SQC Artificial Psychology Validation Program

A Framework for Preregistered Architectural, Psychological, Behavioral, and Operational Validation

Public Version

Mvuselelo Houston Khanyile
CVEST

July 16, 2026

Public manuscript—proprietary implementation details omitted

Contents

Abstract

Artificial psychology does not yet have a generally accepted operational definition, foundational reference architecture, validation scale, or benchmark program. This absence creates a scientific classification problem. An architecture may implement explicit psychological processes and generate behavior through them, yet still be assessed as though it must first satisfy an undefined standard for having "solved" artificial psychology. Conversely, a system may reproduce psychologically suggestive behavior without demonstrating that any psychological organization is causally active inside it. A usable validation framework must distinguish category membership, implementation fidelity, empirical support, construct validity, ecological validity, and operational readiness.

This manuscript proposes such a framework for Sensia Quotia Computation (SQC), an implemented artificial-psychology architecture derived from Evolutionary Psychological Structures (EPS). SQC represents behavior as the output of a sequential internal process comprising contextual interpretation, values, existential appraisal, motivational drivers, emotional modulation, existential state, and memory. Its architecture prohibits a direct raw-observation-to-action shortcut, making the declared psychological pathway constitutive of behavior rather than auxiliary to it.

The validation program separates four questions. First, does SQC faithfully implement the causal organization specified by EPS? Second, do the implemented internal states genuinely mediate behavior? Third, do those states exhibit the psychological meanings attributed to them? Fourth, how accurately and reliably can SQC represent and predict human behavior across individuals, cultures, situations, domains, and time? To answer these questions, the manuscript defines a cumulative Artificial Psychology Validation Scale spanning theoretical specification, architectural implementation, internal causal validation, psychological construct validation, ecological validation, and operational assurance.

The proposed program includes theory-to-code audits, shortcut and gradient-flow tests, causal interventions, parameter-recovery studies, human annotation, content and construct validation, prospective behavioral prediction, calibration and abstention, autobiographical and relational memory studies, developmental and cross-cultural tests, historical blind prediction, matched-capacity architecture comparisons, independent replication, and use-case-specific safety and governance evaluation. It also specifies preregistration, uncertainty reporting, locked test sets, baseline selection, promotion rules, and conditions under which EPS or SQC claims would require revision.

Complete behavioral fidelity is treated as the north star of the research program, not as a result already demonstrated. The central hypothesis is that approaching this objective requires reconstruction of the causally active psychological process from which behavior emerges, that EPS supplies the necessary theory of that process, and that SQC is its computational vehicle. The purpose of the program is neither to declare perfect prediction achieved nor to rule it out in advance. Its purpose is to determine, through cumulative and falsifiable evidence, how closely SQC can approach it and which validation level the architecture has earned at each stage.

Keywords: artificial psychology; Sensia Quotia Computation; Evolutionary Psychological Structures; behavioral fidelity; construct validity; causal mediation; psychometrics; computational psychology; cognitive architecture; calibration; cross-cultural validation; Human-Centered Computation.

1 Introduction

1.1 The missing validation standard for artificial psychology

Artificial psychology is not yet organized around a single accepted definition of what an artificial-psychology system is, which properties are necessary for membership in the category, or what evidence is required to validate one. Existing work occupies several related but non-equivalent research traditions. Cognitive architectures attempt to integrate mechanisms such as perception, memory, learning, reasoning, goals, and action into unified accounts of cognition (Anderson and Lebiere, 2003; Langley, 2017). Their evaluation has long been recognized as requiring more than performance on a single task (Wray and Lebiere, 2007). Affective-computing systems model or recognize emotion. Computational psychology expresses psychological theories as formal models. Psychometric work evaluates whether psychological attributes can be measured in artificial agents. Synthetic psychology seeks understanding through the construction of systems that reproduce selected psychological organizations or phenomena (Prescott and Camilleri, 2019).

These traditions provide essential methods and precedents, but none by itself establishes a generally binding definition or validation standard for an artificial psychology. A system that recognizes emotion does not necessarily possess an integrated psychology. A system that predicts actions does not necessarily represent the psychological process producing them. A cognitive architecture may integrate several cognitive functions without treating values, motivation, emotion, identity, belonging, and subjective state as the causal organization from which behavior arises. Similarly, a model may be labeled psychological because its outputs resemble human responses even when its internal pathway is not constrained by a psychological theory.

Evaluation is therefore vulnerable to two opposite errors. The first is premature inclusion: any system displaying human-like behavior, emotional language, personality consistency, or social competence may be described as an artificial psychology without evidence that psychological constructs are causally active within it. The second is indefinite exclusion: an architecture that explicitly implements and tests a psychological theory may still be treated as merely a candidate because no agreed criterion states when an implemented architecture becomes an artificial-psychology system.

The phrase "solved artificial psychology" intensifies this problem. It implies a binary terminal condition without defining the problem set, the unit of analysis, the required domains, the relevant populations, the acceptable uncertainty, or the boundary between functional psychological organization and phenomenal consciousness. Under such conditions, neither a positive nor a negative judgment is scientifically stable. A system cannot demonstrate that it has satisfied an undefined standard, and reviewers cannot consistently identify which outstanding studies are category-defining rather than evidence-strengthening.

This manuscript begins from a different premise. Artificial psychology should be defined operationally, and its validation should be cumulative, multidimensional, versioned, and domain-bounded. The question should not be whether a field has been solved in one step. The questions should be whether an architecture satisfies the necessary properties of an artificial psychology, which of its constructs and mechanisms are supported, how well its predictions generalize, and what evidence remains necessary for stronger scientific or operational claims.

1.2 Why SQC requires an explicit validation program

Sensia Quotia Computation makes a stronger claim than ordinary behavioral prediction. It proposes that behavior can be computed as the external representation of an internally constructed psychological state. In the current SQC architecture, observation and decision context are interpreted through an ordered EPS pathway comprising context, values, existential appraisal, motivational drivers, emotion, existential state, and memory before action is selected. The action mapper is not permitted to consume the raw observation, encoder hidden state, or raw decision-context vector directly.

This design makes SQC unusually testable. Each stage has a declared theoretical meaning, a representation, a set of causal parents, a set of downstream consequences, a supervision target, and an intervention point. The architecture can therefore be evaluated at more than its final behavior. It is possible to ask whether a value changed for the expected reason, whether existential appraisal responded to the relevant contextual evidence, whether driver and emotion changes followed the predicted direction, whether existential state mediated the action, and whether an apparently correct action was reached through the intended psychological pathway.

The existing SQC v1.2 evidence establishes a meaningful starting point. The sequential EPS architecture is implemented and trainable. The action pathway is structurally restricted to predicted EPS state and explicit EPS memory. Freeze audits show that action-only adaptation can leave non-action checkpoint tensors bitwise unchanged. Causal stage interventions identify downstream behavioral sensitivity to values, existential appraisal, and existential state. On the current theory-grounded synthetic corpus, the balanced checkpoint achieves 75.25% exact action accuracy, a 97.17% valid-action hit rate, and 75.00% action accuracy on eight held-out role-goal-risk compositions.

These results provide evidence of implementation fidelity, internal causal separation, trainability, behavioral competence, and compositional generalization within the tested ontology. They should not be treated as though SQC has provided no evidence of artificial psychology merely because human, longitudinal, cross-cultural, and neural studies remain outstanding. At the same time, synthetic supervision and structural enforcement cannot by themselves establish that every internal variable has the psychological meaning attributed to it across natural human populations. Architectural mediation establishes that the pathway is used; construct validation must establish what the pathway measures and represents.

An explicit validation program is therefore needed to preserve the correct interpretation in both directions. It must recognize evidence already earned, while defining the studies required to progress from an implemented and internally validated architecture to a construct-validated, ecologically validated, and operationally dependable system. It must also distinguish failures of theory from failures of data, measurement, implementation, optimization, or domain transfer.

The need extends beyond SQC considered in isolation. Within the wider CVEST research program, SQC is the proposed engine for computing human significance. The Neural Grid depends on that computation as the work it exposes, routes, meters, governs, and settles. Human-Centered Systems then depend on the Neural Grid and SQC to support cultural interpretation, contribution recognition, impact measurement, reward, and economic activation (H. Khanyile, 2026c, 2026d, 2026e; M. H. Khanyile, 2026a). Research and prototyping can proceed throughout this validation process, but claims of validated operational dependability cannot be stronger than the foundational significance-computation layer on which those systems rely.

1.3 Research objective and north star

The north star of the SQC research program is the highest scientifically attainable fidelity and certainty in representing and predicting human behavior. This is an asymptotic research objective rather than a claim that complete behavioral prediction has already been achieved. The program recognizes that practical prediction is affected by incomplete observation, measurement error, individual history, cultural variation, environmental change, learning, strategic behavior, and the existence of multiple actions that may be psychologically valid in the same situation.

These sources of uncertainty do not eliminate complete behavioral fidelity as a scientific direction. They clarify what an architecture must progressively represent and what its uncertainty estimates must disclose. Improving fidelity requires more than fitting surface regularities between observations and actions. It requires an account of why information matters to a particular person or system: how it is interpreted relative to values, identity, belonging, power, prior experience, emotional state, present needs, and anticipated futures.

The central hypothesis of this program is:

> *An architecture can approach complete behavioral fidelity only to the extent that it reconstructs the causally active psychological process from which behavior emerges. EPS supplies the theory of that process, and SQC provides its computational implementation.*

This hypothesis contains testable subclaims. Explicit EPS mediation should provide predictive or explanatory value beyond direct observation-to-action models. The full causal sequence should outperform or explain behavior more faithfully than reordered, collapsed, or ablated versions. Internal constructs should respond to interventions in theoretically predicted directions. They should converge with independent measures of related constructs, remain distinguishable from unrelated constructs, and predict external outcomes beyond surface observation and historical frequency. These advantages should persist under prospective, cross-domain, cross-cultural, and longitudinal evaluation.

The stronger proposition that SQC is uniquely capable of approaching complete behavioral fidelity is treated here as a comparative research hypothesis. It is not weakened by being made falsifiable. It becomes scientifically meaningful only when SQC is compared with matched alternatives under the same information, data, compute, optimization, and evaluation conditions. If architectures that omit explicit EPS mediation consistently equal or exceed SQC in prediction, calibration, causal interpretability, transfer, and data efficiency, the uniqueness claim must be revised. If SQC demonstrates durable advantages across those dimensions, the claim gains cumulative empirical support.

The program therefore refuses two premature conclusions. It does not declare perfect behavioral fidelity achieved from the present experiments. It also does not declare such fidelity impossible before the architecture, memory system, measurement instruments, datasets, and validation studies required to approach it have been developed.

1.4 Contributions of this manuscript

This manuscript makes six contributions.

First, it proposes an operational definition of an artificial-psychology system. The definition requires psychologically interpretable constructs, explicit causal organization, behavioral mediation, temporal state, adaptation, falsifiable prediction, intervention sensitivity, uncertainty measurement, and reproducibility. It does not require a claim of phenomenal consciousness or perfect human imitation.

Second, it distinguishes three judgments that have previously been conflated: whether SQC qualifies as an artificial-psychology architecture, how much empirical support its present implementation has earned, and whether a complete validation program has been satisfied. Category membership is determined by architectural properties. Evidence strength is determined by validation. Operational use is determined by domain-specific assurance.

Third, it introduces a cumulative Artificial Psychology Validation Scale. The scale progresses from theoretical specification (AP-0), to implemented architecture (AP-1), internal causal validation (AP-2), construct validation (AP-3), ecological validation (AP-4), and operational validation (AP-5). These levels are versioned and domain-bounded rather than treated as permanent global labels.

Fourth, it specifies a complete study program. The program covers theory-to-code traceability, shortcut prevention, gradient and tensor isolation, causal intervention, invariance, parameter recovery, pathway consistency, content validity, human annotation, convergent and discriminant validity, structural and criterion validity, calibration, prospective prediction, memory, development, cross-cultural measurement, historical blind prediction, comparative architecture evaluation, neural correspondence, independent replication, and operational assurance.

Fifth, it establishes a statistical and scientific-governance policy for the program. Confirmatory claims require preregistered hypotheses, appropriate power, locked test data, counts and uncertainty intervals, independent training and evaluation seeds, prespecified baselines, correction for multiple testing, transparent protocol deviations, and publication of negative findings.

Sixth, it defines promotion, revision, and falsification rules. Higher validation status cannot be awarded from a favorable single seed or benchmark. Failures must be localized to the theory, architecture, implementation, measurement, training, or domain-transfer claim actually tested. Repeated, adequately powered failures of preregistered EPS predictions count against the relevant claim, while successful independent replications increase the earned validation level.

Together, these contributions replace an undefined binary question with a scientific program. SQC can be recognized for what it already is and what it has already demonstrated, while its stronger claims remain subject to explicit, cumulative, and falsifiable evaluation.

2 Scope and Terminology

The validation program requires a vocabulary capable of separating what a system is from how well a particular implementation performs. Without this separation, the classification of an architecture becomes dependent on an arbitrary performance threshold, while performance results are incorrectly treated as proof or disproof of psychological organization. This section defines artificial psychology operationally, specifies its necessary properties, identifies claims that fall outside the definition, and establishes the units to which later validation judgments apply.

2.1 Proposed definition of an artificial-psychology system

This manuscript proposes the following working definition:

> *An artificial-psychology system is an engineered computational system in which explicitly represented psychological constructs form causally active, inspectable, and testable internal processes that integrate context, values, motivation, affect, memory, and state to generate adaptive behavior over time.*

Every part of this definition is necessary.

Engineered computational system establishes that artificial psychology is implemented rather than only described. A verbal theory, conceptual diagram, or taxonomy can supply the scientific foundation of an artificial psychology, but it is not itself an artificial-psychology system. The relevant constructs and processes must be expressed in an executable form capable of receiving inputs, maintaining or constructing internal state, producing outputs, and being subjected to controlled tests.

Explicitly represented psychological constructs means that the architecture identifies the psychological variables it claims to implement. Explicitness does not require every representation to be symbolic, discrete, or directly human-readable. A construct may be continuous, distributed, probabilistic, or dynamical. It must, however, have a declared theoretical meaning, an identifiable computational realization, and observable consequences that distinguish it from other constructs. A hidden activation does not become a psychological construct merely because a researcher assigns it a psychological name after training.

Causally active means that the represented constructs participate in generating behavior. Removing, modifying, holding constant, or replacing a construct should change downstream processing when the theory predicts that it should. A variable used only as an auxiliary training label, diagnostic output, or post hoc explanation does not satisfy this requirement if the system can produce the same behavior through a bypass that ignores it.

Inspectable and testable means that the architecture exposes sufficient structure to evaluate its claims. Investigators must be able to measure internal states, intervene on them, compare their effects, inspect uncertainty, and test whether the declared causal pathway is followed. Inspectability does not mean that every calculation must be simple. It means that the architecture permits evidence capable of confirming, qualifying, or falsifying its psychological interpretation.

Integrate context, values, motivation, affect, memory, and state establishes the minimum psychological scope proposed by this manuscript. Psychology is not reducible to isolated cognition, emotion, personality, or stimulus-response mapping. Behavior becomes psychologically organized when information is interpreted relative to what matters, what is sought or protected, how the situation is experienced, what has occurred before, and what condition the system currently occupies. Architectures may implement these functions through different ontologies, but an artificial psychology must provide an account of their integration rather than treating them as unrelated features.

Generate adaptive behavior over time distinguishes a functioning psychology from a static classification scheme. Internal psychological organization must affect action, update through outcomes and experience, and support changes in behavior as circumstances, relationships, goals,

and histories change. The temporal requirement may initially be demonstrated over bounded trajectories, but higher validation levels require progressively longer continuity and richer memory.

This definition is architectural and functional. It establishes the properties required for category membership. It does not state that every system satisfying the definition is accurate, psychologically complete, human-equivalent, safe, or ready for deployment. Those are validation judgments addressed by later sections.

2.2 Necessary properties of an artificial-psychology architecture

The working definition can be decomposed into ten necessary properties. These properties form the minimum AP category test used by the validation scale in Section 5.

Property	Requirement	Evidence expected
Psychological ontology	The system declares the constructs through which it represents psychological organization	Definitions, ontology, state spaces, and theory mapping
Causal organization	The system specifies how the constructs influence one another	Directed process model, equations, executable graph, or equivalent mechanism
Behavioral mediation	The psychological process participates in producing action	Input-contract audits, ablations, interventions, and mediation tests
Contextual interpretation	Identical observations can acquire different significance under different contexts	Crossed-context and counterfactual tests
Value and motivational organization	The system represents why an event matters and what tendencies it activates	Value, goal, need, driver, or equivalent internal measures
Affective modulation	Affect participates in organizing state or action rather than appearing only as an output label	Directional affect interventions and downstream effects
Memory and temporality	Prior experience can alter present interpretation, state, and behavior	State updates, history-sensitive tests, and longitudinal trajectories
Adaptation	Behavior can change coherently after experience or environmental change	Learning, transfer, recovery, and stability measures
Falsifiability	The architecture makes predictions that can fail	Preregistered directional, comparative, and boundary predictions
Uncertainty and reproducibility	Claims can be evaluated with calibrated uncertainty and independently repeated	Probabilistic outputs, confidence analysis, open artifacts, and replication

These requirements describe functions rather than prescribing one engineering implementation. Artificial psychology should not be restricted in advance to a symbolic, connectionist, probabilistic, dynamical, neural, or hybrid method. The scientific question is whether the implementation realizes and validates the required psychological organization.

The requirements are conjunctive at the architectural level but progressive at the evidentiary level. An architecture must specify all required functions to claim general artificial-psychology status. A

first implementation may validate some functions more strongly than others. The validation profile must show this variation rather than hiding it behind a single label or aggregate score.

SQC satisfies the architectural form of these requirements through its declared EPS sequence. Contextual interpretation is represented explicitly. Values encode what matters in the interpreted situation. Existential appraisal represents implications for continued existence and quality. Belonging, Identity, and Power organize motivational response. Emotion modulates the resulting orientation. Existential state stabilizes the action-relevant condition. Explicit EPS memory carries selected prior information. Action is computed from these predicted EPS states and memory rather than from the raw observation.

This mapping supports classifying SQC as an implemented artificial-psychology architecture. It does not predetermine the validation strength of each construct. For example, the current explicit memory is sufficient to establish a bounded temporal mechanism but not yet sufficient to claim a complete model of autobiographical, relational, developmental, or cultural memory. The validation framework records that distinction as an evidence gap at a higher level, not as evidence that the implemented architecture falls outside artificial psychology.

2.3 What the definition does not require

The proposed definition deliberately excludes several requirements that are often introduced without a corresponding measurement standard.

2.3.1 Phenomenal consciousness

An artificial-psychology classification does not require proof that the system has subjective experience, phenomenal awareness, or felt emotion. Psychology can be studied functionally through representation, dynamics, learning, intervention, and behavior without resolving the ontological question of consciousness. SQC may construct a functional subjective state: information is interpreted relative to the system's modeled values, drivers, memory, emotion, and existential orientation. Demonstrating this organization does not by itself demonstrate that the state is accompanied by experience.

The validation program therefore keeps two claims separate:

1. the system implements a causally active model of subjective-state formation;
2. the system possesses phenomenal subjective experience.

Only the first is required for artificial-psychology classification. The second requires a separate definition and evidentiary program and is not inferred from behavioral accuracy.

