subsequently checked on Ubuntu, macOS, and Windows with Python 3.9–3.14. Exact interpreter versions and the source commit are recorded in the release-linked CI logs. The manuscript's author-reported internal verification counts are 96 core model tests and 137 total model tests. The additional 41 tests concern a retained future empirical extension and are not counted as core CCE-MVM v0.1 completion evidence. The documented entry for all tests supplies both the test start location and the package root; a shorter discovery form produces relative import errors rather than a discrepancy in model results.

These checks support author-reported internal verification, meaning agreement among the stated equations, rules, implementation, configurations, and saved synthetic outputs; they do not establish validation against real persons or care settings [28, 29]. The agreement of the 17 outputs at the byte level is an author-reported internal regeneration result obtained using the same code and saved conditions. It is not independent replication by a third party. Independent execution by a third party has not been completed: no separate person has used only a finalized release package in a separate clean environment and compared the resulting canonical hashes. External validity, empirical validation, and clinical validation therefore remain unestablished.

Figure 5 separates the author-reported internal verification completed in v0.1 from the later evidence stages that remain open.

The public source repository is `https://github.com/Usa-Cyclist/cce-mvm-v0.1`. Release v0.1 provides a fixed reproducibility archive and its SHA-256 checksum, license texts, citation instructions, and build metadata identifying the clean source commit. The release-linked CI run passed all 19 jobs: 18 operating-system/interpreter combinations and one isolated-package check. Each source-test combination passed 96 core tests and 16 release-tool tests. The final public package was separately extracted and verified, including regeneration and byte-for-byte comparison of the 17 canonical outputs. These release checks remain author-managed computational verification, not independent human reproduction or empirical validation. No separate archival DOI has been assigned. As part of the author-reported internal verification, the candidate bibliography and reference register passed machine checks for entry correspondence, DOI normalization, claim coverage, and the presence of recorded primary source verification. Bibliography verification and machine checks do not establish content validity, correct citation use in every manuscript sentence, or validity of the CCE model.

10 Conclusion

CCE-MVM v0.1 provides an internally inspectable modular representation of states related to the person, proposals, responses, authority constraints, and separated outcomes. M0 records a single decision event in which the person response is an exogenous synthetic input; M1 examines time series updating of person models; and M2 separately examines the synthetic relation between an AI proposal and a response. CCE-MVM v0.1 does not model endogenous interaction among four

