MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

UNITED STATES: THE GLOBAL HEALTH TECHNOLOGY ASSESSMENT MEMEPLEX OF MEASUREMENT FAILURE

A LOGIT INTERROGATION OF THE KNOWLEDGE BASE OF HEALTH TECHNOLOGY ASSESSMENT

Paul C Langley Ph.D Adjunct Professor, College of Pharmacy, University of Minnesota, Minneapolis, MN

LOGIT WORKING PAPER No 760 FEBRUARY 2026

www.maimonresearch.com

Tucson AZ

THE GLOBAL HEALTH TECHNOLOGY ASSESSMENT MEMEPLEX OF MEASUREMENT FAILURE

A LOGIT INTERROGATION OF THE KNOWLEDGE BASE OF HEALTH TECHNOLOGY ASSESSMENT

Paul C Langley Ph.D., Adjunct Professor, College of Pharmacy, University of Minnesota, Minneapolis MN

ABSTRACT

Health technology assessment (HTA) presents itself as a scientific enterprise grounded in quantitative evaluation of therapeutic impact. This claim presupposes measurement. Arithmetic operations such as multiplication, aggregation, and ratio comparison are admissible only when the underlying constructs satisfy the axioms of representational measurement. This paper reports the results of a systematic interrogation of the global HTA knowledge base using a canonical 24-item representational measurement diagnostic, with categorical endorsement probabilities transformed into normalized logit measures. This diagnostic allows empirical assessment of whether measurement axioms operate as binding constraints within HTA's evaluative framework.

The results reveal a consistent and global pattern. Statements expressing foundational measurement requirement including unidimensionality, admissible scale transformations, and the necessity that measurement precede arithmetic collapse to floor logit values, indicating structural non-possession. Conversely, false propositions endorsing composite utility constructs, QALY aggregation, and simulation-based cost-effectiveness ratios register positive logits, indicating operational acceptance. This inversion demonstrates that HTA performs arithmetic operations on constructs whose measurement properties are neither demonstrated nor required. Quantitative outputs are generated and used to inform decisions affecting therapy access, pricing, and resource allocation despite lacking empirical measurement foundation.

This condition defines the existence of a global HTA memeplex: a self-reinforcing system of numerical belief stabilized through institutional adoption rather than empirical validation. Within this framework, numbers function as administrative instruments rather than representations of measurable quantities. The consequences extend beyond methodology. Measurement absence prevents falsification, arrests the evolution of objective knowledge, and undermines the empirical legitimacy of decisions affecting patient care.

Recovery requires structural reconstruction. Manifest attributes must be evaluated using linear ratio measurement, and latent constructs must be measured using Rasch logit ratio scales. Only by restoring measurement as the foundation of quantitative evaluation can HTA re-enter the domain of normal science and fulfill its duty of care to patients, physicians, and healthcare systems.

INTRODUCTION

Health technology assessment (HTA) presents itself as a scientific enterprise whose central purpose is to quantify therapeutic impact, compare competing interventions, and guide rational allocation of healthcare resources. These objectives presuppose the existence of measurement. Arithmetic operations such as addition, multiplication, and ratio comparison are not self-justifying; they are admissible only when the underlying quantities satisfy the axioms of representational measurement. Unidimensionality, invariance, admissible transformations, and the existence of ratio scale structure are not optional refinements but necessary preconditions for quantification. Without measurement, arithmetic operations lack empirical meaning. Numbers may be generated, manipulated, and reported, but they do not represent quantities. In such circumstances, quantitative outputs become administrative artifacts rather than empirical discoveries. The distinction is decisive. Measurement enables falsification, replication, and the accumulation of objective knowledge. Without measurement, claims cannot be empirically tested. They can be recalculated, but they cannot be wrong.