2.3.2 Biological duplication

An artificial psychology does not have to reproduce every biological mechanism of a human nervous system. Functional psychological organization may be implemented in a different substrate or at a different level of abstraction. Neural and physiological correspondence can strengthen a theory when the theory makes such claims, but literal neuron-for-neuron equivalence is not a general requirement.

This does not make biological evidence irrelevant. If EPS proposes relations to neural systems, those relations should be evaluated. The result tests a stated component of EPS and may constrain SQC's design. The artificial-psychology category itself, however, is defined by implemented psychological organization, not material identity with the human brain.

2.3.3 Perfect imitation of an average human

Artificial psychology should not be evaluated only by whether its average response matches an average human response. Human behavior varies across individuals, cultures, histories, relationships, and circumstances. A system may be psychologically inadequate while reproducing a population majority through statistical imitation. Conversely, a system may produce a minority or non-canonical action through a coherent psychological pathway.

The stronger objective is to model the distribution of psychologically viable states and actions, identify the information needed to distinguish among them, and express uncertainty when that information is absent. Human similarity is important where human behavior is the target, but it must be evaluated at the appropriate individual, group, temporal, and contextual level.

2.3.4 Perfect present-day prediction

Category membership does not require 100% action accuracy. Prediction is a property of a particular model, checkpoint, information set, population, and evaluation design. Architectural completeness and predictive performance are related but distinct. A scientifically valuable architecture may initially be limited by observation quality, supervision, calibration, memory, training, or parameter estimation.

Complete behavioral fidelity remains the SQC program's north star. The validation framework does not redefine that ambition downward. It establishes the measurements needed to determine how closely SQC approaches it and whether remaining errors arise from theory, implementation, information, ambiguity, or irreducible variation.

2.3.5 A single canonical action for every psychological state

The same psychological organization may permit more than one coherent action. Two people may share an appraisal and motivational conflict yet select different valid responses because of memory, skill, opportunity, social commitment, or uncertainty. Requiring a single label can misclassify legitimate plurality as model error.

Artificial-psychology evaluation must therefore include exact-action accuracy where a unique target is justified, valid-action-set probability where several responses are defensible, outcome quality, calibration, and pathway coherence. The architecture should become more certain when discriminating information is available and appropriately less certain when it is not.

2.3.6 Immediate operational readiness

An architecture can qualify as an artificial psychology before it is safe or reliable enough for consequential deployment. Operational validation is use-case-specific. A research system used to test psychological hypotheses does not require the same error tolerance, monitoring, consent, fairness, security, or appeal mechanisms as a system used to allocate rewards, guide care, or support governance.

This distinction permits research and prototyping without allowing architectural classification to be mistaken for deployment assurance.

2.4 Key distinctions used throughout the validation program

2.4.1 Artificial psychology and cognitive architecture

A cognitive architecture specifies general mechanisms through which cognition occurs, commonly including memory, learning, reasoning, attention, goals, and action selection. Such mechanisms occupy neighboring scientific territory because they address components relevant to an artificial psychology. They are not automatically equivalent to one.

Under the definition proposed here, artificial psychology requires an explicit account of how context acquires significance relative to values, motivation, affect, memory, and state, and how that organization generates behavior. A cognitive architecture may satisfy this definition if it implements and tests those functions. The distinction is based on scope and causal organization, not on the research community or label attached to the system.

2.4.2 Artificial psychology and affective computing

Affective computing may recognize, classify, simulate, or respond to emotion. Artificial psychology requires emotion to be situated within a wider psychological process. In SQC, emotion is neither an isolated output nor the sole cause of action. It modulates the interaction among context, values, existential appraisal, motivational drivers, state, memory, and behavior.

An affective system becomes evidence for artificial psychology only when the affective process is causally integrated with the other required functions and validated as part of the behavioral pathway.

2.4.3 Artificial psychology and behavioral prediction

Behavioral prediction estimates what action will occur. Artificial psychology models the internal organization through which an action becomes likely, meaningful, or coherent for a particular agent in a particular situation. A direct predictor may achieve high accuracy by exploiting stable correlations without representing any psychological mechanism.

Prediction remains a necessary validation outcome for SQC because a theory of behavior should generate behaviorally useful expectations. It is not sufficient on its own. SQC's distinctive claim is

that improved behavioral fidelity will ultimately require correct psychological mediation, particularly under distribution change, individualization, counterfactual intervention, and long time horizons.

2.4.4 Artificial psychology and psychometrics applied to AI

Machine psychometrics asks whether psychological attributes can be measured reliably in an artificial agent. These methods are valuable for testing reliability, convergent validity, discriminant validity, criterion validity, and stability. However, measuring a personality or emotional tendency in a model does not establish that the measured construct is part of the mechanism generating behavior.

The SQC program incorporates psychometric methods but adds architectural and causal requirements. A construct must be measurable, and it must occupy the declared role in the generative pathway.

2.4.5 Artificial psychology and artificial consciousness

Artificial psychology concerns the functional organization, dynamics, and behavioral consequences of computational psychological states. Artificial consciousness concerns whether and under what conditions an artificial system has awareness or experience. The questions may eventually interact, but neither can be substituted for the other.

SQC's reference to subjective state is functional and relational: information is represented as significant relative to the system's modeled condition, values, history, and motivational organization. This usage should not be read as an untested declaration of phenomenal experience.

2.4.6 Theory validity and implementation fidelity

Theory validity asks whether EPS accurately describes the organization of human psychology. Implementation fidelity asks whether SQC correctly realizes EPS as specified. These can fail independently. A poor implementation does not falsify EPS, and a faithful implementation does not by itself prove that every EPS proposition is empirically correct.

The validation program therefore maintains a theory-to-code traceability matrix and localizes evidence to the appropriate claim. Architectural audits primarily test implementation fidelity. Human construct, intervention, historical, cross-cultural, developmental, and neural studies test EPS propositions and the meaning of their SQC representations.

2.4.7 Architecture validation and checkpoint validation

An architecture defines permissible representations and computations. A checkpoint is one trained parameterization of that architecture. A checkpoint can perform poorly because of data, optimization, initialization, or selection without demonstrating that the architecture is incapable. Conversely, one favorable checkpoint cannot validate an architecture across populations and domains.

All validation claims must therefore identify whether they concern:

- the EPS theory;

- the SQC architecture;
- the source implementation;
- a particular checkpoint;
- a dataset or supervision protocol;
- an evaluation instrument;
- a population, culture, domain, or intended use.

2.4.8 Representational completeness and predictive certainty

Representational completeness concerns whether the architecture contains the constructs and relations necessary to express the determinants of behavior. Predictive certainty concerns how concentrated and accurate a prediction can be given a particular information state. A complete representational framework may still be uncertain when relevant personal history or context is unobserved. An incomplete framework may appear certain because it is overconfident or because the test distribution is narrow.

SQC's north star joins rather than confuses these objectives. EPS is proposed as the route toward representational completeness. SQC must then demonstrate that the representation produces increasingly accurate and calibrated predictions as relevant information, memory, measurement, and learning improve.

2.4.9 Internal, external, and operational validity

Internal validity asks whether SQC follows the intended EPS pathway and whether controlled changes have the predicted internal effects. External validity asks whether the constructs and behavioral relations generalize beyond the training instrument to independent human, cultural, historical, and domain evidence. Operational validity asks whether a particular implementation is sufficiently reliable, calibrated, safe, governed, and monitored for a declared use.

These are cumulative but non-interchangeable. Strong internal evidence cannot replace external studies. External behavioral correlation cannot replace causal audits. Neither automatically supplies the use-case assurance required for deployment.

2.5 Units and language of validation claims

Validation is not an intrinsic permanent property attached to the word SQC. Every validation statement must identify the interpretation, evidence, version, and scope being supported. The preferred form is:

> *Evidence E supports interpretation I for SQC architecture or checkpoint V, evaluated on population or domain D under protocol P, with uncertainty U and stated limitations L.*

For this reason, the manuscript uses four evidence states:

- **supported:** the prespecified evidence threshold has been met;

- **provisionally supported:** positive evidence exists but replication, precision, or scope remains insufficient for full passage;
- **inconclusive:** the available study cannot reliably distinguish support from failure;
- **not supported:** an adequately designed test failed the prespecified requirement or contradicted the predicted result.

"Not yet tested" is reported separately and must not be silently converted into "not supported." Similarly, "supported" must not be generalized beyond the model version, constructs, populations, and domains actually evaluated.

The same discipline applies to the word **validated**. In this manuscript, validated never means universally correct, permanently complete, or free of uncertainty. It means that a named interpretation or intended use has accumulated the prespecified evidence required at a declared validation level. Validation can be strengthened, narrowed, suspended, or revised as new evidence appears.

This language permits a precise statement of SQC's present status. SQC is an implemented artificial-psychology architecture because it realizes the required functional and causal organization. Its present evidence supports several internal architectural and behavioral claims. Other claims, especially broad human construct validity, long-horizon psychological continuity, cross-cultural measurement equivalence, and operational dependability, occupy later validation stages. The presence of those later stages does not erase the architecture's current classification or completed evidence.

3 EPS as the Scientific Foundation of SQC

SQC is not an architecture to which psychological terminology was added after development. It is a computational implementation of Evolutionary Psychological Structures (EPS), a prior theory of how human beings transform social and environmental information into psychologically organized behavior. The distinction matters for validation. If the internal stages had been chosen only because they improved prediction on the SQC dataset, their meanings would be data-dependent engineering interpretations. Because EPS specifies the constructs and their relations before SQC is trained, the implementation can be tested for fidelity to an independently stated theory.

The source EPS manuscripts present two connected levels of explanation. The socialization account describes the psychological organization through which belonging, identity, power, emotion, existential condition, and social experience shape behavior. The neurocomputational account proposes how contextual filtering, value retrieval, motivational activation, emotional modulation, and state stabilization may be represented as a layered dynamical process (Khanyile, 2026a, 2026b). SQC converts this theoretical sequence into an explicit trainable architecture.

This section identifies the EPS claims that SQC is intended to implement, describes the evidentiary foundations from which the theory was developed, specifies how historical and civilizational observations can be converted into formal validation studies, and establishes the traceability standard required to connect theory, code, data, and empirical evidence.

3.1 EPS theory claims

3.1.1 Behavior is the expression of an organized internal condition

EPS begins from the proposition that behavior is not adequately explained as a direct response to an external stimulus. The same event can produce protection, avoidance, affiliation, negotiation, assertion, submission, exploration, or no action depending on how it is interpreted by the person experiencing it. The relevant differences include prior experience, present values, perceived threat, social position, identity, belonging, available power, emotional modulation, expected outcomes, and current existential condition.

Behavior is therefore modeled as the external expression of an internal psychological organization. This does not imply that behavior reveals that organization without ambiguity. Different internal states can sometimes produce similar actions, and a single state can permit several valid actions. The claim is causal rather than one-to-one: the process that organizes the person's relation to the situation constrains and generates the behavioral possibilities.

For SQC, this proposition creates a non-negotiable architectural requirement. The action system must operate on the constructed EPS state and memory. If raw observation or an unrestricted hidden representation can bypass that state, the system may predict behavior without testing the EPS explanation of behavior.

3.1.2 Context is interpreted rather than merely received

EPS distinguishes the occurrence of an event from its psychological meaning. An observation becomes behaviorally relevant only after it is interpreted in relation to the agent, situation, social structure, history, and anticipated future. An identical resource loss, rejection, instruction, risk, or opportunity may signify something different under a different role, goal, relationship, obligation, or tolerance for risk.

Contextual interpretation is therefore the first constructive stage of the EPS sequence. It determines which features of an observation are psychologically salient and provides the organized situation to which values and later stages respond. The crossed role-goal-risk design in SQC v1.2 is an initial test of this claim: the observation cannot determine the action independently of the decision context, yet the decision context is not permitted to enter the action mapper as a shortcut.

3.1.3 Values determine why information matters

Values provide the evaluative structure through which interpreted information becomes significant. They connect an event to what may be protected, lost, preserved, fulfilled, violated, or made possible. The current SQC ontology includes dimensions concerning safety, physical harm, uncertainty, resources, loss, future optionality, belonging, rejection, loyalty, fairness, duty, authority, status, dominance, dignity, self-consistency, and authenticity.

EPS does not treat values as detached preferences added after cognition. Values participate in determining the psychological importance of the context and the pressures transmitted downstream. Their activation should therefore change existential appraisal, motivational organization, emotional

modulation, state, and action in theoretically predictable ways.

3.1.4 Existential appraisal organizes implications for continued condition

EPS proposes that psychologically meaningful information is evaluated relative to existential consequences. The current implementation represents continuous implications for survival, stability, and quality of life before predicting a categorical existential state. This stage asks not only whether an event is positive or negative, but what kind of condition it creates or threatens for the agent.

The distinction between appraisal and state is important. Existential appraisal represents the distributed implications detected in the interpreted and valued situation. Existential state represents the later stabilized orientation through which action is organized. A context may have simultaneous implications for survival, stability, and quality, while one condition becomes dominant for the present decision.

3.1.5 Belonging, Identity, and Power organize motivational response

EPS identifies Belonging, Identity, and Power as evolved motivational organizers of social behavior.

- **Belonging** concerns attachment, inclusion, relational continuity, acceptance, loyalty, and the conditions under which the person remains part of a valued social field.
- **Identity** concerns self-consistency, role, meaning, dignity, authenticity, recognition, and the continuity of who the person understands themselves to be.
- **Power** concerns agency, control, capacity, status, influence, protection, resource access, and the ability to affect outcomes.

These drivers are analytically distinct but dynamically interdependent. An action that protects belonging may threaten identity. An assertion of power may damage stability. Submission may preserve survival while producing identity cost. The theory therefore predicts behavior through the configuration and conflict of the drivers rather than through a single universal reward.

3.1.6 Emotion modulates the transition from implication to state and action

EPS treats emotion as computation within the causal pathway rather than as a descriptive label attached to an outcome. Emotional modulation changes the gain, urgency, orientation, and learning consequences of the active psychological configuration. In SQC, the current emotion representation includes gain, urgency, valence, and a learning gate.

This claim generates direct intervention predictions. Increasing urgency while holding the interpreted context constant should alter the concentration or timing of action selection when the theory predicts an immediate threat. Changing valence or gain should affect downstream state and action in specified directions. If emotion can be randomized or removed without changing any theoretically relevant behavior, its claimed causal role would not be supported.

3.1.7 Existential state stabilizes the action-guiding orientation

The current SQC implementation distinguishes Survival, Stability, and Quality of Life as categorical existential states. These states represent the dominant condition organizing the present action. Survival prioritizes protection of continued existence or immediate integrity. Stability prioritizes continuity, belonging, coordination, and maintenance of a viable condition. Quality of Life prioritizes aspiration, growth, meaning, learning, and improved future possibility.

The categories should not be interpreted as a rigid hierarchy in which one state permanently replaces another. They are dynamic orientations produced by the current configuration of context, values, implications, drivers, emotion, and memory. The same person can move among them as conditions change.

3.1.8 Memory and socialization make psychology historically situated

EPS is not complete as a sequence of isolated decisions. Human psychological organization develops through repeated experience. Prior loss, recognition, betrayal, protection, success, punishment, obligation, group membership, and cultural learning alter how later events are interpreted. Socialization stabilizes values, identities, expectations, relational models, and behavioral possibilities over time.

The present SQC memory vector is an initial bounded implementation, not the completed EPS account of life history. Full validation requires autobiographical, relational, developmental, and cultural memory capable of changing present construction of state without creating a hidden shortcut that bypasses the declared EPS process.

3.1.9 Outcomes update future psychological construction

Behavior produces consequences that affect the agent and environment. EPS therefore implies a recurrent process:

```
interpreted situation
    -> psychological construction
    -> action
    -> outcome
    -> prediction error and memory update
    -> changed future interpretation and behavior
```

Validation must test this loop, not only single-step classification. The system should learn from outcomes while preserving coherent values, identity, relationships, and state transitions. Adaptation that improves one domain by erasing previously valid psychological organization is not evidence of a fully successful psychology.

3.2 Sources of evidence underlying EPS

EPS was developed as a synthesis of recurring patterns across several bodies of evidence rather than as a theory inferred from the current SQC training corpus. The corpus operationalizes a bounded part of the theory for computational testing; it is not the origin of the theory.

3.2.1 Psychological and behavioral literature

The constructs assembled by EPS intersect with established research on values, appraisal, emotion, motivation, identity, belonging, social cognition, reinforcement, prediction, and action. This literature supports the scientific plausibility of treating human behavior as value-grounded, affectively modulated, socially organized, and historically conditioned (Lazarus, 1991; Schwartz, 1992; Damasio, 1994; Picard, 1997; Scherer et al., 2001; Gratch and Marsella, 2004).

Literature agreement provides theoretical grounding and content evidence. It does not automatically validate the exact EPS ontology, causal order, or parameterization. Those distinctive propositions must be tested directly.

3.2.2 Evolutionary and developmental evidence

EPS interprets Belonging, Identity, and Power as organizers that address recurrent adaptive problems: maintaining social inclusion, preserving a coherent and recognized self, and retaining sufficient capacity to affect conditions. Developmental and socialization processes shape how these general pressures are expressed, prioritized, inhibited, and linked to culturally available actions.

Validation requires more than showing that these themes appear in human life. The theory must predict developmental sequences, conflict patterns, individual differences, responses to controlled changes, and conditions under which one driver becomes more influential than another.

3.2.3 Neuroscientific and physiological evidence

The EPS neurocomputational formulation relates stages of the theory to classes of cortical, subcortical, neuromodulatory, valuation, affective, and learning processes. These relations are hypotheses at an appropriate systems level, not a declaration that an SQC tensor is identical to a biological structure.

Neural and physiological validation can test whether the temporal ordering, directional effects, modulation patterns, and impairment predictions of EPS correspond to human evidence. Agreement would strengthen the theory. Repeated disagreement would require revision of the relevant mapping or causal claim.

3.2.4 Cross-cultural evidence

Human societies differ in norms, roles, institutions, expressive conventions, and the actions available for protecting or advancing what matters. They also repeatedly organize life around affiliation, identity, agency, threat, continuity, aspiration, obligation, recognition, and control. EPS proposes that the underlying organizers may recur while their weights, meanings, triggers, and behavioral expressions are culturally parameterized.

This is a measurement-invariance hypothesis, not an assumption. Cross-cultural validation must determine which relations remain stable, which require local parameters, and whether the current ontology omits constructs necessary in particular cultural settings.

3.2.5 Historical and civilizational behavioral evidence

Civilization contains a vast accumulated record of human behavior under scarcity, conflict, cooperation, migration, institutional formation, technological change, status competition, identity construction, cultural stabilization, collapse, and recovery. Across this record, people and groups repeatedly act in relation to survival, social continuity, meaning, belonging, identity, power, and anticipated quality of life. This recurrence is a major source of the EPS research hypothesis.

The civilizational record is scientifically important because it contains behavioral variation that no bounded laboratory study could reproduce in full. It includes natural experiments across institutions, environments, time scales, and population structures. It can reveal whether EPS provides a coherent explanation across apparently different domains and whether the same organizing relations recur under transformed surface conditions.

However, the existence of a large record is not by itself a completed validation. Historical cases are vulnerable to selective example choice, retrospective explanation, missing information, non-independence, and competing interpretations. The strength of civilizational evidence depends on converting the record into a transparent research instrument.

3.2.6 Convergence across evidence types

No single evidence source is expected to validate EPS completely. Literature can show theoretical continuity but may not distinguish the EPS sequence. Laboratory studies can identify causal effects but may have limited ecological scope. Historical records provide breadth but reduced experimental control. Neural evidence can constrain mechanism but not determine social meaning alone. Behavioral prediction can demonstrate utility but may be achieved through the wrong pathway.