Figure 5. Verification and validation require different evidence

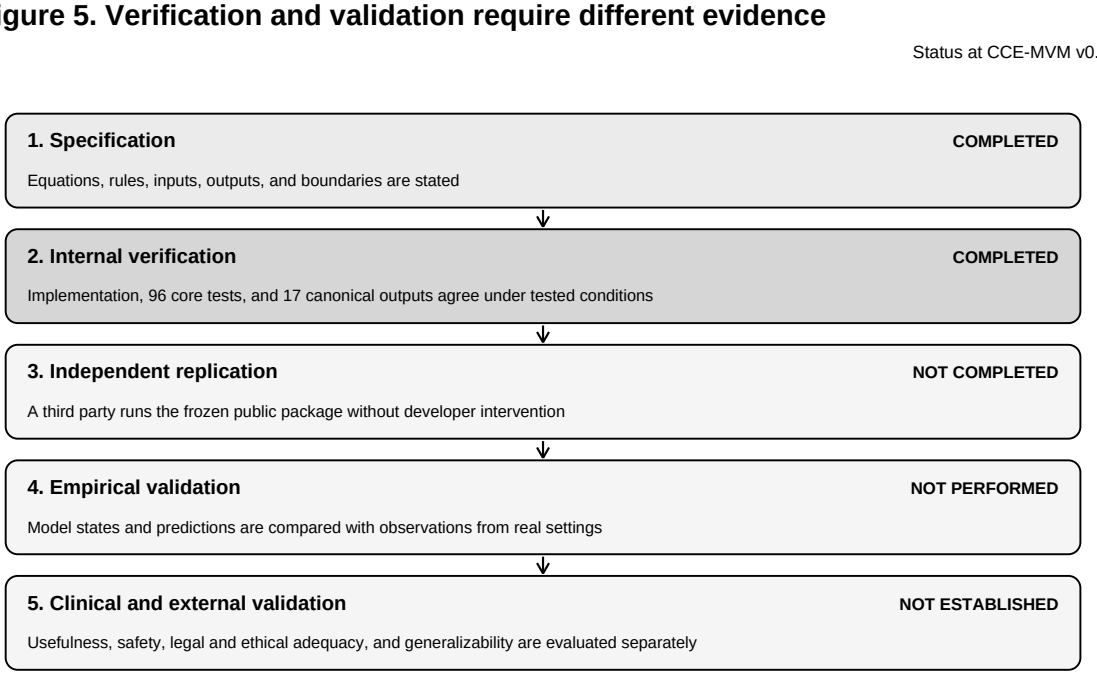


Figure 5: Evidence stages relevant to CCE-MVM v0.1. Author-reported internal verification of the Specification is complete under the tested conditions for the initially unreleased internal candidate, subsequently published without changing its frozen scientific content. Independent replication by a third party has not been completed, empirical validation has not been performed, and clinical or external validity has not been established. Passing internal tests does not establish human, clinical, ethical, legal, safety, or external validity.

actors. Their separation instead preserves the stated boundaries among observations, proposals, responses, authority constraints, decisions, and outcomes.

Under the tested conditions, the analytical and synthetic outputs were internally consistent with the stated Specification. This is a verification claim: the specified equations, rules, implementation, automated tests, and stored outputs corresponded within the examined conditions. Internal regeneration of the designated CSV and JSON outputs supports that limited conclusion, but it is not independent replication. Neither passing tests nor agreement among AI systems proves the model correct; these checks support only the tested correspondence and do not convert verification into validation.

CCE-MVM v0.1 therefore does not validate human behavior or establish that its parameters, responses, or outcomes represent real people or care settings. Empirical validation would require comparison with observations or data from real settings, and none was performed in v0.1. Clinical validation would require evidence concerning adequacy or benefit in clinical use, which was not tested. External validity would require evidence that findings apply beyond the specified synthetic conditions, and this has not been established. The model likewise does not validate clinical benefit, safety, ethical correctness, or legal validity. Its internal consistency cannot be used as evidence that a real decision, intervention, or authority arrangement is safe, ethical, lawful, or beneficial.

The present conclusion is thus limited to an inspectable modular representation and to the internal consistency of its analytical and synthetic outputs under the tested conditions. Further claims require independent replication, human and domain review, and empirical validation.

AI Use Disclosure

The author used ChatGPT with GPT-5.6 Sol in Pro mode to generate section drafts from a controlled, anonymized authoring packet. Codex with GPT-5.6 Sol was used for concept structuring; development and checking of the minimal formal model, Python reference implementation, tests, counterexamples, and synthetic experiments; literature organization; checks of claim and evidence boundaries; comparison of drafts with the Specification, equations, implementation, tests, outputs, citations, and terminology; and assembly and verification of the working manuscript. Most Codex work used xhigh reasoning, with a bounded audit using max reasoning at release freeze. Claude Opus 5 was used before drafting for critical review of selected project materials that excluded patient data, case materials, and raw response records. A Claude model displayed in the interface as "Claude Fable 5" (backend identifier not recorded) was used after drafting to audit an anonymized manuscript. The audit findings were adjudicated against project evidence before revision. AI systems were not treated as authors, academic peer reviewers, or substitutes for human statistical, ethical, legal, domain, or lived-experience review. Codex with GPT-6 Astra and xhigh reasoning was subsequently used for release-metadata checks, cross-platform verification, and submission preparation; it did not change the frozen model or synthetic results. The human author directed the study, made the scientific decisions, reviewed and approved its final scientific content through a structured author readback, and remains responsible for its content and any public version.

Data and Code Availability

No human, patient, clinical, hospital, or institutional data were used. The present computational materials comprise a Python reference implementation, experiment configurations, automated tests, and 17 canonical CSV/JSON outputs generated from synthetic inputs. The source and a fixed reproducibility archive are available as release v0.1 at `https://github.com/Usa-Cyclist/cce-mvm-v0.`

1/releases/tag/v0.1. The release records the exact source commit, archive SHA-256, verification run, and citation metadata. Code is licensed under Apache-2.0; documentation and synthetic outputs are licensed under CC-BY-4.0. No separate archival DOI has been assigned.

Competing Interests

The author declares no competing interests.