Over the past four decades, HTA has constructed an elaborate quantitative architecture centered on composite utility instruments, quality-adjusted life years (QALYs), and simulation-based cost-effectiveness models. These constructs are presented as measurement-based tools capable of expressing therapeutic value on a common numerical scale. Their outputs inform reimbursement decisions, pricing negotiations, and clinical access across publicly funded and private healthcare systems worldwide. Yet the legitimacy of this architecture depends entirely on whether its underlying numerical constructs satisfy the axioms required for measurement. If they do not, the quantitative framework of HTA lacks scientific foundation regardless of its procedural sophistication. Institutional adoption, methodological consensus, and computational complexity cannot substitute for measurement validity. The authority of quantitative claims derives not from their institutional endorsement but from their adherence to representational measurement requirements.

This paper reports the results of a systematic interrogation of the global HTA knowledge base using a canonical 24-item representational measurement diagnostic, with categorical endorsement probabilities transformed into normalized logit measures. This diagnostic allows direct empirical assessment of whether the foundational axioms of measurement operate as binding constraints within HTA's epistemic and operational structures. The analysis has been applied across national HTA systems, operational agencies, academic journals, and multiattribute utility instruments. The results reveal a consistent and striking pattern. Statements expressing the axioms necessary for measurement, such as the requirement that measurement precede arithmetic, that multiplication requires ratio scale properties, and that latent constructs require Rasch transformation to establish invariant measurement, collapse to floor logit values or adjacent negative ranges. Conversely, statements endorsing the arithmetic use of composite utility constructs, aggregation of QALYs, and reliance on simulation-based cost-effectiveness ratios register positive logit values, indicating operational acceptance. This inversion demonstrates that HTA's quantitative framework operates independently of representational measurement constraints.

These findings identify the existence of a global HTA memeplex: a self-reinforcing system of quantitative belief that reproduces itself across jurisdictions, institutions, and professional

communities. Within this system, numerical outputs acquire authority through institutional endorsement rather than measurement validity. The appearance of quantification substitutes for quantification itself. The individual country, agency, and journal-level logit assessments that follow provide direct empirical evidence of this condition within specific operational contexts. This paper establishes the global structure within which those national and institutional profiles must be understood. The question is no longer whether measurement axioms matter. The question is whether HTA, as currently practiced, satisfies them.

SECTION 1: THE CANONICAL LOGIT DIAGNOSTIC: INTERROGATING THE MEASUREMENT FOUNDATIONS OF HTA

Scientific disciplines distinguish themselves from belief systems through their commitment to measurement. Measurement is not the assignment of numbers to observations; it is the demonstration that numerical relations correspond to empirical relations in the attribute being measured. This distinction is foundational. It determines whether arithmetic operations performed on those numbers produce empirically meaningful results or merely numerical artifacts. Representational measurement theory formalizes this requirement by specifying the axioms that must be satisfied before arithmetic operations such as addition, multiplication, and ratio comparison are admissible [1]. These axioms are not matters of convention. They are logical constraints governing the relationship between empirical attributes and their numerical representation. If these constraints are not satisfied, arithmetic operations lose empirical meaning. Numbers may still be manipulated, but they no longer measure anything.

The canonical 24-item representational measurement diagnostic was developed to determine whether these axioms operate as binding constraints within the individual HTA knowledge base. Each statement expresses either a foundational truth required for measurement or a known violation embedded in current HTA practice. Statements such as "measurement must precede arithmetic," "multiplication requires a ratio measure," and "latent traits require Rasch transformation to achieve invariant measurement" express necessary conditions for lawful quantification. Conversely, statements such as "the QALY is a ratio measure," "summations of ordinal responses create ratio measures," and "reference case simulations generate falsifiable claims" represent propositions known to violate representational measurement requirements. Together, these statements form a diagnostic instrument capable of distinguishing between knowledge systems grounded in measurement and those grounded in numerical convention.