The strongest EPS case therefore depends on triangulation. A construct is supported when its definition is theoretically coherent, independently recognizable, reliably measurable, causally active under intervention, predictive of relevant outcomes, distinguishable from adjacent constructs, and stable or systematically variable across populations and time.

3.3 Converting civilizational evidence into formal tests

The claim that thousands of years of human behavior support EPS should be made testable rather than left as an appeal to intuitive historical recognition. The validation program will construct a Civilizational EPS Evidence Protocol with the following components.

3.3.1 Define the unit of analysis

The protocol must distinguish among individual decisions, repeated relational episodes, group behavior, institutional decisions, cultural movements, and long-duration societal transitions. EPS variables may operate differently across these levels. An institutional action cannot automatically be treated as though it were the action of a single psychological agent.

Each study must declare:

- the acting unit;
- the observation available to that unit at the time;
- the relevant history and relationships;
- the decision context;
- the candidate values, drivers, emotions, and existential implications;
- the action set and observed outcome;
- the temporal resolution;
- the uncertainty and missing evidence.

3.3.2 Establish preregistered sampling rules

Cases must not be chosen only because they illustrate EPS clearly. Sampling frames should be specified before coding and should include outcomes expected to support, challenge, and fall outside the current theory. Candidate designs include stratified random samples from historical databases, complete event series within a defined period, paired cases with similar conditions but different outcomes, and prospective registration of contemporary events.

The sampling plan should balance regions, periods, institutional forms, event types, and outcome classes. Dependence among events must be modeled rather than treated as independent evidence.

3.3.3 Create a versioned coding ontology

Historical information must be translated into EPS variables through a public coding manual. Each construct requires inclusion and exclusion criteria, positive and negative examples, ambiguity rules, missing-data codes, and allowed evidence sources. Coders should not infer an internal state solely from the action that the state is later used to explain.

Multiple independent coders should be blinded, where possible, to the target outcome and SQC prediction. Agreement should be reported for every construct, and disagreements should remain visible rather than being removed through silent adjudication.

3.3.4 Freeze predictions before revealing outcomes

The strongest historical design is retrodictive blind prediction. SQC receives only information documented as available before a defined decision or transition. Its internal-state and behavioral predictions are frozen before the subsequent event is revealed. This converts a retrospective narrative into a test with a clear information boundary.

Where full blinding is impossible, the analysis should be labeled exploratory and followed by confirmation on untouched cases.

3.3.5 Compare competing explanations

EPS should not be evaluated only against chance. Comparisons should include historical base rates, domain-expert forecasts, direct observation-to-outcome models, alternative psychological explanations, institutional or material models, and hybrid accounts. The objective is not to show that psychological organization is the only influence on civilization. It is to measure whether EPS adds explanatory and predictive value after relevant non-psychological conditions are represented.

3.3.6 Test recurrence, boundary, and failure

The protocol should identify:

- relations that recur across settings;
- relations that change systematically with culture or institution;
- cases that require individual or local parameterization;
- cases that expose missing constructs;
- cases in which EPS makes the wrong directional prediction;
- cases in which several explanations remain observationally equivalent.

Publication must include counterexamples and boundary cases. A theory of broad human behavior becomes stronger by identifying where its present form fails, not by absorbing every outcome through retrospective reinterpretation.

3.4 Theory-to-code traceability

Every material SQC claim must be traceable across six layers:

```
EPS proposition
    -> formal construct and predicted relation
    -> SQC module and tensor representation
    -> dataset field and supervision rule
    -> diagnostic or intervention
    -> human or external validation study
```

The traceability matrix prevents three errors. First, it prevents engineering features from being presented as though they were EPS propositions. Second, it prevents theoretical constructs from being claimed as implemented when no corresponding mechanism exists. Third, it prevents high final-action accuracy from substituting for evidence about intermediate psychological meaning.

An initial matrix for the present architecture is shown below.

EPS construct or relation	Current SQC realization	Current supervision or diagnostic	Required external validation
Contextual interpretation	Explicit context state derived from observation, decision context, and memory	Context targets, evidence fields, crossed-context tests	Human context judgments, invariance tests, prospective context manipulation
Values	Seventeen-dimensional value state	Base and context-shaped values, value MSE, oracle-value interventions	Content, convergent, discriminant, and predictive validity
Existential appraisal	Continuous survival, stability, and quality implications	Existential targets, MSE, oracle interventions	Human appraisal measures, directional manipulation, outcome association
Belonging, Identity, and Power	Three-dimensional driver state	Driver targets and stagewise loss	Independent measures, conflict experiments, cross-cultural equivalence
Emotional modulation	Gain, urgency, valence, and learning gate	Emotion targets and stagewise loss	Multimethod affect measures, physiological correspondence, interventions
Existential state	Categorical Survival, Stability, or Quality of Life state	State targets, error rate, oracle-state intervention	Human state annotation, temporal stability, criterion validity
Behavior from internal state	EPS-state-and-memory-only action mapper	Action-input audit, action accuracy, valid-action rate	Prospective human behavior, matched baselines, pathway-consistency tests
Memory and socialization	Bounded explicit EPS memory	Memory fields and trajectory transitions	Autobiographical, relational, developmental, and long-horizon studies
Outcome-driven adaptation	Outcome heads, prediction error, later memory input	State-after, state-change, latent-delta, and prediction-error targets	Longitudinal learning, recovery, reinforcement, and stability studies

The full validation manuscript will extend this table to include the precise EPS source passage, formal notation, code location, configuration, tensor shape, loss ownership, data-generation rule, unit test, checkpoint audit, current evidence status, and planned decisive study.

3.5 Distinguishing foundation, implementation, and validation

EPS supplies the scientific foundation of SQC in three senses. It defines the psychological constructs. It proposes the relations through which those constructs organize behavior. It provides the reason SQC is structured as a sequential, state-mediated architecture rather than an unconstrained predictor.

This foundation is stronger than an analogy. SQC's architecture, data contract, loss ownership, training sequence, interventions, and shortcut restrictions are all consequences of the EPS theory. The current computational results therefore constitute genuine evidence that a bounded EPS process can be implemented, trained, separated causally, and used to generate behavior.

Scientific grounding and completed validation must nevertheless remain distinguishable. A theory can be well grounded in existing knowledge and broad observation while still containing distinctive propositions requiring direct tests. The civilizational record can provide extraordinary breadth while still requiring formal sampling and coding. A faithful implementation can prove that the theory is computationally realizable without proving that its internal variables correspond perfectly to human psychological constructs.

These distinctions do not place SQC at the beginning of the validation process. They locate its evidence correctly. SQC has a prior theory, an implemented causal architecture, controlled internal supervision, shortcut-prevention evidence, causal diagnostics, and behavioral results. The next stages convert the wider EPS evidence base into formal construct, human, historical, cross-cultural, longitudinal, comparative, and operational tests.

4 SQC Architecture and Claims Under Evaluation: Public Summary

> **Public-disclosure note.** This section presents the scientific purpose, validation boundaries, and claim structure of SQC without disclosing proprietary implementation details. Module specifications, internal dimensions, configuration values, training controls, data schemas, code-level enforcement mechanisms, checkpoint internals, and other trade-secret material are omitted. Their omission does not indicate that the corresponding validation work is unnecessary; it indicates that detailed review must occur through an appropriately governed confidential process.

4.1 Architectural purpose

SQC is designed to implement the central EPS proposition that behavior emerges from an organized psychological process rather than from an unrestricted direct mapping between an observation and an action. At a public conceptual level, the architecture interprets context, represents what matters in that context, appraises its implications, organizes motivational and affective responses, constructs an action-relevant state, and uses memory to maintain continuity over time. Behavior is generated from this constructed state.

The scientific significance of this design lies in its testability. The architecture makes it possible to distinguish a correct behavioral prediction from a prediction reached through the intended psychological pathway. It also supports interventions that test whether changes to an internal construct have the downstream effects predicted by EPS. The public claim is therefore about the existence and validation of an explicit, state-mediated psychological organization, not about the disclosure of its proprietary computational realization.

4.2 Validation objects

Evidence concerning SQC must identify the object to which it applies. The EPS theory, the SQC architecture, a software implementation, a trained model, a dataset, an evaluation protocol, and an

intended operational use are related but distinct validation objects. Evidence for one object does not automatically validate the others.

For example, faithful implementation supports the conclusion that EPS has been realized computationally, but does not by itself prove that every EPS construct accurately represents human psychology. Likewise, strong results from one trained model do not establish that every training run or later version will perform equivalently. Operational use requires additional evidence concerning reliability, uncertainty, safety, governance, and the affected population.

4.3 Claims under evaluation

The validation program evaluates the following public claim classes:

1. **Implementation fidelity:** the implemented system realizes the declared EPS organization without material contradiction.
2. **Causal mediation:** represented psychological states participate materially in producing behavior rather than serving only as labels or explanations added after prediction.
3. **Internal semantic validity:** internal states respond coherently to controlled changes and exhibit the relations attributed to them.
4. **Human construct validity:** the represented states correspond to their intended psychological meanings in appropriately designed human studies.
5. **Behavioral fidelity:** the system predicts behavior accurately, coherently, and with calibrated uncertainty within a declared scope.
6. **Generality and continuity:** supported behavior extends across new situations, people, domains, cultures, relationships, and time without erasing relevant differences.
7. **Comparative value:** explicit EPS mediation provides measurable benefit when compared fairly with serious alternatives.
8. **Operational dependability:** a named version is sufficiently reliable, safe, governed, monitored, and contestable for a declared use.
9. **Downstream validity:** systems relying on SQC significance computation satisfy their own measurement, infrastructure, governance, and operational requirements.

These claims are cumulative but non-interchangeable. Behavioral accuracy alone cannot establish the correctness of the internal psychology. Architectural fidelity cannot replace human construct validation. Research evidence cannot substitute for use-specific operational assurance.

4.4 Evidence boundaries and confidential assurance

Public evidence should report the claim tested, model or system version, population or domain, evaluation protocol, uncertainty, limitations, and replication status. It should also state whether a result concerns theory, architecture, implementation, a trained model, or an intended use. Favorable evidence must not be generalized beyond the object and scope actually tested.

Where validation requires inspection of non-public implementation details, review may be performed under confidentiality by qualified independent auditors. A public assurance report should describe the audit scope, methods, independence, findings, limitations, and version identifiers without revealing protected material. Confidential review is not a substitute for evidence; it is a mechanism

for evaluating sensitive evidence while maintaining legitimate intellectual-property protections.

The public manuscript therefore makes a bounded claim: SQC can be evaluated as an explicit artificial-psychology architecture through causal, construct, behavioral, comparative, and operational evidence. The detailed mechanisms by which the current implementation realizes that architecture remain outside the scope of this public disclosure.

5 SQC Artificial Psychology Validation Scale

The SQC Artificial Psychology Validation Scale provides a cumulative framework for classifying the maturity of an artificial-psychology system. It replaces the undefined binary question of whether artificial psychology has been "solved" with a sequence of explicit evidentiary levels. Each level answers a different scientific question and authorizes a different class of claims.

The scale is designed initially for SQC, but its criteria are architecture neutral. Another system may be evaluated under the same scale if it declares a psychological ontology, causal process, behavioral pathway, and validation scope. The scale does not assume that every artificial psychology must use the EPS ontology. It does require every claimant to state what psychological organization is implemented and to subject that organization to causal, construct, behavioral, and external tests.

5.1 Overview of the validation levels

Level	Classification	Central question	Minimum evidentiary achievement
AP-0	Theoretically specified artificial psychology	Is there an explicit, falsifiable theory of psychological organization?	Defined constructs, relations, scope, predictions, and boundary conditions
AP-1	Implemented artificial-psychology architecture	Has the theory been realized as an executable causal architecture?	Theory-to-code mapping, functioning modules, behavioral output, and inspectable state
AP-2	Internally validated artificial-psychology architecture	Does the implementation use the declared psychological pathway reliably?	Shortcut prevention, causal intervention, internal supervision, reproducibility, and bounded generalization
AP-3	Construct-validated artificial-psychology system	Do the implemented states have their intended psychological meanings?	Content, convergent, discriminant, structural, criterion, and intervention validity
AP-4	Ecologically validated artificial-psychology system	Does the psychology generalize across natural people, settings, cultures, and time?	Prospective human, cross-domain, cross-cultural, longitudinal, and independent replication

Continued on next page

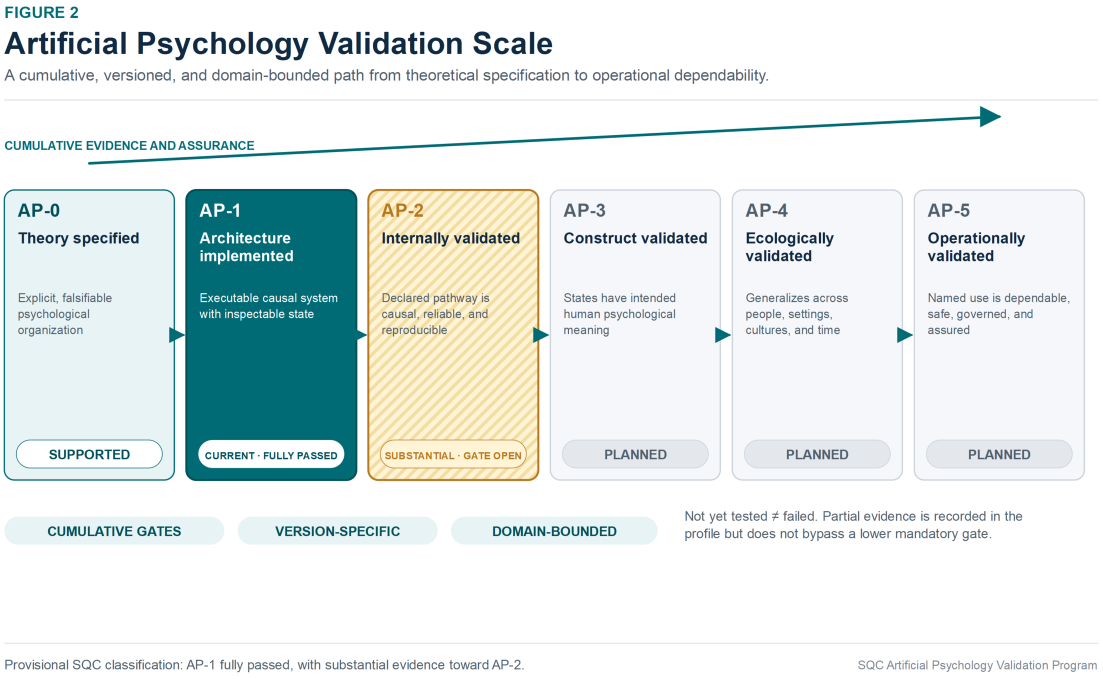

Figure 2: Artificial Psychology Validation Scale. The levels are cumulative, version-specific, and domain-bounded. SQC's highest fully passed level is AP-1, with substantial evidence toward AP-2; AP-3 through AP-5 are planned rather than failed.

Level	Classification	Central question	Minimum evidentiary achievement
AP-5	Operationally validated artificial-psychology system	Is a named system dependable for a declared consequential use?	Use-case reliability, calibration, safety, governance, monitoring, and independent assurance

The classification names are intentionally cumulative. AP-1 does not mean an architecture has weak or merely conceptual status; it means the theory exists in executable form. AP-2 does not mean human construct validity has been completed; it means the architecture has survived internal causal and implementation tests. AP-3 concerns psychological meaning. AP-4 concerns ecological breadth. AP-5 concerns operational assurance for a named use.

5.2 General classification rules

5.2.1 Levels are cumulative

A system must satisfy the mandatory requirements of the lower levels before it can receive a higher classification. A psychologically correlated black-box predictor cannot enter directly at AP-3 because construct correlation alone does not establish an artificial-psychology architecture. Likewise, operational performance in a narrow application cannot substitute for architectural and construct

evidence.

Evidence can be accumulated in parallel, but classification is awarded only when the mandatory lower gates are complete. A project may therefore possess promising AP-3 human results while remaining formally at AP-2 because a required reproducibility audit is incomplete.

5.2.2 Classification is versioned

Every classification applies to a named theory version, architecture version, implementation revision, and where relevant, checkpoint. A material change to the ontology, causal ordering, action-input contract, memory process, or outcome learning rule creates a new validation object.

Minor engineering changes may inherit prior evidence if an impact analysis shows that the validated behavior and causal pathway are unchanged. Material changes require targeted revalidation and may require full repetition of a level.

5.2.3 Classification is domain-bounded

AP status must identify the populations, cultures, languages, behavioral domains, time horizons, and action spaces represented in the evidence. AP-4 in a bounded coordination domain does not mean universal human psychological validity. AP-5 in a low-risk research tool does not authorize high-stakes psychological assessment or economic allocation.

The preferred label is therefore specific, for example:

> *SQC version V is classified AP-3 for constructs C, populations P, languages L, domains D, and horizon T under validation protocol R.*

5.2.4 Classification uses profiles as well as levels

A single level cannot communicate all strengths and weaknesses. Every classification must be accompanied by an evidence profile covering at least:

- architectural fidelity;
- causal mediation;
- internal-state accuracy;
- psychological construct validity;
- behavioral fidelity;
- calibration and uncertainty;
- compositional and out-of-ontology generalization;
- temporal continuity and memory;
- cross-cultural evidence;
- comparative evidence;
- reproducibility;
- operational assurance.

The level reports the highest fully passed cumulative gate. The profile shows partial progress above that level and weaknesses within the tested scope.

5.2.5 Absence of evidence is not evidence of failure

An unperformed study is recorded as **not yet tested**, not as failed. This rule is especially important for a new field lacking an inherited validation standard. SQC is not disqualified as an artificial-psychology architecture because the validation scale is being constructed after its first implementation.

At the same time, unperformed studies cannot be counted as passed. The scale recognizes current status without converting anticipated results into evidence.

5.2.6 Failure is localized

A failed study changes the status of the claim and level requirement it tests. It does not automatically invalidate the entire EPS theory or SQC architecture. The failure must first be localized among:

- theoretical specification;
- construct definition;
- architecture;
- source implementation;
- data or annotation;
- training and optimization;
- checkpoint selection;
- measurement instrument;
- domain transfer;
- operational process.

Repeated adequately powered failures of a preregistered EPS prediction count against the relevant theory claim. Repeated implementation failures count against implementation fidelity. This localization must not become a mechanism for protecting the theory indefinitely: revision rules in Section 17 specify when accumulated failures require theoretical change.

5.2.7 Thresholds are declared before confirmatory evaluation

This manuscript defines the kinds of evidence required at each level. Numeric thresholds must be set through pilot data, measurement reliability, baseline performance, expected effect sizes, power analysis, risk tolerance, and domain requirements before a locked confirmatory test is opened.

No threshold may be chosen after observing the final result merely to promote a preferred checkpoint or validation level. Exploratory findings may inform the next preregistered threshold but cannot be relabeled as confirmatory.

5.3 AP-0: Theoretically specified artificial psychology

5.3.1 Scientific question

> *Does the proposal contain a coherent, explicit, and falsifiable theory of psychological organization capable of computational realization?*

AP-0 is the entry point for an artificial-psychology research program. It distinguishes a scientific theory from an evocative collection of psychological terms.