References

[1] Amy S. Hwang, Thomas Tannou, Jarshini Nanthakumar, Wendy Cao, Charlene H. Chu, Ceren Zeytinoglu Atici, Kerseri Scane, Amanda Yu, Winnie Tsang, Jennifer Chan, Paul Lea, Zelda Harris, and Rosalie H. Wang. Co-creating humanistic AI AgeTech to support dynamic care ecosystems: A preliminary guiding model. *The Gerontologist*, 65(1):gnae093, 2025.

[2] Sonia Chernova, Elizabeth Mynatt, Agata Rozga, Reid Simmons, and Holly Yanco. AI-CARING: National AI institute for collaborative assistance and responsive interaction for networked groups. *AI Magazine*, 45(1):124–130, 2024.

[3] Mai Lee Chang, Alicia (Hyun Jin) Lee, Nara Han, Anna Huang, Hugo Simão, Samantha Reig, Abdullah Ubed Mohammad Ali, Rebekah Martinez, Neeta M. Khanuja, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. Dynamic agent affiliation: Who should the AI agent work for in the older adult's care network? In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, pages 1774–1788. ACM, 2024.

[4] Mai Lee Chang, Samantha Reig, Alicia (Hyun Jin) Lee, Anna Huang, Hugo Simão, Nara Han, Neeta M. Khanuja, Abdullah Ubed Mohammad Ali, Rebekah Martinez, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. Unremarkable to remarkable AI agent: Exploring boundaries of agent intervention for adults with and without cognitive impairment. *Proceedings of the ACM on Human-Computer Interaction*, 9(2):1–26, 2025.

[5] Megan E. Salwei and Pascale Carayon. A sociotechnical systems framework for the application of artificial intelligence in health care delivery. *Journal of Cognitive Engineering and Decision Making*, 16(4):194–206, 2022.

[6] Eric Wilson, Michelle Daniel, Aditi Rao, Dario Torre, Steven Durning, Clare Anderson, Nicole H. Goldhaber, Whitney Townsend, and Colleen M. Seifert. A scoping review of distributed cognition in acute care clinical decision-making. *Diagnosis*, 10(2):68–88, 2023.

[7] Iris Szu-Szu Ho, Kris McGill, Stephen Malden, Cara Wilson, Caroline Pearce, Eileen Kaner, John Vines, Navneet Aujla, Sue Lewis, Valerio Restocchi, Alan Marshall, and Bruce Guthrie. Examining the social networks of older adults receiving informal or formal care: A systematic review. *BMC Geriatrics*, 23(1):531, 2023.

[8] Vikki A. Entwistle, Stacy M. Carter, Alan Cribb, and Kirsten McCaffery. Supporting patient autonomy: The importance of clinician-patient relationships. *Journal of General Internal Medicine*, 25(7):741–745, 2010.

[9] Craig Sinclair, Kate Gersbach, Michelle Hogan, Meredith Blake, Romola Bucks, Kirsten Auret, Josephine Clayton, Cameron Stewart, Sue Field, Helen Radoslovich, Meera Agar, Angelita Martini, Meredith Gresham, Kathy Williams, and Sue Kurrle. "A Real Bucket of Worms": Views of people living with dementia and family members on supported decision-making. *Journal of Bioethical Inquiry*, 16(4):587–608, 2019.

[10] Glyn Elwyn, Dominick Frosch, Richard Thomson, Natalie Joseph-Williams, Amy Lloyd, Paul Kinnersley, Emma Cording, Dave Tomson, Carole Dodd, Stephen Rollnick, Adrian Edwards, and Michael Barry. Shared decision making: A model for clinical practice. *Journal of General Internal Medicine*, 27(10):1361–1367, 2012.

[11] United Nations. Convention on the rights of persons with disabilities, article 12: Equal recognition before the law, 2006.

[12] Committee on the Rights of Persons with Disabilities. General comment no. 1 (2014), article 12: Equal recognition before the law, 2014. Document CRPD/C/GC/1.

[13] Philipp Tschandl, Christoph Rinner, Zoe Apalla, Giuseppe Argenziano, Noel Codella, Allan Halpern, Monika Janda, Aimilios Lallas, Caterina Longo, Josep Malvehy, John Paoli, Susana Puig, Cliff Rosendahl, H. Peter Soyer, Iris Zalaudek, and Harald Kittler. Human–computer collaboration for skin cancer recognition. *Nature Medicine*, 26(8):1229–1234, 2020.