The diagnostic does not rely on subjective interpretation. Each statement is assigned a categorical endorsement probability derived from interrogation of the HTA knowledge base within the large language model. Within a defined knowledge base, the interrogation process operates by establishing explicit epistemic boundaries defined by documents, methods papers, guidelines, instrument development reports, applied evaluations, and methodological commentaries that collectively define what the system knows, accepts, and reproduces. These boundaries are not abstract; they are operational. They delimit the universe of admissible claims and methodological assumptions. Each canonical statement in the 24-item diagnostic functions as a probe directed at this bounded corpus. The question is not whether a statement is philosophically correct in isolation, but whether the knowledge base itself contains, reinforces, or contradicts that statement as an operational principle. The interrogation therefore asks a strictly internal question: if one were trained solely within this corpus, what belief about measurement would one acquire? The categorical probability assigned to each statement represents the degree to which the knowledge base reinforces that statement as part of its operational logic.

The assignment of categorical probability reflects the structural consistency with which a statement is supported, contradicted, or excluded by the knowledge base. When the corpus repeatedly employs arithmetic operations requiring ratio measurement such as multiplication, aggregation, or ratio comparison without establishing ratio scale properties, it implicitly endorses

the false proposition that such arithmetic is admissible. This generates high categorical probability for the false statement and correspondingly low probability for its true counterpart. Conversely, when a principle is absent such as the requirement for Rasch transformation to convert ordinal observations into invariant interval measures the knowledge base does not reinforce that principle as a necessary condition for quantitative claims. The categorical probability therefore falls to the floor, reflecting non-possession rather than explicit rejection. The probability is thus not a frequency count of explicit statements, but a structural inference drawn from the internal coherence of the knowledge system: what must be believed for the system's quantitative operations to remain internally consistent.

This process may be understood as an epistemic inversion test structured around both true and false canonical statements. Each true statement represents a necessary axiom of representational measurement, while each false statement represents a known violation of those axioms. The interrogation examines whether the knowledge base behaves as if the true statements were binding constraints and the false statements were inadmissible, or whether this relationship is inverted. When a true statement receives low categorical probability, this indicates non-possession: the knowledge base does not operationally recognize the axiom as a necessary precondition for quantitative claims. When a false statement receives high categorical probability, this indicates active endorsement: the knowledge base behaves as if the violation were legitimate. The categorical probability therefore quantifies both correct and incorrect endorsement. It captures the degree to which true measurement axioms are reinforced or ignored, and the degree to which false measurement beliefs are rejected or institutionalized. The resulting probability profile provides a structural measure of epistemic alignment or inversion. It reveals whether the knowledge base operates within the lawful constraints of measurement or within a framework that substitutes scoring, convention, and administrative arithmetic for measurement-valid quantity.

A critical point in interpretation is that these categorical probabilities are not statistical frequencies derived from sampling. Such frequencies would be impossible because the object of analysis is not a population of respondents but a bounded knowledge base defined by its documents, methodological guidance, published analyses, and operational rules. The probability assignment is therefore an epistemic classification, not a sampling estimate. It represents the degree to which the knowledge base, when interrogated as a structured corpus, behaves as if a given canonical statement is operationally true or false. The assignment reflects internal coherence, consistency of application, and structural reinforcement; not counting.

Accordingly, the importance of the results lies not in any single categorical probability, but in the pattern across statements. When true axioms of representational measurement repeatedly receive low probabilities while false statements receive high probabilities, this reveals systematic inversion. The knowledge base is not merely inconsistent; it is structurally misaligned with the requirements of measurement. Conversely, when true statements receive high probabilities and false statements receive low probabilities, this indicates structural possession of measurement principles. The interpretation is therefore architectural rather than numerical. It assesses whether the evaluative framework is grounded in lawful measurement or operates independently of it.

For that reason, there is no overall score and none is required. Measurement validity is not a continuous trait that can be meaningfully summed across axioms. The axioms are conjunctive

requirements: each must be satisfied. Failure of any axiom invalidates the admissibility of arithmetic operations on the affected construct. The diagnostic therefore operates as a structural profile, not a scalar index. Its purpose is to reveal whether the knowledge base functions within the constraints of representational measurement or whether those constraints are absent as operational determinants.