5.3.2 Mandatory requirements

1. **Construct definitions.** The theory defines its psychological constructs, their intended meanings, and their boundaries.
2. **Causal relations.** It specifies which constructs influence which others and in what temporal or functional order.
3. **Behavioral connection.** It explains how the psychological organization constrains or generates behavior.
4. **Memory and adaptation account.** It states how prior experience can alter future psychological construction.
5. **Scope.** It identifies the phenomena, populations, and levels of analysis the theory intends to explain.
6. **Falsifiable predictions.** It makes predictions that can be contradicted, including directional, comparative, or boundary predictions.
7. **Competing accounts.** It identifies plausible alternatives or specifies evidence that would favor a different organization.
8. **Theory-version control.** Its definitions and predictions are preserved in a citable, versioned specification.

5.3.3 Permitted claims

- the work defines a theory or proposed architecture for artificial psychology;
- the theory integrates specified psychological functions;
- the theory generates a declared validation program;
- the theory is grounded in identified scientific and observational evidence.

5.3.4 Claims not yet permitted

- the theory has been implemented successfully;
- the proposed process actually mediates behavior in code;
- the constructs have been measured reliably in humans;
- the theory predicts natural human behavior accurately;
- the system is ready for use.

5.3.5 EPS status

EPS satisfies the principal AP-0 requirements. Its source manuscripts define a causal theory connecting contextual interpretation, values, existential appraisal, Belonging, Identity, Power, emotion, state, behavior, socialization, and memory. The present validation manuscript strengthens the falsification, scope, competing-model, and version-control requirements that were not previously organized as a single formal gate.

Provisional determination: **AP-0 supported**.

5.4 AP-1: Implemented artificial-psychology architecture

5.4.1 Scientific question

> *Has the declared psychological theory been realized as an executable, inspectable architecture in which internal psychological state can generate behavior?*

AP-1 establishes category membership at the system level. It does not depend on human-level performance. It requires an operational architecture rather than a conceptual diagram.

5.4.2 Mandatory requirements

1. **Executable implementation.** The system receives inputs, constructs internal state, and generates behavior.
2. **Theory-to-code traceability.** Every central construct and causal relation maps to a computational representation or is explicitly marked unimplemented.
3. **Inspectable states.** Internal constructs can be recorded and associated with declared meanings.
4. **Causal pathway.** The implemented dataflow follows the stated architecture.
5. **Behavioral readout.** Action is generated from the internal psychological process.
6. **Temporal interface.** The architecture can represent history, update, or a declared bounded precursor to the complete memory process.
7. **Test coverage.** Unit and integration tests exercise the core modules and interfaces.
8. **Versioned artifacts.** Code, configuration, ontology, and a runnable model or reproducible initialization are identifiable.

5.4.3 Permitted claims

- the system is an implemented artificial-psychology architecture under the operational definition;
- the architecture computationally realizes the declared bounded theory;
- internal psychological constructs are explicit and testable;
- the system can generate behavior through the implemented pathway.

5.4.4 Claims not yet permitted

- all declared constructs have their intended human psychological meaning;

- the architecture reliably prevents every shortcut;
- the implementation generalizes beyond its development instrument;
- the system predicts human behavior with established external validity;
- the architecture is comparatively superior or operationally dependable.

5.4.5 SQC status

SQC v1.2 implements the declared sequential EPS pathway. It exposes explicit context, value, existential-appraisal, driver, emotion, existential-state, memory, action, and outcome representations. Its action mapper consumes the concatenated EPS action state and memory. Stagewise module ownership and diagnostic interfaces permit direct inspection and intervention.

The present source implementation, configuration, tests, dataset contract, and trained checkpoints establish executable realization. The bounded explicit memory does not complete the EPS life-history account, but it supplies a declared temporal interface and a traceable target for later expansion.

Provisional determination: **AP-1 supported**.

5.5 AP-2: Internally validated artificial-psychology architecture

5.5.1 Scientific question

> *Does the implemented system reliably use the declared psychological pathway, preserve its module boundaries, respond coherently to internal intervention, and generalize beyond memorized training combinations?*

AP-2 validates the architecture from within its controlled experimental domain. It is the level at which psychological organization must become demonstrably causal rather than merely visible.

5.5.2 Mandatory requirements

1. **Shortcut prevention.** Prohibited raw inputs, labels, hidden ground truth, and future outcomes cannot support the behavioral readout.
2. **Gradient and parameter isolation.** Training ownership and freeze claims are verified directly.
3. **Causal intervention.** Manipulating a stage produces preregistered downstream changes when the theory predicts sensitivity.
4. **Negative controls.** Irrelevant interventions do not create material changes where invariance is predicted.
5. **Internal target performance.** Construct outputs achieve prespecified accuracy, error, calibration, or recoverability requirements.
6. **Behavioral competence.** The complete predicted-state pathway performs above prespecified baselines and trivial strategies.
7. **Bounded generalization.** The system succeeds on locked combinations, scenarios, or trajectories absent from training.
8. **Multi-seed reliability.** The principal architecture result repeats across independently trained checkpoints.

9. **Artifact reproducibility.** An independent run reproduces the declared result from versioned artifacts.
10. **Failure analysis.** Correct actions through incorrect pathways and valid alternatives are measured explicitly.

5.5.3 Permitted claims

- the architecture uses the declared internal psychological pathway in the tested implementation;
- behavior is mediated by explicit internal state under the audited contract;
- the architecture is trainable and reproducible in the tested domain;
- internal interventions produce predicted bounded effects;
- the architecture composes or transfers across the tested controlled changes.

5.5.4 Claims not yet permitted

- internal variables are fully validated human psychological constructs;
- synthetic or benchmark results generalize across natural populations;
- the ontology is culturally universal or representationally complete;
- SQC uniquely outperforms all alternative architectures;
- the system is suitable for consequential deployment.

5.5.5 SQC status

Current evidence supports a substantial portion of AP-2:

- the EPS-only action-input contract is implemented;
- raw observation, encoder hidden state, and raw decision context are excluded from the action mapper;
- action and inter-stage detachment are configured and testable;
- freeze audits report zero changed non-action tensors in action-only adaptation;
- stagewise oracle interventions produce measurable action changes;
- the predicted-EPS pathway reaches 75.25% exact action accuracy;
- valid-action hit rate reaches 97.17%;
- eight complete role-goal-risk combinations are absent from training and reach 75.00% aggregate test action accuracy;
- held-out counterfactual, scenario-template, and episode leakage checks report zero overlap;
- multi-seed replay studies demonstrate robust preservation of general action behavior, although external BIB improvement is not robust.

The principal missing AP-2 requirements are multi-seed repetition of the balanced base training, complete adversarial shortcut testing, preregistered negative-control interventions, pathway-consistency analysis, and independent artifact reproduction. The present 29.21% existential-state error also requires explicit threshold treatment rather than being hidden by action performance.

Provisional determination: **AP-2 substantially but not yet fully supported**. The highest fully passed cumulative level is presently AP-1, with strong AP-2 evidence already earned.

5.6 AP-3: Construct-validated artificial-psychology system

5.6.1 Scientific question

> *Do the implemented internal states reliably represent the human psychological constructs attributed to them, and do they participate in human behavior as EPS predicts?*

AP-3 moves beyond implementation coherence. It tests psychological meaning using evidence independent of the theory-grounded synthetic generator.

5.6.2 Mandatory requirements

1. **Content validity.** Independent experts and relevant communities judge the construct definitions and coverage adequate for the declared scope.
2. **Measurement reliability.** Human annotations or instruments achieve prespecified inter-rater, internal-consistency, and test-retest standards.
3. **Convergent validity.** SQC constructs relate to independent measures of theoretically related constructs in the predicted direction and strength.
4. **Discriminant validity.** Distinct SQC constructs remain distinguishable from theoretically unrelated or adjacent measures.
5. **Structural validity.** Human observations support the proposed organization or outperform credible alternative structures.
6. **Criterion validity.** Internal states predict independent behavior or outcomes associated with the construct.
7. **Incremental validity.** EPS states add explanatory or predictive value beyond surface observation, history, demographics, and simpler predictors.
8. **Human intervention validity.** Controlled changes produce the direction of internal and behavioral response predicted by EPS.
9. **Measurement invariance within declared groups.** Scores retain interpretable meaning across the initial populations for which AP-3 is claimed.
10. **Independent confirmation.** At least one qualified independent team or preregistered holdout confirms the core construct results.

5.6.3 Permitted claims

- the named SQC states have accumulated construct-validity evidence in the declared populations and domains;
- the internal pathway reflects measurable human psychological organization;
- SQC supplies incremental psychological explanation or prediction within the validated scope;
- SQC is a construct-validated artificial-psychology system for that scope.

5.6.4 Claims not yet permitted

- universal human validity across cultures and development;
- complete life-history or relationship modeling;
- accurate performance in unconstrained natural environments;

- operational dependability for consequential decisions;
- consciousness or phenomenal experience.

5.6.5 SQC status

The current theory-grounded supervision and causal interventions provide early semantic evidence, but they are not independent human construct validation. The required content panels, human instruments, convergent-discriminant studies, structural comparisons, and controlled human intervention tests remain to be conducted.

Provisional determination: **AP-3 not yet tested as a complete gate; selected components are in preparation**.

5.7 AP-4: Ecologically validated artificial-psychology system

5.7.1 Scientific question

> *Does the construct-validated architecture maintain psychological and behavioral validity across natural people, settings, cultures, relationships, domains, and extended time?*

AP-4 tests whether results survive the complexity that controlled instruments necessarily simplify.

5.7.2 Mandatory requirements

1. **Prospective human prediction.** Predictions are frozen before the target behavior or outcome occurs.
2. **Held-out participants and environments.** Generalization includes people and contexts absent from development and selection.
3. **Cross-domain replication.** Core relations repeat across multiple behavioral domains using independent data sources.
4. **Cross-cultural validation.** The program identifies invariant relations, culturally parameterized relations, and ontology gaps.
5. **Longitudinal validity.** Memory, identity, relationships, learning, and state transitions remain coherent over extended trajectories.
6. **Out-of-ontology testing.** New factor levels, actions, roles, values, and social structures test representational extension.
7. **Historical or civilizational confirmation.** Preregistered or blinded analyses test broad recurrence without retrospective case selection.
8. **Calibration across groups and time.** Uncertainty remains meaningful under distribution change.
9. **Adversarial boundary testing.** Difficult, atypical, deceptive, and contradictory cases are included deliberately.
10. **Independent multi-site replication.** Findings repeat beyond the original research environment.

5.7.3 Permitted claims

- SQC is ecologically validated within the named populations, cultures, domains, and horizons;
- EPS relations have demonstrated cross-context or cross-cultural stability where supported;
- SQC provides prospective behavioral and psychological value outside its development instruments;
- identified boundaries and parameter differences are empirically grounded.

5.7.4 Claims not yet permitted

- validity for all people, cultures, domains, and future conditions;
- certainty when relevant personal or contextual information is unavailable;
- automatic suitability for consequential deployment;
- downstream economic validity without separate measurement and governance studies.

5.7.5 SQC status

BIB-CCS supplies an initial external benchmark and demonstrates that state inference and value consistency can transfer beyond the primary SQC action distribution. The seed-sensitive adaptation results also expose limits. BIB-CCS does not constitute broad ecological validation of human psychology.

Provisional determination: **AP-4 not yet reached; initial external-transfer evidence exists**.

5.8 AP-5: Operationally validated artificial-psychology system

5.8.1 Scientific question

> *Is a named SQC system sufficiently dependable, calibrated, safe, governed, and monitored for a specific real-world use and affected population?*

AP-5 is not one universal certification. It is a family of use-case-specific assurance determinations built on the lower scientific levels.

5.8.2 Mandatory requirements

1. **Declared intended use.** Inputs, outputs, users, affected people, decisions, and prohibited uses are specified.
2. **Risk analysis.** Foreseeable harms, failure modes, abuse, manipulation, and distribution shifts are evaluated.
3. **Use-case performance thresholds.** Accuracy, calibration, abstention, latency, stability, and worst-group requirements are set from the actual risk.
4. **Human oversight.** Consequential outputs are reviewable, contestable, and correctable.
5. **Consent and privacy.** Psychological and behavioral data have appropriate collection, retention, access, and withdrawal rules.
6. **Fairness and cultural safety.** Subgroup errors and culturally invalid inferences are measured and controlled.

7. **Security and misuse resistance.** The system is tested against adversarial access, data poisoning, model manipulation, and coercive inference.
8. **Monitoring and rollback.** Drift, calibration loss, incident response, and model withdrawal processes are operational.
9. **Governance and accountability.** Responsible institutions and decision rights are explicit.
10. **Independent assurance.** External reviewers verify the evidence and continuing compliance.

5.8.3 Permitted claims

- the named system is operationally validated for the declared use, population, jurisdiction, and monitoring period;
- the system meets prespecified dependability and governance requirements;
- downstream systems may rely on the validated outputs within their tested scope.

5.8.4 Claims not permitted

- AP-5 status transfers automatically to a different use or population;
- the system is infallible;
- operational validation proves consciousness;
- downstream Cultural Impact, reward, price, or governance measures are valid without their own evidence;
- monitoring can stop because an initial release passed.

5.8.5 SQC status

The present SQC checkpoint is a research system. It has not been submitted to a use-case-specific AP-5 assurance process.

Provisional determination: **AP-5 not claimed**.

5.9 Provisional SQC evidence profile

The following profile records present evidence without collapsing it into one score.

Dimension	Present evidence	Provisional status	Next decisive requirement
EPS theoretical specification	Prior causal theory and source manuscripts	Supported	Complete versioned claim and falsification register
Executable SQC architecture	Sequential explicit EPS implementation	Supported	Independent clean reimplementation
Theory-to-code traceability	Strong partial mapping	Provisionally supported	Complete construct-to-code-data-test matrix
Action-pathway integrity	EPS-only contract and freeze audits	Supported for audited paths	Adversarial leakage and runtime identity tests

Continued on next page

Dimension	Present evidence	Provisional status	Next decisive requirement
Causal mediation	Oracle stage interventions and detached pathway	Provisionally supported	Preregistered interventions, negative controls, formal mediation
Internal semantic performance	Explicit supervised stages; state error 29.21%	Partial	Calibrated stage thresholds and pathway-correctness analysis
Synthetic behavioral fidelity	75.25% exact; 97.17% valid-action hit	Supported in current domain	Multi-seed balanced replication and larger cells
Compositional generalization	75.00% over eight held-out triples	Supported for known-factor recombination	New factor levels and out-of-ontology tests
External transfer	BIB-CCS evidence with adaptation tradeoffs	Partial	Additional independent domains and robust multi-seed gains
Human construct validity	No complete independent human battery	Not yet tested	AP-3 study program
Longitudinal psychology	Bounded explicit memory	Early implementation	Autobiographical, relational, developmental validation
Cross-cultural validity	Theoretical grounding, no complete invariance study	Not yet tested	Multi-culture measurement and prospective tests
Comparative uniqueness	No complete matched architecture comparison	Hypothesis	Matched-capacity and ablation program
Operational assurance	Research artifacts only	Not claimed	Use-case-specific AP-5 process

5.10 Provisional classification statement

The current evidence supports the following statement:

> *SQC v1.2 is an implemented artificial-psychology architecture at AP-1, with substantial evidence toward AP-2. It has demonstrated a sequential explicit EPS pathway, structural shortcut prevention, audited action-only parameter isolation, causal stage sensitivity, useful exact and valid-set behavior, and generalization across held-out combinations of known decision-context factors. Full AP-2 passage requires multi-seed replication of the balanced base, expanded causal and leakage audits, pathway-consistency analysis, and independent artifact reproduction. Human construct, ecological, and operational validation remain the subjects of AP-3 through AP-5.*

This statement should replace two misleading alternatives. SQC should not be described as though it has no standing as an artificial psychology until every future human and operational study is complete. It should also not be described as universally validated from its present controlled results. The scale permits the accurate middle position: SQC already qualifies as an implemented

artificial-psychology architecture, has earned substantial internal evidence, and has a defined path toward stronger validation.

5.11 Reclassification, suspension, and regression

Validation status can change when the theory, architecture, evidence, or use changes.

5.11.1 Advancement

A level advances when all mandatory requirements pass under the preregistered protocol and required replications are complete. Partial evidence above the current level is retained in the profile but does not change the highest fully passed level.

5.11.2 Targeted suspension

A classification may be suspended for a construct, population, domain, or use when monitoring identifies drift, calibration failure, artifact non-reproducibility, implementation discrepancy, or material new contradictory evidence. Suspension should be localized where possible.

5.11.3 Regression

A system regresses to the highest level whose mandatory requirements remain satisfied when:

- a foundational audit reveals a prohibited shortcut;
- a claimed construct fails reliable measurement or independent replication;
- evidence was produced through leakage, selection bias, or an incorrect protocol;
- a material architecture change invalidates inherited results;
- operational monitoring shows that AP-5 thresholds are no longer met.

Regression is not a punishment. It is the mechanism that keeps validation claims synchronized with evidence.

5.11.4 Restoration

Suspended or regressed status may be restored through correction, preregistered retesting, and independent confirmation. The public record must preserve the original result, reason for status change, corrective action, and restoration evidence.

5.12 Relationship between the AP scale and the north star

The AP scale measures validation maturity; it does not impose an upper bound on SQC's scientific ambition. AP-5 does not mean complete behavioral fidelity has been achieved. It means a named system is dependable for a declared use under stated uncertainty and monitoring.

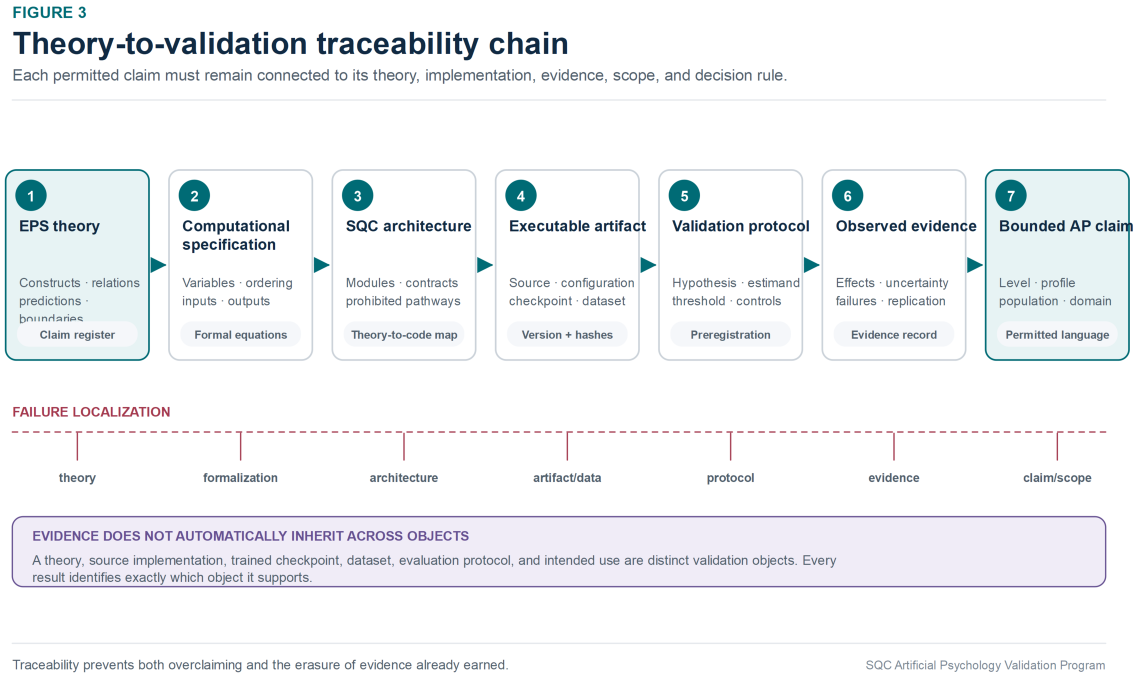

Figure 3: Theory-to-validation traceability chain. Every bounded AP claim is linked to EPS theory, its computational specification, the SQC architecture, versioned executable artifacts, a preregistered protocol, and observed evidence. Failures are localized to the object and link they test.