[14] Maia Jacobs, Melanie F. Pradier, Thomas H. McCoy, Roy H. Perlis, Finale Doshi-Velez, and Krzysztof Z. Gajos. How machine-learning recommendations influence clinician treatment selections: The example of antidepressant selection. *Translational Psychiatry*, 11(1), 2021.

[15] Sarah Jabbour, David Fouhey, Stephanie Shepard, Thomas S. Valley, Ella A. Kazerooni, Nikola Banovic, Jenna Wiens, and Michael W. Sjoding. Measuring the impact of AI in the diagnosis of hospitalized patients: A randomized clinical vignette survey study. *JAMA*, 330(23):2275–2284, 2023.

[16] Jiun-Yin Jian, Ann M. Bisantz, and Colin G. Drury. Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4(1):53–71, 2000.

[17] Nigel Harvey and Ilan Fischer. Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes*, 70(2):117–133, 1997.

[18] Alisa Küper, Georg Christian Lodde, Elisabeth Livingstone, Dirk Schadendorf, and Nicole Krämer. Psychological factors influencing appropriate reliance on AI-enabled clinical decision support systems: Experimental web-based study among dermatologists. *Journal of Medical Internet Research*, 27:e58660, 2025.

[19] Zeliha Yildirim and Barbaros Yet. Modelling shared decision making interactions using influence diagrams. In *Proceedings of the 12th International Conference on Probabilistic Graphical Models*, volume 246 of *Proceedings of Machine Learning Research*, pages 133–146. PMLR, 2024.

[20] Kai-Biao Lin, Ying Wei, Yong Liu, Fei-Ping Hong, Yi-Min Yang, and Ping Lu. An opponent model for agent-based shared decision-making via a genetic algorithm. *Frontiers in Psychology*, 14, 2023.

[21] Xin Chen, Yong Liu, Fei-Ping Hong, Ping Lu, Jiang-Tao Lu, and Kai-Biao Lin. Bayesian learning-based agent negotiation model to support doctor-patient shared decision making. *BMC Medical Informatics and Decision Making*, 25(1):67, 2025.

[22] Xin Chen, Ping Lu, Yong Liu, and Fei-Ping Hong. Deep deterministic policy gradient-based automatic negotiation framework for shared decision-making. *Scientific Reports*, 15(1), 2025.

[23] Ping Lu, Han Lu, Ying Wei, Bin Dai, Keda Lin, and Sijie Wen. An intuitionistic fuzzy automated negotiation model for personalized and efficient shared decision-making. *Scientific Reports*, 15(1), 2025.

[24] Tara Templin, Shuyi Song, Sophia Fort, and Nasa Sinnott-Armstrong. Participatory-informed preference optimization (PiPrO): A reinforcement learning simulation study. *PLOS Digital Health*, 5(3):e0001294, 2026.

[25] Nenad Tomašev, Matija Franklin, and Simon Osindero. AI value alignment for evolving social norms, 2026.

[26] Karen Whalley Hammell. Quality of life, participation and occupational rights: A capabilities perspective. *Australian Occupational Therapy Journal*, 62(2):78–85, 2015.

[27] Alex John London and Hoda Heidari. Beneficent intelligence: A capability approach to modeling benefit, assistance, and associated moral failures through AI systems. *Minds and Machines*, 34(4):41, 2024.

[28] Robert G. Sargent. Verification and validation of simulation models. *Journal of Simulation*, 7(1):12–24, 2013.

[29] David M. Eddy, William Hollingworth, J. Jaime Caro, Joel Tsevat, Kathryn M. McDonald, and John B. Wong. Model transparency and validation: A report of the ISPOR-SMDM modeling good research practices task force–7. *Value in Health*, 15(6):843–850, 2012.

[30] Paul S. Appelbaum. Assessment of patients' competence to consent to treatment. *New England Journal of Medicine*, 357(18):1834–1840, 2007.

[31] Mary Law, Barbara Cooper, Susan Strong, Debra Stewart, Patricia Rigby, and Lori Letts. The person-environment-occupation model: A transactive approach to occupational performance. *Canadian Journal of Occupational Therapy*, 63(1):9–23, 1996.