These probabilities are then transformed into normalized logits, producing a linear interval scale representing the structural presence or absence of each measurement principle. The logit transformation is critical. It converts categorical endorsement probabilities into a form that supports meaningful comparison across statements, institutions, and jurisdictions. Positive logits indicate endorsement or structural presence within the knowledge base. Negative logits indicate absence or exclusion. Floor values at −2.50 represent effective non-possession: the proposition does not operate as a binding constraint within the evaluative framework.

This approach differs fundamentally from narrative critique. Narrative critique may identify conceptual inconsistencies, but it cannot quantify the structural status of measurement principles within a defined knowledge base. The logit diagnostic provides empirical evidence. It demonstrates not merely that measurement axioms are debated, misunderstood, or inconsistently applied, but whether they function operationally as constraints on quantitative practice. This distinction is decisive. A principle may be acknowledged rhetorically while remaining absent operationally. The logit diagnostic captures operational presence, not rhetorical recognition.

The application of this diagnostic across national HTA systems, operational agencies, academic journals, and utility instrument frameworks reveal a consistent structural pattern. Measurement-valid propositions repeatedly collapse to negative logits or floor values. Statements requiring unidimensionality, ratio scale structure, and Rasch transformation register at −2.50 or adjacent ranges [2] . These values indicate structural exclusion. They demonstrate that measurement axioms do not function as prerequisites for quantitative claims within the range of HTA knowledge bases. Arithmetic operations are performed independently of measurement validation.

Conversely, statements endorsing constructs that violate measurement axioms consistently register positive logits. Propositions asserting that QALYs can be aggregated, that composite utility scores support multiplication, and that simulation outputs constitute evaluable evidence show strong structural presence. These findings demonstrate that the HTA knowledge base operates on a numerical architecture detached from measurement requirements. Numbers are generated, manipulated, and compared without demonstration that they represent quantities.

The significance of this pattern extends beyond methodological critique. Measurement is the mechanism through which science distinguishes between belief and knowledge. Measurement-valid constructs can be falsified. They can be replicated. They allow cumulative knowledge development. Measurement-invalid constructs cannot. Their numerical outputs may be internally consistent, but they lack empirical referents. They cannot be proven wrong because they do not measure anything. Their persistence reflects institutional stabilization rather than empirical validation.

The canonical logit diagnostic therefore provides for the first time direct empirical evidence regarding the epistemic status of HTA's quantitative framework. It demonstrates whether HTA operates within the tradition of measurement-based science or within a self-reinforcing numerical pseudo-science belief system. The results, as the subsequent national and institutional profiles demonstrate, consistently identify the latter. The axioms of representational measurement do not operate as binding constraints within the global HTA knowledge base. Instead, arithmetic operations are performed on constructs whose measurement properties are neither demonstrated nor required. The implications of this finding define the central problem addressed in this paper; HTA is bunk not science.

Taken together, the systematic non-endorsement of true measurement axioms and the strong endorsement of false measurement propositions define what may properly be described, in Dawkins' sense, as a global HTA false measurement memeplex [3]. Richard Dawkins introduced the concept of the meme to describe a unit of cultural transmission; an idea, practice, or belief that replicates itself by propagating through communication, institutional reinforcement, and imitation rather than empirical verification. A memeplex is a mutually reinforcing cluster of such memes that stabilizes itself by creating an internally coherent belief system. Its persistence depends not on correspondence with reality, but on its ability to reproduce within host institutions and resist displacement.