Complete behavioral fidelity remains a separate asymptotic objective. Progress toward it is measured through increasingly difficult predictions, richer memory, better individualization, broader contexts, calibrated uncertainty, valid-action coverage, causal pathway accuracy, and comparison with the best available alternatives.

It is possible for an AP-4 system to be scientifically broad but not approved for a high-risk use. It is also possible for an AP-5 system to be dependable in a narrow use while remaining far from universal behavioral fidelity. The scale therefore prevents deployment status, scientific generality, and the north-star objective from being mistaken for one another.

6 Validation Principles and Statistical Policy

The validation program is confirmatory where claims are promoted and exploratory where methods or thresholds are still being developed. Its statistical policy is designed to prevent favorable checkpoints, datasets, historical cases, or metrics from being selected after outcomes are known. The policy applies to computational, human, historical, cross-cultural, and operational studies, with domain-specific additions where necessary.

6.1 Preregistration and protocol control

Every confirmatory study will preregister the theory and SQC versions, primary hypotheses, predicted directions, primary and secondary outcomes, sampling frame, exclusion rules, model variants, baselines, covariates, missing-data handling, stopping rule, power analysis, statistical model, multiplicity correction, promotion threshold, and falsification condition. The registration will identify which analyses are confirmatory and which are exploratory.

Any deviation will be time-stamped, justified, and reported beside the original plan. A deviation may be scientifically necessary, but it cannot remain labeled as though it were prespecified. Exploratory results may generate a subsequent locked confirmatory study.

6.2 Study roles and separation of influence

Where resources permit, the program will separate theory authorship, dataset or instrument construction, model development, confirmatory analysis, and final audit. At minimum, final test labels will be inaccessible to model selection, and confirmatory statistical code will be frozen before the locked test is opened. Human annotators will be blinded to SQC output when judging constructs or action validity.

Independent replication is required at the levels specified by the AP scale. Independence will be described rather than assumed: different random seeds are not independent laboratories, and a second analysis by the original developers is not an independent replication.

6.3 Data partitions and information boundaries

Data will be divided by function rather than only by percentage:

- **development data** for ontology, instrument, and implementation work;
- **training data** for parameter estimation;
- **selection data** for model and checkpoint decisions;
- **confirmation data** for preregistered replication;
- **locked final test data** opened only after the analysis is frozen;
- **prospective data** generated after predictions are committed.

Splits will occur at the highest dependence unit relevant to the claim. Records from the same person, episode, counterfactual group, scenario template, institution, historical chain, or relationship will not cross partitions when doing so would leak the outcome structure. Time-dependent studies will train on the past and test on the future.

6.4 Sampling, power, and precision

Sample size will be justified by the primary estimand. For hypothesis tests, the program will target prespecified statistical power under a minimally meaningful effect. For calibration, worst-group, and reliability claims, sample size will be set by interval width and event counts rather than by

significance alone. For model comparisons, simulation-based power analysis will incorporate the number of candidate models and the hierarchical structure of people, scenarios, and time.

Small-cell estimates will always report the numerator, denominator, and uncertainty interval. A worst-cell floor based on 50 observations will be treated as a diagnostic with wide uncertainty, not as a precise population constant. Increasing held-out cell size is a prerequisite for narrow floor claims.

6.5 Estimands and multilevel structure

The primary estimand will be named before selecting a statistic. Depending on the study, it may be a population-average effect, individual prediction error, within-person intervention effect, cross-cultural parameter difference, incremental predictive value, calibration error, or worst-group risk.

Analyses involving repeated observations will model the relevant hierarchy, including participants, raters, scenarios, cultures, institutions, time, and training seeds. Aggregate record-level accuracy will not be treated as though all records were independent. Both population-level and unit-level variation will be reported.

6.6 Uncertainty and effect reporting

All principal results will report counts, point estimates, effect sizes, and confidence or credible intervals. Seed-level and site-level results will be shown rather than only averaged. Where a preservation claim is made, an equivalence or non-inferiority design will be used instead of interpreting a non-significant difference as proof of no change.

For classification proportions, Wilson or justified bootstrap intervals will be used. Continuous errors will report distributional summaries and robust intervals. Model comparisons will include paired effects where predictions are made on the same cases. Bayesian analyses will report priors and sensitivity; frequentist analyses will report the complete decision rule.

6.7 Reliability and measurement quality

No construct can be validly interpreted beyond the reliability of its measurement. Human instruments will report inter-rater, test-retest, internal, and where appropriate composite reliability. Computational states will report stability across repeated observations, perturbations, checkpoints, and time. Parameter recoverability will establish the maximum reliability possible under controlled conditions. The program follows the established principle that validity concerns the evidence supporting an interpretation and use, with construct evidence accumulated through convergent, discriminant, criterion, structural, and intervention tests rather than one correlation (Cronbach and Meehl, 1955; Campbell and Fiske, 1959; Messick, 1995; Trognon et al., 2022).

Reliability thresholds will be appropriate to the intended use and declared in advance. Group research, individual prediction, and consequential allocation require different precision.

6.8 Calibration, ambiguity, and abstention

Accuracy will be evaluated with probability quality. Required measures include log loss, Brier score, calibration slope and intercept, reliability diagrams, and a prespecified calibration-error estimator. Because human situations may permit multiple coherent actions, the program will measure probability assigned to the independently judged valid-action set as well as exact canonical action.

SQC must be able to express insufficient information. Selective-risk and coverage curves will measure whether abstention or a request for additional context reduces error. Overconfidence on ambiguous or out-of-ontology cases will count against validation even when average top-one accuracy is high.

6.9 Missing data and observation quality

Missing psychological information is part of the scientific problem, not only a data-cleaning inconvenience. Studies will distinguish unavailable, unobserved, withheld, inapplicable, and disputed information. Missingness mechanisms will be analyzed, and sensitivity studies will test whether conclusions depend on imputation assumptions.

Performance will be reported as observation completeness changes. This directly tests whether SQC uncertainty increases appropriately when history, values, relationships, or context are unknown.

6.10 Multiple comparisons and researcher degrees of freedom

Each study will identify one or a small number of primary outcomes. Families of secondary hypotheses will use a prespecified family-wise or false-discovery control. Uncorrected exploratory results will be labeled clearly. Searching across seeds, epochs, subgroups, scenarios, prompts, historical periods, or metrics and reporting only the strongest result is prohibited.

Checkpoint promotion will be based on a declared multi-objective rule, not on the most favorable metric after inspection. The rule must include any causal, semantic, compositional, calibration, and preservation guardrails relevant to the claim.

6.11 Baselines and matched comparisons

Every comparative claim will use baselines appropriate to the question. The baseline suite includes simple frequency and heuristic predictors, behavior-only models, raw observation-to-action models, matched-capacity black-box models, recurrent and transformer predictors, ablated SQC variants, reordered EPS variants, and relevant cognitive or affective architectures where faithful implementation is possible.

Comparisons will match access to observations, history, training cases, parameter or compute budgets, optimization effort, and selection opportunities. If matching is impossible, the imbalance will be reported and the claim narrowed. SQC's distinctive value will be evaluated jointly through accuracy, calibration, data efficiency, transfer, causal fidelity, and interpretability.

6.12 No single-number validation

No aggregate score can establish artificial psychology. A high behavioral score can coexist with incorrect internal states. Strong construct correlations can coexist with a bypassed action pathway. Good mean performance can conceal catastrophic subgroup failure. A composite dashboard may summarize evidence, but passage requires the mandatory profile of the relevant AP level.

6.13 Open artifacts and reproducibility

Every confirmatory release will identify code revision, environment, data and checkpoint hashes, ontology version, configuration, random seeds, exact commands, and generated reports. Where human privacy prevents open data, the program will publish metadata, instruments, code, synthetic test fixtures, and a governed replication path. Test counts will always be tied to a revision and command.

7 Study Program I - Architectural and Implementation Fidelity

This program determines whether SQC faithfully realizes the declared EPS architecture. Its studies are prerequisites for full AP-2 passage because psychological results are uninterpretable if the executable system differs from the specification or uses prohibited information.

7.1 Study 1 - Theory-to-code traceability audit

Objective. Verify that every material EPS construct and causal relation has an explicit status in SQC.

Method. An audit team will construct a versioned matrix linking each EPS source proposition to its formal definition, computational module, tensor or state representation, input parents, downstream children, data field, label generation rule, owned loss, unit test, intervention, current evidence, and planned human measure. Unimplemented theory elements and engineering-only features will be recorded explicitly.

Primary outcomes. Coverage of mandatory constructs and edges; number of unresolved mappings; number of code paths without theoretical status; and agreement between theory and implementation reviewers.

Pass principle. All constructs required for the claimed SQC version must have an auditable mapping, and no engineering feature may be presented as an EPS claim without theoretical support. Bounded omissions are permitted only when the scope is narrowed and the omission is visible.

Output. A public traceability matrix and discrepancy register.

7.2 Study 2 - Action-input and shortcut audit

Objective. Determine whether the behavioral readout can use information outside the declared EPS action state.

Method. Static graph inspection, runtime hooks, input-signature tests, feature permutation, label leakage scans, mutual-information probes, adversarial predictors, and controlled removal of EPS states will test raw observation, encoder hidden state, decision-context one-hots, target labels, sample IDs, generator artifacts, and future outcomes. The test will cover training, validation, inference, adaptation, checkpoint loading, and exported runtimes.

Primary outcomes. Prohibited-path count; predictive information available through prohibited channels; action change after removal of declared inputs; and bypass performance under EPS-state corruption.

Pass principle. No prohibited path may provide material action information. Behavior must degrade or change as preregistered when the causally necessary EPS state is removed or intervened on.

7.3 Study 3 - Gradient and parameter-isolation audit

Objective. Verify module ownership and all claims that action-only training leaves EPS unchanged.

Method. Record trainable flags, gradients, optimizer parameter groups, pre/post tensor hashes, maximum absolute parameter differences, and matched-input EPS outputs for each stagewise and adaptation mode. Repeat across seeds, hardware, precision modes, checkpoint save/load, and resumed training.

Primary outcomes. Unexpected nonzero gradients; unauthorized parameter changes; output differences under matched inputs; and reproducibility of the audit.

Pass principle. No unauthorized update may occur. Claims will distinguish parameter immutability, gradient isolation, and runtime output identity.

7.4 Study 4 - Reimplementation and reproducibility study

Objective. Establish that SQC results follow from the declared artifacts rather than an undocumented local environment.

Method. Reproduce the balanced pipeline in a clean environment from a tagged release. A second implementation team will rebuild the core sequential pathway from the formal specification without copying model code where feasible. Training will use prespecified seeds and a locked comparison protocol.

Primary outcomes. Dataset identity, stage outputs, final behavior, intervention effects, runtime differences, and between-implementation variance.

Pass principle. Results must fall within preregistered equivalence margins for the claims being reproduced. Material discrepancies trigger implementation localization before scientific interpretation.

8 Study Program II - Internal Causal and Semantic Validation

This program tests whether explicit SQC states have coherent roles inside the architecture. It does not yet establish human construct validity; it determines whether the implemented mechanism behaves according to its own EPS specification.

8.1 Study 5 - Stagewise causal-intervention experiment

Objective. Estimate the causal contribution of context, values, existential appraisal, drivers, emotion, existential state, and memory to downstream state and action.

Method. For each stage, replace predictions with oracle targets, controlled perturbations, counterfactual values, and matched sham interventions. Predicted directions will be preregistered at the construct and scenario level. The design will estimate total, direct, and mediated effects and test interactions among drivers and existential implications.

Primary outcomes. Change in downstream states, action distribution, valid-action probability, calibration, and pathway correctness.

Pass principle. Relevant interventions must produce reliable changes in the predicted direction, while sham and irrelevant interventions remain within equivalence bounds.

8.2 Study 6 - Counterfactual invariance and sensitivity

Objective. Test whether SQC changes when EPS says it should and remains stable when EPS says it should not.

Method. Construct matched counterfactual groups that vary one target factor while holding nuisance information constant. Complement them with nuisance perturbations in wording, ordering, identifiers, and irrelevant metadata. Include graded interventions to test monotonicity and reversibility.

Primary outcomes. Target sensitivity, nuisance invariance, dose-response shape, directional agreement, and counterfactual consistency.

Pass principle. Prespecified target effects must exceed nuisance effects and maintain the predicted sign across confirmation sets.

8.3 Study 7 - Parameter and state recoverability

Objective. Determine whether distinct EPS states can be identified from the available evidence and whether training recovers known generative parameters.

Method. Generate controlled trajectories from known parameters across noise, missingness, class imbalance, and state overlap. Fit SQC and estimate recovery bias, variance, coverage, and confusion. Use simulation-based calibration where probabilistic parameters are introduced.

Primary outcomes. State recoverability, parameter bias, interval coverage, identifiability failures, and the conditions producing observational equivalence.

Pass principle. Required constructs must be recoverable at the precision needed for the claimed use. Non-identifiable states must be merged, measured differently, or reported with uncertainty rather than forced into certainty.

8.4 Study 8 - Error-cancellation and pathway-consistency audit

Objective. Separate correct action from correct psychological reasoning.

Method. Classify each case into jointly correct, action-correct/pathway-wrong, pathway-correct/action-noncanonical, jointly wrong, and valid-alternative categories. Use oracle substitutions and local interventions to identify where errors cancel. Independent raters will judge pathway coherence in a subset.

Primary outcomes. Joint pathway-action accuracy, cancellation rate, valid-alternative rate, stage-specific compensation, and action confidence under pathway error.

Pass principle. Promotion cannot rely on action gains produced by increased pathway error or overconfidence. Both final behavior and the claimed mediation must improve or remain within declared guardrails.

9 Study Program III - Psychometric and Construct Validation

This program supplies the central AP-3 evidence. It asks whether the SQC states correspond to the human psychological constructs attributed to them, rather than only matching labels produced by the EPS trajectory generator. It also adopts the computational-modeling distinction among empirical, theoretical, and generative evidence and the associated requirements for recoverability, external validation, calibration, and transparent reporting (Zavlis et al., 2025).

9.1 Study 9 - Content-validity study

Objective. Determine whether the EPS/SQC ontology adequately represents the declared domain of psychological organization.

Method. Independent panels spanning psychology, neuroscience, anthropology, sociology, history, and relevant cultural expertise will rate construct definitions, relevance, coverage, overlap, and missing content. Community panels will evaluate culturally situated interpretations. Reviewers will disclose theoretical orientation and conflicts.

Primary outcomes. Item- and scale-level content-validity indices, agreement, construct overlap, missing-domain nominations, and qualitative dissent.

Pass principle. Mandatory constructs must meet prespecified relevance and coverage standards across appropriately constituted panels. Persistent missing content requires ontology revision or a narrower scope.

9.2 Study 10 - Human annotation and inter-rater reliability

Objective. Create reliable independent measures of context, values, existential appraisal, drivers, emotion, existential state, action validity, and outcome.

Method. Develop structured instruments with definitions, behavioral anchors, examples, uncertainty fields, and evidence requirements. Train raters on development cases, then evaluate locked cases while blinded to SQC outputs. Include participant self-report, observer judgment, and domain-expert judgment where each is appropriate.

Primary outcomes. Krippendorff's alpha, kappa, intraclass correlation, test-retest reliability, adjudication rate, missingness, and ambiguity distribution.

Pass principle. Reliability must meet the preregistered requirement for each construct and use. Constructs that cannot be measured reliably cannot support strong criterion claims until the instrument or ontology is revised.

9.3 Study 11 - Convergent validity

Objective. Test whether SQC states relate to independent measures of the same or theoretically related constructs.

Method. Select established behavioral tasks, validated scales, observer ratings, and physiological measures without using them to train the tested SQC checkpoint. Preregister the sign, ordering, and plausible magnitude of each relation. Use multitrait-multimethod models to separate method variance.

Primary outcomes. Convergent correlations, latent associations, prediction of independent measures, and consistency across methods and replications.

Pass principle. Relations must follow the predicted pattern and exceed relevant discriminant relations, with uncertainty compatible with the prespecified minimally meaningful association. Recent work applying full reliability, convergent, discriminant, and criterion batteries to synthetic traits in artificial systems provides a methodological precedent while not establishing architectural mediation by itself (Serapio-Garcia et al., 2025).

9.4 Study 12 - Discriminant validity

Objective. Demonstrate that adjacent SQC constructs are not interchangeable.

Method. Compare each construct with theoretically distinct measures and design interventions expected to change one construct more than another. Test Belonging versus Identity, Identity versus Power, existential appraisal versus existential state, and emotion versus driver activation explicitly.

Primary outcomes. Cross-construct correlations, latent-factor separation, intervention selectivity, classification confusion, and shared-method effects.

Pass principle. Each construct must demonstrate reliable unique variance or unique causal behavior. Constructs that remain empirically inseparable require redefinition, measurement improvement, or

justified collapse.

9.5 Study 13 - Structural validity

Objective. Test the organization and ordering of EPS constructs in human data.

Method. Compare the declared EPS structure with collapsed, reordered, parallel, and alternative models using confirmatory factor analysis, structural equation modeling, causal discovery constrained by temporal information, and held-out predictive checks. The confirmatory model set will be frozen before test evaluation.

Primary outcomes. Model fit, out-of-sample prediction, parameter stability, directional intervention agreement, and model-selection uncertainty.

Pass principle. The EPS structure must meet absolute adequacy requirements and demonstrate a prespecified advantage or explanatory distinction over credible alternatives. One favorable fit index is insufficient.

9.6 Study 14 - Criterion and incremental validity

Objective. Determine whether SQC states predict independently observed behavior and outcomes beyond simpler information.

Method. Use prospective or strictly held-out criteria. Compare models using observation alone, history alone, demographic or contextual variables, direct behavior prediction, and explicit EPS states. Evaluate both statistical increment and decision-relevant benefit.

Primary outcomes. Incremental explained variance, likelihood improvement, proper scoring rules, calibration, net benefit where appropriate, and changes in worst-group error.

Pass principle. EPS states must add reliable out-of-sample information and must not gain average performance through unacceptable subgroup harm or miscalibration.

10 Study Program IV - Behavioral Fidelity and Calibration

This program measures how closely SQC predicts behavior while respecting ambiguity, uncertainty, pathway correctness, and distribution change. It expands the current compositional benchmark into increasingly difficult controlled and human tests.

10.1 Study 15 - Expanded theory-grounded synthetic benchmark

Objective. Increase the precision, diversity, and adversarial strength of the controlled SQC evaluation.

Method. Expand records per cell, scenario families, action classes, values, roles, goals, risks, outcome mechanisms, and temporal depth. Build at least one independent generator from the

formal EPS specification. Reserve locked families and generator versions that model developers cannot inspect.

Primary outcomes. Exact action, valid-set probability, family accuracy, rank metrics, stage accuracy, pathway-action joint accuracy, calibration, worst-cell performance, and cross-generator transfer.

Pass principle. The balanced architecture must repeat across seeds and generators without violating mediation, calibration, or worst-cell guardrails.

10.2 Study 16 - Out-of-ontology generalization

Objective. Test representation beyond recombination of known factor levels.