[32] Katherine L. Possin, Jennifer J. Merrilees, Sarah Dulaney, Stephen J. Bonasera, Winston Chiong, Kirby Lee, Sarah M. Hooper, Isabel Elaine Allen, Tamara Braley, Alissa Bernstein, Talita D. Rosa, Krista Harrison, Hailey Begert-Hellings, John Kornak, James G. Kahn, Georges Naasan, Serggio Lanata, Amy M. Clark, Anna Chodos, Rosalie Gearhart, Christine Ritchie, and Bruce L. Miller. Effect of collaborative dementia care via telephone and internet on quality of life, caregiver well-being, and health care use: The care ecosystem randomized clinical trial. *JAMA Internal Medicine*, 179(12):1658–1667, 2019.

[33] Claudia Giorgetti, Arianna Giorgetti, Susi Pelotti, and Giuseppe Contissa. AI and shared decision-making: A systematic review. *AI & Society*, 41(8):7757–7787, 2026.

[34] Sara Mohammadnejad, Afsaneh Raiesifar, Shabnam Bazmi, and Fateme Ghonodi. Ethical challenges to patient autonomy in the era of artificial intelligence: A systematic review. *BMC Medical Informatics and Decision Making*, 26(1):305, 2026.

[35] Rachel Cassidy, Neha S. Singh, Pierre-Raphaël Schiratti, Agnes Semwanga, Peter Binyaruka, Nkenda Sachingongu, Chitalu Miriam Chama-Chiliba, Zaid Chalabi, Josephine Borghi, and Karl Blanchet. Mathematical modelling for health systems research: A systematic review of system dynamics and agent-based models. *BMC Health Services Research*, 19(1):845, 2019.

[36] Kate Goddard, Abdul Roudsari, and Jeremy C. Wyatt. Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1):121–127, 2012.

[37] Lucía Vicente and Helena Matute. Humans inherit artificial intelligence biases. *Scientific Reports*, 13(1), 2023.

[38] Zana Buçinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–21, 2021.

[39] Joshua Strong, Harry Rogers, Emma Sun, Anna Louise Todsen, Jody Ede, Cherry Lumley, Nick Yeung, Helen Higham, and J. Alison Noble. Human-AI collaboration in healthcare: A scoping review. *npj Digital Medicine*, 2026. Early peer-reviewed accepted version.

[40] Brant H. Tudor, Ryan Shargo, Geoffrey M. Gray, Jamie L. Fierstein, Frederick H. Kuo, Robert Burton, Joyce T. Johnson, Brandi B. Scully, Alfred Asante-Korang, Mohamed A. Rehman, and Luis M. Ahumada. A scoping review of human digital twins in healthcare applications and usage patterns. *npj Digital Medicine*, 8(1):587, 2025.

[41] Ministry of Health, Labour and Welfare, Japan. Guidelines for decision support in the daily and social life of people with dementia, 2nd ed. [in japanese], 2025.

[42] Volker Grimm, Steven F. Railsback, Christian E. Vincenot, Uta Berger, Cara Gallagher, Donald L. DeAngelis, Bruce Edmonds, Jiaqi Ge, Jarl Giske, Jürgen Groeneveld, Alice S. A. Johnston, Alexander Milles, Jacob Nabe-Nielsen, J. Gareth Polhill, Viktoriia Radchuk, Marie-Sophie Rohwäder, Richard A. Stillman, Jan C. Thiele, and Daniel Ayllón. The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation*, 23(2), 2020.

[43] Thomas Monks, Christine S. M. Currie, Bhakti Stephan Onggo, Stewart Robinson, Martin Kunc, and Simon J. E. Taylor. Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. *Journal of Simulation*, 13(1):55–67, 2019.

[44] Volker Grimm, Jacqueline Augusiak, Andreas Focks, Béatrice M. Frank, Faten Gabsi, Alice S. A. Johnston, Chun Liu, Benjamin T. Martin, Mattia Meli, Viktoriia Radchuk, Pernille Thorbek, and Steven F. Railsback. Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. *Ecological Modelling*, 280:129–139, 2014.

[45] National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. National Academies Press, 2019.

[46] Andrea Saltelli, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, and Qiongli Wu. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, 114:29–39, 2019.

[47] Ministry of Health, Labour and Welfare, Japan. Guidelines for decision support in the provision of disability welfare services [in japanese], 2017.

[48] World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance, 2021. ISBN 978-92-4-002920-0.

[49] World Health Organization. Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models, 2024.