The HTA memeplex can be understood as a relativist belief system in which the legitimacy of quantitative claims derives not from their correspondence with measurable attributes of reality, but from their acceptance within an institutional community. In such a framework, success is judged by the ability to mobilize consensus, through guidelines, publication norms, and professional endorsement, rather than by satisfying the axioms of representational measurement. Numerical constructs such as utilities, QALYs, and cost-effectiveness ratios acquire authority because they are widely used, not because they have been demonstrated to possess invariant unit structure or admissible scale properties. This is the defining feature of epistemic relativism: truth becomes a function of agreement rather than measurement. By contrast, normal science proceeds in the opposite direction. Measurement establishes quantities whose properties are independent of belief, authority, or convention. Claims are exposed to falsification because they refer to attributes that can be observed, measured, and replicated. Within a relativist memeplex, however, constructs that lack measurement validity can persist indefinitely, because their survival depends on institutional reinforcement rather than empirical exposure. The global HTA framework exhibits precisely this structure. Its central quantitative constructs are sustained by methodological consensus, professional training, and administrative adoption, even as the logit evidence demonstrates the systematic absence of the measurement axioms required to support arithmetic. The result is a numerically elaborate but epistemically closed and barren system in which rhetorical coherence substitutes for measurement, and consensus substitutes for truth.

The HTA false measurement memeplex exhibits precisely these characteristics. At its core are several interlocking propositions: that multiattribute utility scores constitute measurable quantities; that preference-weighted composite indices can be multiplied by time to produce QALYs; that cost-per-QALY ratios are valid quantitative objects; and that simulation outputs derived from these constructs represent empirical evidence. None of these propositions satisfies the axioms of representational measurement. Yet they are propagated as methodological

requirements, embedded in guidelines, taught in academic programs, enforced by reimbursement agencies, and reproduced in peer-reviewed journals. Their survival depends not on measurement validity, but on institutional replication.

The logit diagnostic profile provides direct empirical evidence of this barren memeplex structure. True canonical statements, such as the requirement that measurement precede arithmetic, that multiplication requires ratio scale properties, and that latent constructs require Rasch transformation to achieve invariant measurement all collapse to floor values. This indicates absence of possession. These propositions do not function as operational constraints. In contrast, false statements such as the claim that QALYs are ratio measures, that summated ordinal scores can be treated as arithmetic quantities, and that simulation models generate evaluable empirical claims receive strong endorsement. This inversion demonstrates that the knowledge base is not neutral with respect to measurement axioms. It actively stabilizes propositions that violate them while excluding propositions that would enforce them.

This pattern is the signature of a memeplex rather than a scientific framework. In science, false propositions are eliminated through falsification, and measurement-valid constructs replace measurement-invalid predecessors. In a memeplex, however, replication occurs through institutional authority and pedagogical continuity. Students are trained to apply QALYs. Analysts are required to present cost-effectiveness ratios. Agencies mandate their use in reimbursement submissions. Journals publish analyses based on these constructs. Each replication reinforces the others, creating a closed system in which the memeplex sustains itself independent of falsification.

This relativist structure is not merely an abstract philosophical characterization; it is visible in concrete institutional artifacts. If independent evidence were required beyond the logit assessments, one need look no further than the widespread endorsement of the CHEERS 2022 reporting guidance [4] . CHEERS 2022 does not require that submitted claims satisfy the axioms of representational measurement. It does not require demonstration of unidimensionality, invariant unit structure, or admissible scale transformations. It does not require that quantities combined arithmetically possess ratio scale properties or dimensional homogeneity. Instead, it provides a standardized reporting framework for modeled cost-effectiveness analyses built upon composite utility constructs and simulation outputs. In doing so, it implicitly defines modeled reference case simulation as an acceptable basis for endlessly repetitive quantitative claims, regardless of whether the underlying numerical objects constitute measures.

This is the defining characteristic of a memeplex. The CHEERS framework stabilizes and reproduces a belief system by specifying how its outputs should be presented, not whether they are measurement-valid. It ensures that submissions conform to established narrative and structural expectations, thereby reinforcing institutional legitimacy. The focus shifts from measurement validity to reporting completeness. Analysts are instructed to describe model structure, assumptions, sensitivity analyses, and uncertainty intervals, but not to demonstrate that the quantities being manipulated possess lawful measurement properties. The result is a self-reinforcing evaluative environment in which the appearance of quantitative rigor substitutes for measurement discipline.