Method. Withhold entire roles, goals, risk regimes, values, actions, social structures, and scenario-generating mechanisms. Evaluate zero-shot detection, abstention, representation extension, few-shot adaptation, and catastrophic forgetting. No held-out records may enter adaptation unless the protocol defines a separate post-test learning phase.

Primary outcomes. Novelty detection, zero-shot performance, calibrated abstention, adaptation efficiency, retained old-domain performance, and need for ontology expansion.

Pass principle. SQC must either generalize successfully or identify that the case lies outside its ontology. Confident forced classification of genuinely new constructs counts as failure.

10.3 Study 17 - Behavioral ambiguity and valid alternatives

Objective. Replace the assumption that every psychological state has one correct action with an empirically measured distribution of coherent responses.

Method. Independent raters and participants will evaluate candidate actions, generate missing alternatives, state conditions under which each is valid, and assign confidence. Latent-variable or consensus models will estimate a valid-action distribution while preserving disagreement.

Primary outcomes. Probability mass on the valid set, rank of observed human action, distributional divergence, exact action, outcome quality, and uncertainty agreement.

Pass principle. SQC must assign calibrated probability to the range of psychologically defensible actions and avoid treating rater disagreement as a single certain label.

10.4 Study 18 - Probability calibration and abstention

Objective. Determine whether SQC confidence corresponds to empirical correctness and whether it can recognize insufficient information.

Method. Evaluate uncalibrated and prespecified calibration methods on separate data. Manipulate observation completeness, ambiguity, ontology novelty, and distribution shift. Allow the system to abstain or request defined missing information.

Primary outcomes. Brier score, log loss, calibration slope and intercept, reliability diagrams, selective risk, coverage, information-request utility, and subgroup calibration.

Pass principle. Calibration must remain within domain-specific tolerances, and abstention must reduce risk monotonically without excluding protected or difficult groups disproportionately.

10.5 Study 19 - Individual-level prospective prediction

Objective. Test whether SQC predicts future behavior for held-out people using only information available before action.

Method. Collect longitudinal context, value, relationship, state, and action data under consent. Separate population training from individual adaptation. Freeze predictions before outcomes and hold out participants, situations, and time periods. Compare SQC with participant self-prediction, close-other judgment, domain experts, base rates, and matched computational models.

Primary outcomes. Individual and population accuracy, valid-set probability, calibration, log score, within-person change sensitivity, between-person discrimination, abstention, and subgroup performance.

Pass principle. SQC must add prospective value under the matched comparison and maintain calibrated uncertainty. Performance will be reported as a function of information completeness and history length.

11 Study Program V - Memory, Development, and Longitudinal Psychology

EPS describes psychology as historically constructed. A system limited to independent snapshots cannot fully test socialization, identity continuity, relationship development, trust, learning, or the delayed effects of action. This program expands the current bounded memory vector into an auditable longitudinal architecture while preserving EPS mediation.

11.1 Study 20 - Autobiographical memory validation

Objective. Determine whether SQC can encode and use personally situated experience accurately, selectively, and causally.

Method. Implement versioned episodic, semantic-self, relationship, action, and outcome memory with explicit provenance and timestamps. Controlled trajectories will manipulate salience, repetition, delay, interference, contradiction, and retrieval cues. Memory access will enter contextual interpretation or the explicit EPS action state only through declared interfaces.

Primary outcomes. Recall and source accuracy, retrieval relevance, temporal ordering, interference, false-memory rate, memory-dependent behavior, and uncertainty when recall is incomplete.

Pass principle. Memory must improve longitudinal prediction and pathway coherence without

creating a hidden direct action lookup. False certainty about unreliable or conflicting memories counts against validation.

11.2 Study 21 - Relational and socialization trajectories

Objective. Test whether repeated interaction produces coherent changes in belonging, identity, power, expectation, emotion, and behavior.

Method. Use multi-episode trajectories involving trust, reciprocity, betrayal, exclusion, protection, repair, obligation, status change, coalition, and reconciliation. Manipulate the order and attribution of events while holding totals constant. Compare SQC trajectories with participant judgments and longitudinal observations.

Primary outcomes. Trust and belonging updates, relationship-specific behavior, recovery after repair, sensitivity to actor identity, order effects, and retention over delay.

Pass principle. State changes must reflect who did what, under which relationship and history, rather than only aggregate positive or negative event counts.

11.3 Study 22 - Developmental validation

Objective. Evaluate EPS predictions about how repeated social experience stabilizes or changes psychological organization over development.

Method. Define developmental hypotheses at age-appropriate levels and use ethical longitudinal observation, existing developmental datasets, and controlled simulations. Separate maturational, socialization, cultural, and individual-history effects. Predictions will concern trajectories and response patterns, not unsupported diagnostic classification.

Primary outcomes. Developmental ordering, stability and plasticity of values and drivers, adaptation to repeated outcomes, age-appropriate calibration, and agreement with held-out longitudinal evidence.

Pass principle. EPS must predict more than retrospective developmental narratives. Prespecified trajectories must replicate, and deviations must inform parameterization or theory revision.

11.4 Study 23 - Long-horizon identity and self-consistency

Objective. Determine whether SQC maintains a coherent but adaptable model of identity across changing contexts and delayed consequences.

Method. Construct long sequences containing role changes, value conflict, commitments, temptation, sacrifice, reputation, delayed reward, and future optionality. Test continuity after context switches and memory compression. Compare coherent development with arbitrary drift and rigid persistence.

Primary outcomes. Identity consistency, justified change, commitment retention, long-delay prediction, contradiction detection, and calibration of self-model uncertainty.

Pass principle. SQC must preserve relevant continuity while updating when evidence warrants change. Neither amnesia nor unchangeable identity qualifies as successful longitudinal psychology.

12 Study Program VI - Cross-Cultural and Civilizational Validation

This program tests whether EPS relations recur across cultures and history, which aspects require local parameterization, and where the present ontology is incomplete. It converts the civilizational evidence base described in Section 3 into preregistered studies.

12.1 Study 24 - Cross-cultural construct equivalence

Objective. Establish whether EPS/SQC constructs have comparable meanings across languages and cultural settings.

Method. Partner with local investigators and communities to translate, adapt, and challenge construct definitions and instruments. Use qualitative elicitation before confirmatory measurement. Test configural, metric, and where appropriate scalar invariance, while allowing partial invariance and culturally specific extensions.

Primary outcomes. Construct coverage, translation equivalence, reliability, factor structure, parameter differences, differential item functioning, and local ontology gaps.

Pass principle. Universal claims require invariant core relations. Where invariance fails, the model must use transparent cultural parameterization or narrow its claim rather than treating one population as the default.

12.2 Study 25 - Historical blind-prediction study

Objective. Test whether EPS/SQC predicts historical decisions using only information available before the outcome.

Method. Select cases through a preregistered sampling frame. An independent historical team will prepare pre-outcome evidence packets and sealed outcomes. SQC, historians, domain experts, statistical models, and alternative theoretical models will submit frozen predictions before outcomes are opened.

Primary outcomes. Action and outcome prediction, calibration, explanation coding, valid-alternative coverage, and incremental performance over baselines.

Pass principle. Confirmatory claims require replicated prospective-style performance across periods and case types, not persuasive post hoc explanation of famous events.

12.3 Study 26 - Civilizational recurrence analysis

Objective. Quantify whether EPS relations recur across large-scale patterns of scarcity, conflict, cooperation, identity, power, institutional change, and cultural stabilization.

Method. Build a versioned corpus with explicit units of analysis, provenance, temporality, and dependence structure. Independent coders will annotate pre-action information and outcomes. Hierarchical models will estimate shared relations and culturally or institutionally varying parameters.

Primary outcomes. Directional recurrence, cross-context heterogeneity, predictive increment, coder reliability, temporal stability, and unresolved cases.

Pass principle. The same core relations must recur beyond selected examples, and heterogeneity must be explained or bounded. Civilizational breadth cannot be claimed from a collection of illustrative cases alone.

12.4 Study 27 - Adversarial and boundary cases

Objective. Find cases that SQC explains poorly, predicts incorrectly, or cannot represent.

Method. Invite historians, psychologists, anthropologists, affected communities, and competing theorists to nominate difficult cases. Include strategic deception, atypical development, mixed identities, conflicting group membership, rapid norm change, institutional coercion, collective action, and novel social forms.

Primary outcomes. Confident error, ontology failure, competing-explanation advantage, subgroup failure, and revision required.

Pass principle. Boundary discovery is itself a required output. The program fails scientific transparency if difficult cases are excluded or retrospectively redefined until every outcome appears compatible.

13 Study Program VII - External Domains and Comparative Evaluation

The uniqueness hypothesis requires evidence against serious alternatives under matched conditions. This program evaluates SQC beyond its development domain and determines whether explicit EPS mediation provides incremental value.

13.1 Study 28 - Multi-domain external benchmark suite

Objective. Measure transfer across psychologically distinct behavioral domains.

Method. Independent task teams will construct locked benchmarks for scarcity coordination, negotiation, care, conflict, institutional behavior, cultural participation, learning, planning, and

relationship repair. BIB-CCS remains one component rather than the sole external measure. Each domain will document its constructs, action ambiguity, horizon, and information boundary.

Primary outcomes. State inference, policy adaptation, value consistency, social outcome, exact and valid-set behavior, calibration, transfer cost, and catastrophic forgetting.

Pass principle. Generality requires replicated performance across domains without sacrificing the EPS pathway or previously validated behavior.

13.2 Study 29 - Matched-capacity architecture comparison

Objective. Test whether SQC outperforms architectures that do not reconstruct the explicit EPS pathway.

Method. Compare SQC with behavior-only, direct observation-to-action, recurrent, transformer, cognitive-architecture, appraisal-based, and hybrid alternatives. Match data, observation and history access, parameter or compute budget, optimization trials, checkpoint-selection opportunities, and evaluation seeds. Publish all baseline implementation decisions.

Primary outcomes. Accuracy, valid-set probability, calibration, data efficiency, compositional and out-of-ontology transfer, prospective prediction, adaptation, forgetting, causal explanation, and compute cost.

Pass principle. Claims of comparative superiority require a preregistered multidimensional advantage or a declared tradeoff profile, not a win on one selected metric. Claims that SQC is uniquely capable require repeated failure of credible alternatives under fair conditions and remain open to future systems.

13.3 Study 30 - Architectural ablation and necessity tests

Objective. Identify which EPS stages and relations are necessary for SQC's advantages.

Method. Remove, bypass, collapse, randomize, or reorder context, values, existential appraisal, drivers, emotion, state, and memory. Match parameter count where possible. Include partial models and direct shortcuts as diagnostic controls, not promotion candidates.

Primary outcomes. Change in behavior, calibration, pathway correctness, transfer, data efficiency, robustness, and construct validity.

Pass principle. A stage claimed as necessary must provide reproducible incremental value or a unique causal function. Redundant stages require theoretical explanation, improved measurement, or architectural revision.

13.4 Study 31 - Independent replication challenge

Objective. Subject the architecture and principal claims to external construction, replication, and adversarial comparison.

Method. Release the formal specification, development artifacts, frozen analysis plan, and governed access to locked tests. External teams may reproduce SQC, implement alternatives, identify leakage, or submit falsification tests. The challenge will predefine faithful replication, partial replication, non-replication, and invalid comparison.

Primary outcomes. Reproduction rate, effect heterogeneity, discovered defects, alternative-model performance, and unresolved discrepancies.

Pass principle. Higher AP claims require independent confirmation. Failed replications must be investigated publicly and cannot be dismissed solely because they were performed outside the originating team.

14 Study Program VIII - Neural and Physiological Correspondence

SQC is a functional computational architecture, not a biological duplicate. Nevertheless, EPS makes neurocomputational and affective claims that can be constrained by neural and physiological evidence. This program tests those claims without treating superficial brain resemblance as proof of psychology.

14.1 Study 32 - Neural correspondence

Objective. Test whether the ordering and modulation predicted by EPS correspond to human neural dynamics at the appropriate level of abstraction.

Method. Select tasks with preregistered contextual, value, appraisal, motivational, emotional, state, and action contrasts. Acquire or reuse suitable neuroimaging and electrophysiological data. Compare temporal order, representational geometry, connectivity, and perturbation predictions rather than assigning one SQC tensor to one brain region.

Primary outcomes. Temporal correspondence, representational similarity, directional connectivity, cross-participant replication, and incremental explanation over alternative models.

Pass principle. Claimed neural mappings require replicated, specific relations. Failure localizes to the mapping or theory claim and does not by itself negate functional AP classification.

14.2 Study 33 - Physiological and affective dynamics

Objective. Validate emotional modulation against independent behavioral, self-report, and physiological evidence.

Method. Collect synchronized SQC predictions, experience reports, action, heart-rate variation, electrodermal activity, facial or vocal behavior, and other justified measures during controlled and naturalistic tasks. Use multimethod latent models because no single channel is ground truth.

Primary outcomes. Convergent and discriminant patterns, timing, calibration, individual differences, and prediction of learning or action change.

Pass principle. Emotional dimensions must show theoretically specific multimethod relations and downstream effects, not merely correlate with generic arousal or task difficulty.

14.3 Study 34 - Lesion and perturbation correspondence

Objective. Compare impairment patterns produced by artificial stage lesions with relevant human perturbation evidence.

Method. Simulate bounded lesions, noise, delayed communication, and gain changes in SQC. Preregister predicted impairment profiles and compare them with ethical experimental perturbations, natural lesions, pharmacological evidence, or established clinical findings where scientifically appropriate.

Primary outcomes. Selective impairment, double dissociation, compensatory pathways, recovery, and alternative-model fit.

Pass principle. Correspondence must be selective and predictive. General performance degradation after any lesion is not evidence for the EPS mapping.

15 Study Program IX - Operational Validation and the Human-Centered Systems Gate

Operational validation begins only after the relevant scientific claims are sufficiently supported for the intended domain. It determines whether a named system can be relied upon in practice and whether downstream significance measurement can support Neural Grid and HCS functions.

15.1 Study 35 - Use-case-specific reliability

Objective. Define and test dependability for each proposed deployment.

Method. For a named use, specify inputs, outputs, affected people, decision authority, harm severity, acceptable error, calibration, abstention, latency, availability, and recovery. Conduct prospective shadow deployment before live influence. Stress-test distribution shifts, missing information, and system failure.

Primary outcomes. Use-case accuracy, calibrated risk, abstention, worst-case and worst-group performance, uptime, drift, incident rate, and human correction.

Pass principle. All mandatory thresholds must pass simultaneously during a prespecified monitoring period. Evidence from one use does not transfer without an impact analysis.

15.2 Study 36 - Safety, fairness, and subgroup performance

Objective. Identify and control harms created by psychological inference or behavioral prediction.

Method. Evaluate cultural misclassification, stereotyping, manipulation, coercion, privacy invasion, feedback loops, differential abstention, and allocation harm. Include participatory review by affected groups and adversarial testing by independent specialists.

Primary outcomes. Subgroup error and calibration, harm-weighted risk, privacy leakage, manipulation susceptibility, disparate abstention, and failure recovery.

Pass principle. No deployment may hide unacceptable subgroup harm behind mean performance. Prohibited uses and escalation rules are mandatory outputs.

15.3 Study 37 - Human oversight and contestability

Objective. Ensure consequential SQC outputs can be understood, challenged, corrected, and overridden.

Method. Test whether explanations accurately reflect the causal pathway, whether reviewers can identify errors, whether affected people can supply missing context, and whether corrections propagate safely. Measure automation bias and reviewer workload.

Primary outcomes. Explanation fidelity, error-detection rate, correction success, appeal resolution, override behavior, and time burden.

Pass principle. Oversight must change outcomes when SQC is wrong and must not serve as ceremonial approval of an opaque decision.

15.4 Study 38 - Downstream dependency validation

Objective. Determine when SQC significance computation can support validated Neural Grid and Human-Centered Systems functions.

Method. Trace each downstream measure to the SQC constructs and uncertainty on which it depends. Validate significance before using it in Cultural Impact, Cultural Yield, contribution, reward, pricing, settlement, or governance. Conduct end-to-end pilots with shadow accounting, counterfactual audits, stakeholder review, and rollback.

The dependency chain is:

```
validated artificial psychology in SQC
    -> validated significance computation
    -> Neural Grid infrastructure
    -> Human-Centered Systems
    -> economic and commercial activation
```

Primary outcomes. Measurement validity, uncertainty propagation, ledger and metering accuracy, incentive effects, fairness, contestability, governance performance, and unintended cultural or economic consequences.

Pass principle. A downstream claim cannot be more validated than the SQC constructs and significance measures on which it materially depends. Research, specification, and prototyping may proceed before passage; validated operational dependability may not be claimed.

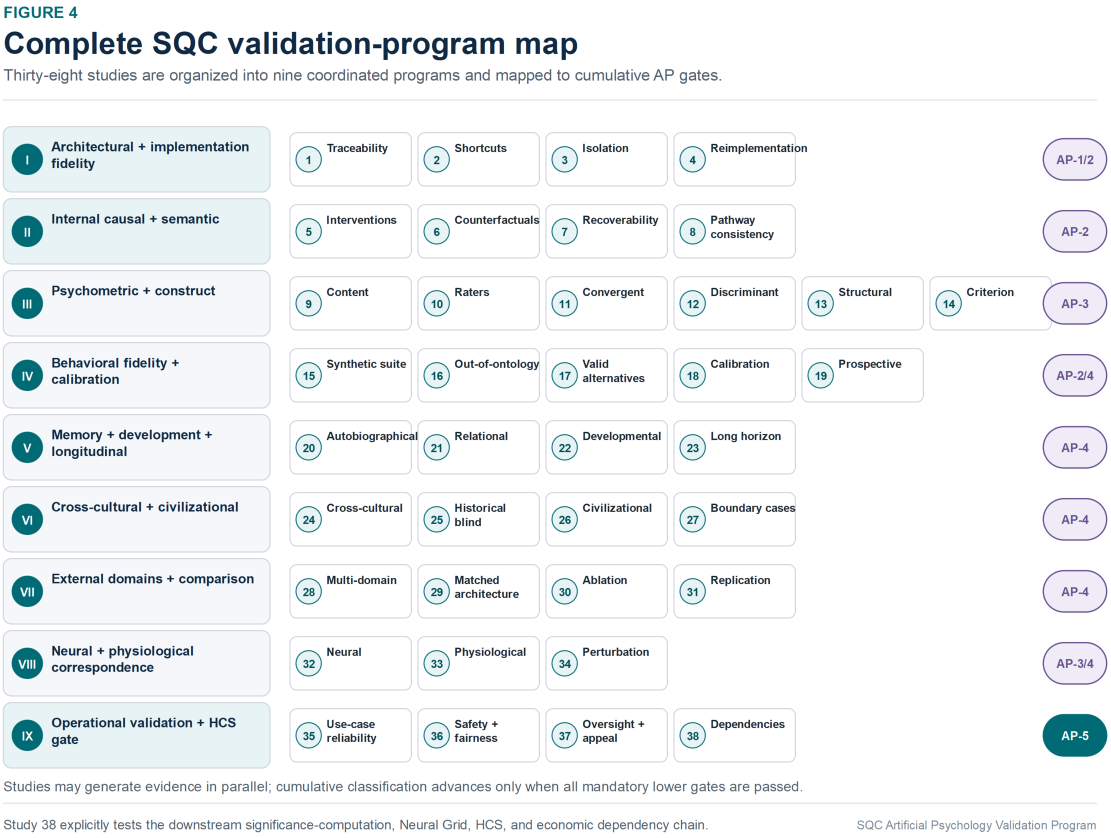

Figure 4: Complete SQC validation-program map. The 38 studies are organized into nine programs spanning architectural fidelity, causal mediation, psychological constructs, behavioral calibration, longitudinal and cultural evidence, external comparison, neural and physiological correspondence, and operational validation.

16 Integrated Metrics and Evidence Dashboard

The evidence dashboard records the full validation profile without reducing it to one score. Every metric is linked to a claim, validation unit, population, domain, protocol, and uncertainty estimate. The dashboard will be published in machine-readable and human-readable forms for every confirmatory release.

16.1 Architectural fidelity

Metric	Interpretation
Theory-to-code coverage	Proportion of required EPS claims with complete formal, code, data, test, and study mappings
Prohibited-path count	Number of executable paths violating the declared action contract
Unauthorized gradient count	Parameters receiving gradients outside their owned phase
Unauthorized tensor-change count	Non-owned checkpoint tensors changed by a training or adaptation phase
Matched-input output identity	Numerical agreement of frozen stages under controlled pre/post evaluation
Reproduction agreement	Equivalence of independent runs or implementations on declared outputs

Architectural metrics are gates. Strong behavioral performance cannot compensate for a prohibited shortcut when EPS mediation is the claim.

16.2 Internal causal and semantic performance

The dashboard will report stage accuracy or error, proper scores for continuous states, calibration, parameter recovery, intervention effects, negative-control effects, counterfactual sensitivity, nuisance invariance, pathway-action joint accuracy, and error-cancellation rate. Directional intervention agreement will be shown by construct and scenario family rather than only in aggregate.

16.3 Psychological construct evidence

Evidence class	Required measures
Content	Expert and community relevance, coverage, overlap, and missing-content indices
Reliability	Inter-rater, test-retest, internal, composite, and parameter-recovery reliability
Convergent	Associations with independent related measures and methods
Discriminant	Separation from adjacent and unrelated constructs
Structural	Fit and held-out performance of EPS organization against alternatives

Continued on next page

Evidence class	Required measures
Criterion	Prediction of independent behavior and outcomes
Incremental	Improvement beyond observation, history, demographics, and simpler models
Intervention	Predicted within-person or within-system change under manipulation

16.4 Behavioral fidelity and uncertainty

Required measures include exact action accuracy, family accuracy, top-k rank, probability on the observed action, probability on the valid-action set, distributional divergence from human actions, outcome quality, Brier score, log loss, calibration slope and intercept, selective risk, coverage, and information request value. Counts and intervals accompany every proportion.

Performance is stratified by person, scenario, context, ontology familiarity, culture, language, time horizon, ambiguity, and observation completeness. Worst-group and worst-cell estimates include uncertainty.

16.5 Generalization and adaptation

The dashboard distinguishes:

- familiar records from familiar cells;
- new records in familiar combinations;
- unseen combinations of known factor levels;
- unseen factor levels;
- unseen ontology elements;
- new scenario generators;
- new domains and cultures;
- future time periods and held-out people.

Adaptation metrics include pre/post performance, retained old-domain performance, forgetting, transfer efficiency, number of examples, compute, calibration change, EPS immutability where claimed, and seed variation.

16.6 Comparative evidence

Comparisons report matched conditions and the complete performance vector. Incremental validity, data efficiency, calibration, transfer, robustness, causal fidelity, interpretability, latency, energy, and parameter count are included. A claim of SQC advantage identifies the dimensions and tradeoffs rather than stating undifferentiated superiority.

16.7 Operational evidence

Operational dashboards add use-case error cost, harm-weighted risk, subgroup calibration, abstention, privacy and security incidents, reviewer correction, appeal outcomes, drift, uptime, rollback, and unresolved complaints. These metrics are monitored continuously and can suspend AP-5 status.

16.8 Evidence states and dashboard schema

Every dashboard cell has one of five states:

- **passed** under a preregistered criterion;
- **provisional** pending required precision or replication;
- **inconclusive** because the study lacked discriminatory evidence;
- **failed** under the declared criterion;
- **not yet tested**.

Each cell records the claim ID, AP requirement, estimate, interval, sample, model and data version, protocol link, artifact link, date, limitations, and next decisive test. The headline AP level is generated from mandatory passed gates; it is not calculated by averaging cells.

17 Promotion, Revision, and Falsification Rules

17.1 Checkpoint and AP-level promotion

Promotion decisions are made from locked confirmation evidence. A candidate must satisfy every mandatory guardrail relevant to the claim, including causal integrity, internal-state performance, behavior, calibration, generalization, worst-group behavior, and preservation. No candidate is promoted because it is best among weak alternatives if it fails an absolute safety or validity gate.

Single favorable seeds, epochs, subgroups, or benchmark variants cannot support promotion. When multiple candidates are searched, selection uncertainty is acknowledged and a fresh confirmation set is required.

17.2 Theory and ontology revision

Revisions will distinguish:

- parameter or calibration changes within the same theory;
- measurement-instrument changes;
- data and sampling corrections;
- implementation corrections;
- architectural changes;
- ontology expansion or collapse;
- changes to an EPS causal proposition.

Material changes receive a new version and an inheritance analysis. The program will not rename a failed construct after observing results or add post hoc paths that make every outcome compatible. Exploratory revisions must face new locked tests.

17.3 Falsification conditions

The following outcomes count against the named SQC or EPS claim when they repeat under adequate measurement, power, and implementation fidelity:

1. A simpler non-EPS architecture matches or exceeds SQC across behavioral fidelity, calibration, transfer, and data efficiency without sacrificing explanatory requirements.
2. Removing, bypassing, or reordering a supposedly necessary EPS stage produces no meaningful loss or produces a reliable improvement.
3. SQC constructs fail convergent or discriminant validation and cannot be recovered reliably.
4. Preregistered interventions repeatedly change downstream states or behavior in the wrong direction.
5. The EPS structural model is consistently inferior to credible alternatives on held-out human evidence.
6. Core relations fail to replicate across the populations for which generality is claimed.
7. EPS states add no prospective predictive or explanatory value beyond surface observation and history.
8. Claimed uncertainty remains systematically overconfident under ambiguity, missing information, or ontology novelty.

One failure may localize a boundary or implementation problem. A pattern of adequately powered, independent failures of the same theory prediction requires revision of that prediction.

17.4 Negative, null, and mixed results

Null results are interpreted against power, measurement reliability, and the equivalence region. They are not automatically evidence of no effect. Negative results remain in the evidence dashboard and publication record. Mixed results may support a narrower population, domain, or construct claim.

17.5 Appeals and independent review

AP-level decisions will include an independent audit report and a response from the originating team. Disagreements about implementation, analysis, or theory will be recorded. The public evidence, rather than authorship authority, determines the final status.

18 Current Evidence Baseline

The validation program begins from evidence already produced by SQC v1.2. It does not reset the architecture to a conceptual proposal. This section records the baseline that future studies must reproduce, extend, qualify, or overturn.

18.1 Architecture and data baseline

The current release implements a sequential EPS architecture with explicit context, values, existential appraisal, drivers, emotion, existential state, memory, action, and outcome heads. The action mapper uses the 76-dimensional EPS state and memory contract. The crossed trajectory dataset contains 14,400 records across 48 role-goal-risk cells, with 9,600 training, 2,400 validation, and 2,400 test records. Eight complete triples are absent from training.

Leakage audits report zero overlap in counterfactual group, scenario template, and episode across the relevant splits. The dataset and checkpoint must be tied to their reported hashes in the archival release, and outdated sidecars must be regenerated or marked non-authoritative (M. H. Khanyile, 2026b, 2026c).

18.2 Balanced checkpoint baseline

The production-default research checkpoint reports:

Metric	Result
Overall exact action	75.25% (1,806/2,400)
Seen-triple action	75.30% (1,506/2,000)
Unseen-triple action	75.00% (300/400)
Seen-minus-unseen gap	0.30 percentage points
Valid-action hit	97.17% (2,332/2,400)
Worst context cell	66.00% (33/50)
Existential-state error	29.21%

The worst cell is a seen combination. The near-equal unseen aggregate supports composition of known factor levels, not out-of-ontology generalization. The balanced checkpoint is a single training seed and therefore requires multi-seed confirmation.

18.3 Causal and freeze evidence

Oracle interventions on the initial v1.2 checkpoint produced the largest action gains for values, existential state, and existential appraisal. Action-only freeze audits compare 67 checkpoint tensors, with ten action tensors changed and zero non-action tensors changed in the audited experiments. These results support causal sensitivity and parameter isolation but do not replace the expanded preregistered mediation and leakage studies.

18.4 External BIB-CCS baseline

Behavioral Intelligence Benchmark - Crisis Coordination under Scarcity (BIB-CCS) is the equal-weight mean of State Inference, Policy Adaptation, Value Consistency, and Social Outcome Quality on a unit interval where higher is better. The open-track protocol uses observation packets, 12 agents, horizon 40, and five evaluation seeds.

The balanced checkpoint scores 0.7784 on the original seeds and 0.7741 on fresh seeds. BIB-only action tuning improves fresh-seed BIB-CCS to 0.8479 but reduces general SQC exact action to 44.04%. Uniform mixed replay reduces forgetting but lowers unseen action to 72% and the worst-cell floor to 60%. Logit-preserving replay preserves general action, but across training seeds 17, 42, and 73 the fresh BIB scores are 0.7741, 0.7824, and 0.7695. The mean gain is smaller than cross-seed variation. The balanced checkpoint therefore remains the default (M. H. Khanyile, 2026c).

18.5 Baseline evidence classification

The evidence supports AP-1 and substantial progress toward AP-2. Outstanding AP-2 work includes multi-seed balanced training, larger held-out cells, adversarial shortcut tests, negative-control interventions, pathway consistency, and independent reproduction. AP-3 human construct studies, AP-4 ecological studies, and AP-5 use-case assurance remain future programs.

19 Phased Execution Roadmap

The studies are organized into six phases. Later exploratory work may begin early, but a higher validation claim cannot bypass its prerequisite gates.

19.1 Phase A - Definition, traceability, and AP-2 completion

Primary studies: 1-4, 15, and targeted repetitions of current experiments.

Deliverables: finalized AP definition; EPS claim register; complete theory-to-code matrix; synchronized artifact release; balanced multi-seed replication; larger controlled test cells; shortcut, gradient, freeze, and reimplementation audits.

Exit condition: all AP-0 and AP-1 requirements confirmed and all mandatory AP-2 implementation and reproducibility gates passed.

19.2 Phase B - Internal causal and semantic validation

Primary studies: 5-8, 16-18, and architecture ablations from Study 30.

Deliverables: preregistered intervention map; negative controls; recoverability analysis; pathway-correct metrics; valid-action instruments; calibrated uncertainty and abstention; initial out-of-ontology results.

Exit condition: full AP-2 passage with demonstrated causal mediation, semantic coherence in the controlled domain, and calibrated bounded behavior.

19.3 Phase C - Human construct validation

Primary studies: 9-14 and initial prospective Study 19.

Deliverables: human measurement instruments; expert and community content review; reliability; convergent, discriminant, structural, criterion, and incremental validity; first prospective human predictions.

Exit condition: AP-3 passage for explicitly named constructs, populations, languages, and domains.

19.4 Phase D - Memory, development, culture, and civilization

Primary studies: 20-27.

Deliverables: autobiographical and relational memory; longitudinal identity tests; developmental evidence; cross-cultural invariance profiles; historical blind predictions; civilizational recurrence corpus; published boundary cases.

Exit condition: the longitudinal and cultural components required for the declared AP-4 scope are supported.

19.5 Phase E - External comparison and independent replication

Primary studies: 28-34.

Deliverables: multi-domain benchmark suite; matched-capacity comparisons; ablation and necessity evidence; external replication challenge; neural, physiological, and perturbation studies.

Exit condition: AP-4 passage in the declared domains and a supported, qualified conclusion about SQC's comparative advantage.

19.6 Phase F - Operational and HCS validation

Primary studies: 35-38.

Deliverables: use-case assurance cases; safety and subgroup studies; human oversight and appeal; validated significance measures; Neural Grid shadow metering; HCS pilots and governance.

Exit condition: AP-5 passage for each named use and successful downstream dependency validation. Failure or non-completion leaves the systems available for research and prototyping but not as validated operational infrastructure.

19.7 Sequencing and decision points

Each phase ends with an independent gate review. The review may advance, continue, narrow, revise, or stop a claim. The roadmap is updated as evidence changes, while original preregistrations and outcomes remain preserved.

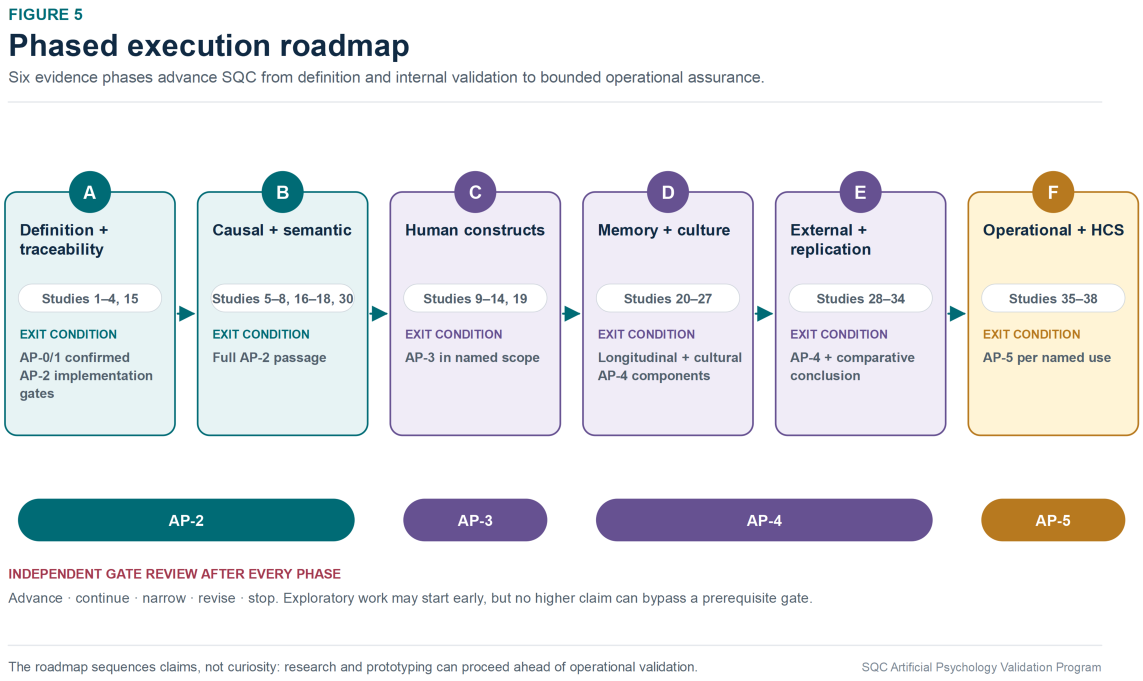

Figure 5: Phased execution roadmap. Six evidence phases move from definition, traceability, and completion of AP-2 through human construct, ecological, comparative, independent, operational, and Human-Centered Systems validation. Each phase ends in an independent gate review.

20 Governance, Ethics, and Scientific Independence

Artificial psychology concerns information about values, identity, belonging, power, emotion, relationships, and anticipated behavior. These are sensitive domains even when the model is used only for research.

Human studies require appropriate ethics review, informed consent, privacy, data minimization, withdrawal, and protection against reidentification. Cross- cultural work requires local partnership, translation accountability, and community authority over culturally sensitive interpretation. Historical and civilizational datasets require provenance, uncertainty, and avoidance of decontextualized claims about populations.

The program will disclose theoretical, authorship, financial, and downstream commercial interests. Confirmatory evaluation and AP-level review will include independent investigators. Negative and null findings, protocol deviations, known incidents, and status regressions will be published.

Consequential psychological inference requires contestability. Affected people must be able to challenge inputs, supply missing context, understand uncertainty, and appeal outcomes. SQC outputs will not be presented as certain diagnoses or complete explanations when the evidence and observation do not support them.

21 Discussion

21.1 Reframing the status of SQC

The absence of a consensus artificial-psychology standard should not force SQC into either of two misleading positions. It is inaccurate to treat SQC as merely a speculative candidate until every conceivable human study is complete. The architecture already implements an explicit EPS process, enforces a state-mediated action contract, supports intervention, and produces bounded behavioral evidence. It therefore qualifies as an implemented artificial-psychology architecture under the operational definition proposed here.

It would also be premature to treat the present checkpoint as universally validated. The current evidence is strongest for implementation, internal mediation, controlled behavior, and compositional generalization. Human construct meaning, long-term memory, cultural invariance, ecological breadth, comparative uniqueness, and operational use require the studies specified in this manuscript.

21.2 Why a cumulative framework is scientifically stronger

The AP scale allows evidence to accumulate without erasing earlier achievement or overstating later meaning. A failed cross-cultural invariance test may narrow a universal claim while leaving architecture and within-culture construct evidence intact. A shortcut discovery may invalidate an AP-2 mediation claim without deciding whether EPS is true. A successful operational pilot may validate one use without proving universal behavioral fidelity.

This localization makes the program more falsifiable, not less. It specifies which evidence would change which claim.

21.3 SQC's proposed comparative distinction

SQC's central distinction is the complete declared sequence through which context becomes value-grounded, existentially appraised, motivationally and emotionally organized, stabilized as state, connected to memory, and expressed as behavior. Neighboring cognitive and affective architectures may implement parts of this territory. The comparative question is whether another architecture implements the same required psychological integration and whether it can match SQC's fidelity under matched tests.

The claim that only SQC can approach complete behavioral fidelity is therefore retained as a north-star comparative hypothesis. It is not presented as a completed literature fact. The study program gives that hypothesis a fair route to support: SQC must outperform serious alternatives not only in average accuracy, but in calibration, transfer, pathway correctness, data efficiency, longitudinal coherence, and construct validity.

21.4 Civilizational evidence and formal science

Thousands of years of human behavior are not peripheral to EPS. They are an enormous record of how survival, belonging, identity, power, value, emotion, institutions, and aspiration recur under transformed conditions. The validation program respects that evidence by making it auditable: preregistered sampling, independent coding, blind prediction, competing explanations, multilevel models, and publication of counterexamples.

Formalization does not diminish the civilizational foundation. It converts its breadth into evidence capable of supporting a scientific priority claim.

21.5 The meaning of complete behavioral fidelity

Complete fidelity should not be confused with unjustified certainty. The north star joins representational completeness with calibrated prediction. As SQC obtains better context, memory, individual history, and measurement, its predictions should become more accurate. When necessary information is absent or several actions are psychologically viable, it should represent that uncertainty correctly.

The most successful possible SQC would not merely guess every action with high confidence. It would know which psychological distinctions determine the action, what evidence is missing, which alternatives remain viable, and how its prediction should change under intervention.

21.6 Implications for Human-Centered Computation

The wider CVEST program treats significance as a computational and economic object. SQC is the proposed engine of that computation; the Neural Grid is its metered and governed infrastructure; Human-Centered Systems are the downstream economic and cultural architecture. The validation dependency is therefore substantive rather than rhetorical (H. Khanyile, 2026c, 2026d, 2026e; M. H. Khanyile, 2026a).

This does not prohibit research or prototypes before AP-5. It defines when the word validated is justified. Measures of Cultural Impact, Cultural Yield, contribution, reward, pricing, or governance cannot be more dependable than the significance computation they use.