More importantly, CHEERS 2022 functions as a transmission mechanism. It enables journals, reviewers, and agencies to propagate the reference case simulation framework across jurisdictions and institutional contexts. By providing a checklist for acceptable reporting, it ensures that successive generations of analysts reproduce the same quantitative architecture. Composite utility indices are multiplied by time. Costs, expressed as monetary aggregates of heterogeneous resource inputs, are divided by composite effectiveness scores. Simulation models generate precise numerical outputs extending decades into hypothetical futures. These outputs are then presented as evidence, even though they originate entirely from assumptions applied to constructs lacking demonstrated measurement validity.

The authority of CHEERS 2022 therefore derives not from its correspondence to measurement science, but from its role in coordinating institutional belief. It creates uniformity of presentation, which in turn creates the appearance of scientific consensus. Yet consensus is not measurement. The axioms of representational measurement are not subject to negotiation or majority vote. They define the conditions under which numbers represent empirical attributes. A reporting framework that ignores these conditions cannot transform non-measures into measures by standardizing their description.

The logit evidence and the CHEERS framework thus converge on the same conclusion. The memeplex is sustained not by empirical validation but by institutional reinforcement. CHEERS 2022 does not correct the absence of measurement foundations; it codifies procedures for operating without them. Its success lies in its ability to mobilize agreement, coordinate practice, and reproduce the evaluative framework across journals and agencies. In this sense, it exemplifies the relativist principle: numerical claims achieve legitimacy through acceptance, not through measurement.

This is why the persistence of the reference case simulation framework cannot be attributed to ignorance alone. The measurement axioms have long been established and are readily accessible from Stevens seminal paper on the scales of measurement and allowable arithmetic operations published in 1946 [5]. Their continued exclusion reflects the stabilizing force of a memeplex whose survival depends on the continued production, publication, and acceptance of numerical constructs that function administratively but lack measurement validity.

The consequence is epistemic stabilization without measurement foundation. Numerical outputs acquire authority through repetition and institutional endorsement rather than through demonstration of lawful measurement structure. The memeplex persists because it is administratively useful and culturally entrenched, not because it satisfies the axioms that define quantitative science.

Dawkins emphasized that memeplexes can persist even when they are false, provided they are effective at replication. The global HTA false measurement memeplex exemplifies this principle. Its defining feature is not mere error, but structured resistance to the axioms of measurement that would expose that error. The logit profile makes this visible. It reveals a globally replicated evaluative architecture in which false measurement propositions are stabilized, propagated, and enforced as if they constituted scientific knowledge. Any concept of commitment to the evolution

of what Popper describes as the evolution of objective knowledge is outside their cognitive boundary [6].

SECTION 2: THE GLOBAL LOGIT PROFILE: A SYSTEMATIC INVERSION OF MEASUREMENT AXIOMS

The application of the canonical 24-item representational measurement diagnostic across national HTA systems, operational agencies, academic journals, and utility instrument frameworks reveal a pattern that is both consistent and decisive. The axioms that define measurement do not function as binding constraints within the global HTA knowledge base. Instead, they are systematically absent, while propositions that violate those axioms are structurally embedded. This pattern is not confined to individual countries, institutions, or journals. It is global in scope. It reflects not isolated misunderstanding but a stabilized epistemic architecture in which numerical constructs are treated as measures despite lacking the structural properties required for measurement.

The most striking feature of the global logit profile is the repeated collapse of foundational measurement axioms to floor values of −2.50 or adjacent negative ranges. Statements asserting that measurement must precede arithmetic, that multiplication requires ratio scale structure, and that latent constructs require Rasch transformation to establish invariant measurement repeatedly register at these floor values. These values have a precise interpretation. They do not indicate disagreement or controversy. They indicate non-possession. The principles expressed by these statements do not operate as structural constraints within the HTA knowledge base. They do not function as prerequisites for performing arithmetic operations. Instead, arithmetic is performed independently of measurement validation. The numerical architecture of HTA proceeds without satisfying the conditions that make quantification possible.