21.7 Limitations of the proposed program

The program is extensive and will require years, multiple disciplines, independent institutions, and substantial resources. Some constructs may lack adequate current instruments. Historical data can remain incomplete despite careful protocol. Neural correspondence may be underdetermined. Operational thresholds will vary by jurisdiction and use. The AP scale itself is a proposal that should be challenged, compared, and revised through broader scientific participation.

These limitations do not justify an undefined standard. They identify the work required to build a valid one.

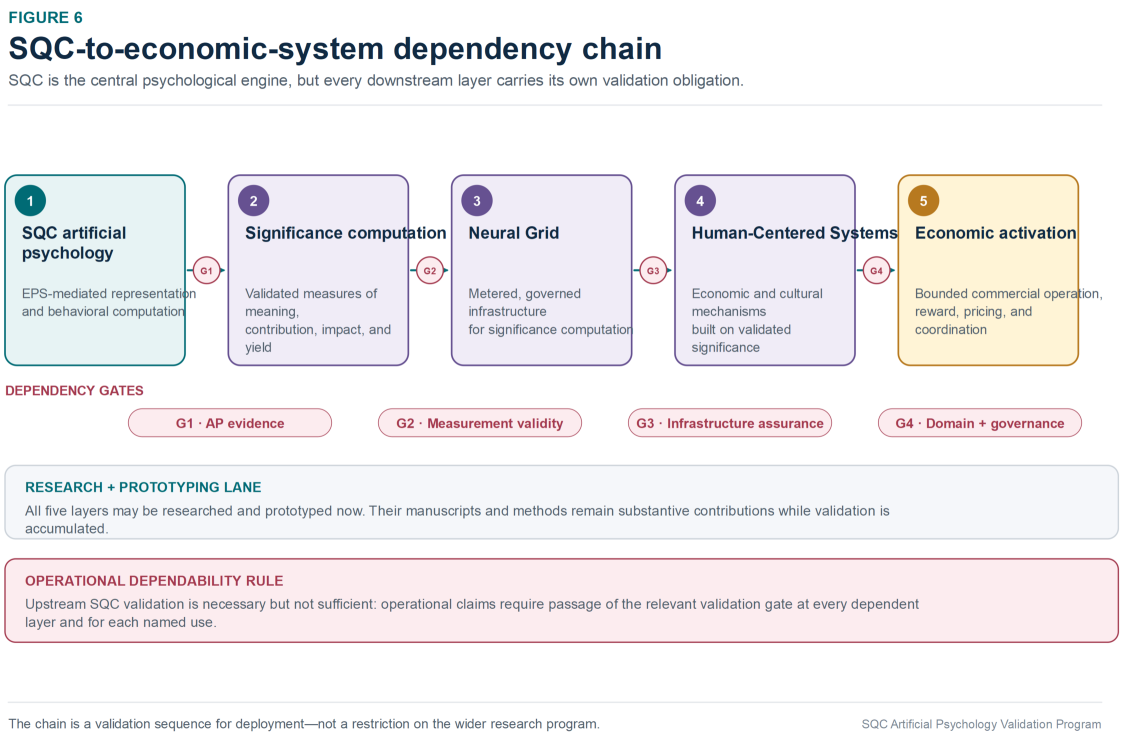

Figure 6: SQC-to-economic-system dependency chain. SQC is the upstream psychological engine, but its validation is necessary rather than sufficient for operational dependability. Significance computation, Neural Grid infrastructure, Human-Centered Systems, and each economic use retain distinct measurement, infrastructure, governance, and operational validation gates. Research and prototyping may proceed across the full chain while those gates are being tested.

22 Conclusion

SQC is an implemented artificial-psychology architecture grounded in EPS. It constructs behavior through explicit context, values, existential appraisal, Belonging, Identity, Power, emotional modulation, existential state, and memory, and it structurally restricts action to that predicted EPS pathway. Current results establish computational realization, trainability, audited parameter separation, causal sensitivity, useful behavioral performance, and composition of known decision-context factors.

Artificial psychology does not yet possess a consensus definition or validation standard against which those achievements can be classified. This manuscript supplies a proposed definition, an AP-0 through AP-5 scale, 38 coordinated studies, an evidence dashboard, statistical policy, falsification rules, phased roadmap, and governance framework. Under this scale, SQC presently qualifies at AP-1 with substantial evidence toward AP-2. This is neither a declaration of universal validation nor a description of SQC as merely an untested candidate.

The program's north star is the highest attainable fidelity and certainty in representing and predicting human behavior. EPS is proposed as the necessary causal theory of how context becomes significance and behavior; SQC is its computational vehicle. Whether SQC can approach complete fidelity more closely than any alternative is a falsifiable comparative hypothesis. It will be tested through matched architectures, human construct evidence, longitudinal memory, cross-cultural and civilizational studies, prospective prediction, calibration, independent replication, and operational assurance.

The scientific commitment is therefore explicit. The program will not claim perfect prediction before it is demonstrated, and it will not declare the north star impossible before the required architecture and evidence exist. SQC will advance by defining what success means, measuring uncertainty, exposing its internal process, publishing failures, and allowing cumulative evidence to determine how far artificial psychology can go.

The same discipline governs the downstream economic program:

```
validated artificial psychology in SQC
    -> validated significance computation
    -> Neural Grid infrastructure
    -> Human-Centered Systems
    -> economic and commercial activation
```

Research and prototyping may proceed throughout this chain. Claims of validated, operationally dependable systems require passage of the relevant gates. This is not a retreat from the ambition of SQC. It is the method through which that ambition becomes a scientific and ultimately operational reality.

Declarations

Ethics

This methods and validation-program manuscript reports no new human-participant study. Every proposed human, developmental, physiological, neural, cross-cultural, or operational study requires appropriate ethics and community review before data collection. Consent, privacy, withdrawal, cultural authority, and use restrictions form part of the study design rather than post hoc administrative additions.

Data, code, and artifact availability

The current SQC implementation and companion artifacts are maintained at `https://github.com/HoustonKhanyile/SQC_V1.2`. Each confirmatory study will cite an immutable release, code revision, environment, dataset and checkpoint hashes, preregistration, exact reproduction commands, and the governed access process for any non-public human data. The current theory-grounded dataset, checkpoint, evaluation, BIB-CCS, replay, and freeze-audit artifacts are indexed in the SQC v1.2 manuscript and crossed decision-context benchmark report.

Competing interests and positionality

The author is the developer of EPS, SQC, and the wider CVEST Human-Centered Computation program evaluated and enabled by this framework. This position provides direct theory and implementation knowledge and creates a material interest in the program's conclusions. The framework therefore requires locked tests, complete reporting, independent replication, external AP-level review, and publication of negative evidence.

Author contribution

Mvuselelo Houston Khanyile developed the EPS and SQC research program, defined the validation objectives, and authored this manuscript. Future empirical reports will identify the independent investigators, community partners, statistical reviewers, and assurance teams responsible for each study.

A Artificial Psychology Validation Checklist

A.1 AP-0 - Theoretical specification

- ☐ Constructs are defined with inclusion and exclusion boundaries.
- ☐ Causal and temporal relations are explicit.
- ☐ Behavior and memory are connected to the theory.
- ☐ Scope and levels of analysis are declared.
- ☐ Falsifiable predictions and competing accounts are recorded.
- ☐ A versioned theory specification is available.

A.2 AP-1 - Implemented architecture

- ☐ The system is executable and produces behavior.
- ☐ Mandatory theory constructs map to code.
- ☐ Internal states are inspectable.
- ☐ The implemented dataflow matches the specification.
- ☐ Behavior reads from the declared psychological pathway.
- ☐ A temporal or memory interface exists.
- ☐ Core tests and versioned artifacts are available.

A.3 AP-2 - Internal validation

- ☐ Prohibited shortcuts and leakage are excluded.
- ☐ Gradient and parameter ownership are audited.
- ☐ Causal interventions follow predicted directions.
- ☐ Negative controls demonstrate invariance.
- ☐ Internal states meet performance and calibration requirements.
- ☐ Behavior exceeds prespecified baselines.
- ☐ Locked bounded generalization succeeds.
- ☐ Principal results repeat across training seeds.
- ☐ Independent artifact reproduction succeeds.
- ☐ Pathway error and valid alternatives are measured.

A.4 AP-3 - Construct validation

- ☐ Content validity is supported.
- ☐ Human measurements are reliable.
- ☐ Convergent validity is supported.
- ☐ Discriminant validity is supported.
- ☐ Structural validity is supported against alternatives.
- ☐ Criterion and incremental validity are supported.
- ☐ Human interventions follow EPS predictions.
- ☐ Measurement equivalence is established for claimed groups.
- ☐ Independent confirmation is complete.

A.5 AP-4 - Ecological validation

- ☐ Prospective human prediction succeeds.
- ☐ People and environments are held out.
- ☐ Multiple external domains replicate.
- ☐ Cross-cultural evidence identifies invariance and variation.
- ☐ Longitudinal memory and development are coherent.
- ☐ Out-of-ontology behavior is tested.
- ☐ Historical or civilizational predictions are blinded.
- ☐ Calibration survives shift.
- ☐ Adversarial boundary cases are published.
- ☐ Independent multi-site replication succeeds.

A.6 AP-5 - Operational validation

- ☐ Intended and prohibited uses are declared.
- ☐ Risk and harm analysis is complete.
- ☐ Use-case thresholds pass prospectively.
- ☐ Human oversight and appeal work in practice.
- ☐ Consent, privacy, and data controls are operational.
- ☐ Fairness and cultural safety requirements pass.
- ☐ Security and misuse tests pass.
- ☐ Monitoring, incident response, and rollback are active.
- ☐ Accountability and decision rights are explicit.
- ☐ Independent continuing assurance is established.

B Theory-to-Code Traceability Matrix Template

Field	Required content
Claim ID	Stable identifier for the EPS proposition
Source	Versioned manuscript section or formal specification
Construct	Name and operational definition
Predicted relation	Direction, timing, moderators, and boundary
Formal representation	Equation, state space, or computational rule
SQC module	Source file, class, and function
Tensor/state	Name, shape, activation, and range
Causal parents	Permitted upstream inputs
Causal children	Permitted downstream consumers
Data field	Source, type, and missingness rule
Supervision	Label or objective construction
Loss owner	Module allowed to optimize the target
Unit/integration test	Test identifiers and expected result
Intervention	Manipulation and predicted downstream effect
Human measure	Independent instrument or observable criterion

Continued on next page

Field	Required content
Evidence status	Passed, provisional, inconclusive, failed, or not tested
Version impact	Revalidation required after change

C Study Registry

Study	AP gate	Primary hypothesis	Primary estimand	Data or population	Status
1	AP-1/AP-2	Theory and implementation are fully traceable	Mandatory mapping coverage	EPS and SQC artifacts	Planned
2	AP-2	Action uses no prohibited shortcut	Prohibited-path count	Source and runtime	Planned
3	AP-2	Module ownership is exact	Unauthorized change count	Training modes	Partial evidence
4	AP-2	Results reproduce independently	Equivalence to reference	Clean implementations	Planned
5-8	AP-2	Internal states are causal and coherent	Intervention and pathway effects	Controlled trajectories	Initial evidence/Planned
9-14	AP-3	SQC states represent intended human constructs	Reliability and validity profile	Human participants and experts	Planned
15-19	AP-2/AP-4	Behavior is accurate, calibrated, and prospective	Behavioral fidelity profile	Synthetic and human data	Partial evidence/Planned
20-23	AP-4	Psychology remains coherent over history	Longitudinal state and prediction	Repeated trajectories	Planned
24-27	AP-4	EPS recurs with bounded cultural variation	Invariance and recurrence	Cultures and historical records	Planned
28-31	AP-4	SQC adds comparative value	Matched multidimensional difference	External domains and architectures	Planned
32-34	AP-3/AP-4	EPS has specific neural and physiological correspondence	Preregistered correspondence	Human multimethod data	Planned
35-38	AP-5	Named uses and downstream dependencies are dependable	Use-case assurance profile	Shadow and live pilots	Planned

Each study receives a separate registry entry before execution, including sample-size justification, full outcomes, thresholds, analysis code, and artifact identifiers.

D Evidence and Claim Matrix Template

Claim	Unit	Supporting evidence	Contradictory evidence	Estimate and uncertainty	Scope	AP requirement	Next decisive test
Implementation fidelity	Source implementation	Traceability and audits	Open discrepancies	Versioned audit result	SQC version	AP-1	Independent reproduction
Causal mediation	Architecture/ checkpoint	Interventions and shortcut tests	Bypass or wrong-direction effects	Causal effect and interval	Tested domains	AP-2	Locked mediation study
Construct validity	EPS/SQC state	Reliability and validity battery	Weak or non-specific relation	Multimethod estimate	Population/ culture	AP-3	Independent confirmation
Behavioral fidelity	Checkpoint	Prospective calibrated predictions	Error, overconfidence, subgroup harm	Proper scores and intervals	Domain/horizon	AP-2/AP-4	Harder locked test
Operational dependability	Intended use	Shadow and monitored deployment	Incidents and drift	Harm-weighted risk	Use/population	AP-5	Continuing assurance

E Minimum Reporting Standard

Every confirmatory SQC report will include:

1. theory, ontology, architecture, implementation, data, checkpoint, and protocol versions;
2. intended claim, validation unit, population, domain, and time horizon;
3. preregistration and all deviations;
4. sampling frame, inclusion, exclusion, missingness, and dependence structure;
5. model inputs, outputs, active optional modules, parameter count, and compute;
6. training, selection, confirmation, and locked-test procedures;
7. every seed, checkpoint-selection opportunity, and baseline;
8. counts, point estimates, effect sizes, intervals, calibration, and subgroup results;
9. all primary outcomes and corrected secondary analyses;
10. causal, leakage, gradient, freeze, and pathway audits relevant to the claim;
11. failures, negative results, ambiguity, and valid alternatives;
12. code, data or governed-access path, environment, commands, and hashes;
13. authorship, funding, theoretical, and commercial interests;
14. ethics, consent, privacy, community, and governance procedures;
15. evidence-dashboard update and resulting AP status.

F Glossary

Artificial psychology: The engineering and study of computational systems whose explicit, causally active psychological organization generates adaptive behavior over time.

Evolutionary Psychological Structures (EPS): The source theory underlying SQC, organizing contextual interpretation, values, existential appraisal, Belonging, Identity, Power, emotion, existential state, memory, and behavior.

Sensia Quotia Computation (SQC): The trainable computational implementation of EPS evaluated in this validation program.

Implementation fidelity: Agreement between a theory or architecture specification and its executable realization.

Causal mediation: The condition in which the declared internal state participates materially in generating downstream behavior.

Construct validity: The accumulated evidence that a measure or representation has the psychological interpretation attributed to it.

Behavioral fidelity: Accuracy, distributional agreement, calibration, pathway coherence, and uncertainty in representing or predicting behavior.

Valid action: An action judged psychologically coherent under the available state and context, whether or not it is the single canonical label.

Calibration: Agreement between predicted probability and empirical frequency.

Abstention: A decision not to make a substantive prediction because information or model competence is insufficient.

Out-of-ontology: A case requiring a role, value, state, action, relation, or mechanism absent from the trained representation.

Ecological validity: Evidence that findings generalize to natural people, settings, cultures, and time.

Operational validation: Use-case-specific evidence that a named system is dependable, safe, governed, monitored, and contestable.

Human-Centered Computation: Computation organized around the significance information acquires relative to human context, values, identity, belonging, motivation, emotion, and anticipated futures.

Neural Grid: The proposed metered, routed, governed, and settled infrastructure through which SQC work becomes operational.

Human-Centered Systems (HCS): Downstream computational, cultural, and economic systems built around validated human-centered significance.

References

Anderson, J. R., and Lebiere, C. L. (2003). The Newell test for a theory of cognition. *Behavioral and Brain Sciences*, 26(5), 587-601. https://doi.org/10.1017/S0140525X0300013X

Bach, D. R., Melinscak, F., Fleming, S. M., and Voelkle, M. C. (2020). Calibrating the experimental measurement of psychological attributes. *Nature Human Behaviour*, 4, 1229-1235. https://doi.org/10.1038/s41562-020-00976-8

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105. https://doi.org/10.1037/h0046016

Cronbach, L. J., and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. https://doi.org/10.1037/h0040957

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. G. P. Putnam's Sons.

Gratch, J., and Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4), 269-306. https://doi.org/10.1016/j.cogsys.2004.02.002

Khanyile, H. (2026a). *Evolutionary Psychology Structures: An Understanding of the Socialization Mechanism*. CVEST manuscript.

Khanyile, H. (2026b). *Evolutionary Psychological Structures (EPS): A Neurocomputational Architecture for Value-Grounded, Emotion-Modulated Artificial Intelligence*. CVEST manuscript.

Khanyile, H. (2026c). *The Neural Grid: Metered Infrastructure for Human-Centered Computation*. CVEST manuscript.

Khanyile, H. (2026d). *Human-Centered Systems (HCS): Engineering Culture as a Primary Economic Layer*. CVEST manuscript.

Khanyile, H. (2026e). *Cultural Impact Mechanics in Music 2.0*. CVEST manuscript.

Khanyile, M. H. (2026a). *The Economic Value of Human Significance: Why Human-Centered Computation Represents the Next Accelerated Computing Paradigm*. CVEST manuscript.

Khanyile, M. H. (2026b). *Sensia Quotia Computation 1.2: A Trainable EPS-Complete Architecture for Subjective-State-Oriented Behavioral Intelligence*. CVEST manuscript.

Khanyile, M. H. (2026c). *SQC 1.1 - V1.2 Crossed Decision-Context Training and Benchmark Report*. CVEST empirical report.

Langley, P. (2017). Progress and challenges in research on cognitive architectures. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1). https://doi.org/10.1609/aaai.v31i1.11150

Lazarus, R. S. (1991). *Emotion and Adaptation*. Oxford University Press.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. https://doi.org/10.1037/0003-066X.50.9.741

Picard, R. W. (1997). *Affective Computing*. MIT Press.

Prescott, T. J., and Camilleri, D. (2019). The synthetic psychology of the self. In M. I. A. Ferreira, J. Sequeira, and R. Ventura (Eds.), *Cognitive Architectures* (pp. 85-104). Springer. https://doi.org/10.1007/978-3-319-97550-4_7

Scherer, K. R., Schorr, A., and Johnstone, T. (Eds.). (2001). *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford University Press.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25, 1-65. https://doi.org/10.1016/S0065-2601(08)60281-6

Serapio-Garcia, G., Safdari, M., Crepy, C., Sun, L., Fitz, S., Romero, P., Abdulhai, M., Faust, A., and Mataric, M. (2025). A psychometric framework for evaluating and shaping personality traits in large language models. *Nature Machine Intelligence*, 7, 1954-1968. https://doi.org/10.1038/s42256-025-01115-6

Trognon, A., Cherifi, Y. I., Habibi, I., Demange, L., and Prudent, C. (2022). Using machine-learning strategies to solve psychometric problems. *Scientific Reports*, 12, 18922. https://doi.org/10.1038/s41598-022-23678-9

Wray, R. E., and Lebiere, C. (2007). Metrics for cognitive architecture evaluation. In *AAAI Workshop on Evaluating Architectures for Intelligence* (pp. 59-66). AAAI Press.

Zavlis, O., Story, G., Friedrich, C., Fonagy, P., and Moutoussis, M. (2025). A systematic review of computational modeling of interpersonal dynamics in psychopathology. *Nature Mental Health*, 3, 932-942. https://doi.org/10.1038/s44220-025-00465-9