The absence of Rasch transformation is particularly revealing. Latent constructs such as need fulfilment, functional status, and symptom burden cannot be directly observed. Measurement of such constructs requires transformation of ordinal observations into interval measures while preserving invariance across persons and items. Rasch measurement provides the only lawful transformation model capable of achieving this [7]. Its absence indicates that latent constructs within HTA are not measured. They are scored. Numerical outputs derived from preference-weighted multiattribute instruments represent aggregated ordinal responses rather than invariant quantities. Yet these scores are routinely treated as if they possess interval or ratio scale properties. They are multiplied, divided, aggregated, and compared across therapies. The logit profile demonstrates that the measurement transformation required to justify these operations is structurally absent.

At the same time, false propositions that violate representational measurement axioms consistently register strong positive logits. Statements asserting that the QALY is a ratio measure, that composite utility scores support arithmetic operations, and that cost-effectiveness simulation outputs constitute evaluable evidence show widespread structural endorsement. These positive logits indicate operational presence. They demonstrate that these propositions function as binding elements within the HTA evaluative framework. They are not peripheral beliefs. They are central operational assumptions. They define how therapies are evaluated, compared, and ranked.

This combination of negative logits for true measurement axioms and positive logits for false measurement propositions constitutes a systematic inversion. The principles required for measurement are excluded, while propositions that violate those principles are embedded. Arithmetic operations are performed on constructs whose scale properties do not support

arithmetic. Composite utility scores are treated as if they were invariant quantities. Cost-effectiveness ratios are constructed from numerators and denominators that lack dimensional homogeneity. Simulation outputs derived from these constructs are treated as empirical evidence despite being functions of model assumptions rather than measurements.

The global consistency of this pattern is critical. The inversion appears across jurisdictions with different healthcare systems, governance structures, and policy traditions. It appears in academic journals, national HTA guidelines, and operational decision frameworks. It appears in countries widely regarded as methodological leaders and in those adopting established reference case models. The logit profiles demonstrate structural convergence. This convergence cannot be attributed to local misunderstanding or isolated error. It reflects diffusion of a shared evaluative architecture, a memeplex (as described above) within which quantitative practice is defined independently of measurement requirements.

This convergence also demonstrates that procedural sophistication does not imply measurement validity. Many HTA systems exhibit high levels of technical competence. Their guidelines are detailed. Their models are computationally complex. Their analyses are internally coherent. Yet the logit evidence shows that computational sophistication coexists with measurement absence. Numerical complexity cannot compensate for failure to satisfy representational measurement axioms. Arithmetic operations remain inadmissible regardless of model sophistication if the underlying quantities do not possess appropriate scale properties.

For members of the HTA memeplex, the proposition that there are only two admissible forms of measurement, a linear ratio scale for manifest attributes and a Rasch logit ratio scale for latent constructs, does not function as a contested scientific claim. It does not provoke counterargument, methodological debate, or empirical challenge. Instead, it typically produces incomprehension. The reason is structural. This proposition lies outside the conceptual architecture within which the memeplex operates. Its evaluative framework assumes, as foundational premises, the legitimacy of composite utility indices, preference-weighted scores, and cost-effectiveness ratios. These assumptions define what counts as meaningful quantification. A claim that challenges these premises at the level of measurement axioms does not enter the system as a hypothesis to be examined; it fails to register as a meaningful statement at all.

This reaction is characteristic of closed epistemic systems. Within such systems, propositions are not evaluated solely on empirical or logical grounds, but on their compatibility with the internal structure of accepted beliefs. The memeplex defines, implicitly and explicitly, the boundaries of admissible discourse. Constructs such as the QALY are treated as self-evidently quantitative because they are embedded in guidelines, teaching materials, and institutional procedures. By contrast, the assertion that arithmetic operations require ratio scale properties, or that latent constructs must be transformed through Rasch measurement to achieve invariant unit structure, falls outside these established conventions. It is not rejected through falsification; it is excluded through non-recognition. The conceptual vocabulary required to interpret the proposition has not been incorporated into the knowledge base.

This exclusion has profound implications for the evolution of objective knowledge within HTA. Scientific progress depends on the ability of a discipline to recognize and evaluate foundational

challenges to its measurement framework. When such challenges cannot be recognized as meaningful propositions, the possibility of falsification disappears. The system becomes self-referential. Its constructs persist not because they have survived empirical testing, but because the criteria required to test them are absent from the evaluative framework. In this sense, the assertion that only linear ratio and Rasch logit ratio scales constitute lawful measurement does not threaten the memeplex; it bypasses it. It exposes a structural boundary beyond which the memeplex cannot interpret, evaluate, or respond.

The global logit profile therefore identifies a structural condition rather than a methodological disagreement. The HTA knowledge base operates within a numerical framework that does not satisfy the requirements of measurement. Its quantitative outputs derive authority from institutional endorsement rather than empirical evaluation. This condition explains the persistence of composite utility constructs, cost-per-QALY ratios, and simulation-based evaluation despite their failure to satisfy representational measurement requirements. These constructs function effectively within the administrative logic of HTA. They provide standardized metrics. They enable procedural consistency. They facilitate comparative decision making. But their operational utility does not confer measurement validity.

The implications extend beyond methodological correctness. Measurement is the mechanism through which scientific knowledge evolves. Measurement allows claims to be falsified, replicated, and refined. Without measurement, numerical outputs cannot be empirically evaluated. They cannot be proven wrong because they do not correspond to measured quantities. They can be recalculated under different assumptions, but recalculation is not falsification. It is variation within a closed computational system.

The global logit profile demonstrates that HTA, as currently practiced, operates within such a closed and barren system. Its quantitative constructs are stabilized through institutional adoption rather than measurement validation. The appearance of quantification substitutes for quantification itself. The numbers produced by HTA function as administrative instruments rather than scientific measures. This inversion defines the structural condition that subsequent national and institutional analyses document in detail.

SECTION 3: FROM MEASUREMENT TO MEMEPLEX: HOW FALSE QUANTIFICATION BECAME INSTITUTIONALIZED

The global logit profile demonstrates not merely the absence of measurement axioms but their systematic replacement by an alternative operational logic. This logic does not emerge accidentally. It reflects the institutionalization of a numerical framework whose authority derives from procedural adoption rather than empirical validation. The result is the emergence, as described above, of a memeplex: a self-reinforcing system of quantitative belief that stabilizes itself across institutions, jurisdictions, and professional communities. Within this memeplex, numbers function as symbols of authority rather than representations of measured quantities. Their legitimacy is assumed rather than demonstrated. Their persistence reflects institutional reproduction rather than scientific validation.

The defining characteristic of a memeplex is self-reinforcement. Once numerical constructs such as the QALY and cost-effectiveness ratios become embedded within HTA guidelines, they acquire operational authority independent of their measurement properties. Agencies require their use. Journals publish analyses based upon them. Academic programs teach them as standard evaluative tools. Manufacturers construct submissions around them. Each institutional component reinforces the others. The constructs persist not because their measurement validity has been demonstrated, but because they are required for participation within the evaluative system. Their authority becomes circular. They are used because they are accepted, and they are accepted because they are used.

This institutional stabilization produces an epistemic closure. Measurement axioms cease to function as constraints because the evaluative framework does not require demonstration of measurement validity. Instead, quantitative constructs are accepted based on procedural conformity. Analysts demonstrate competence by correctly applying established scoring algorithms and simulation models. Agencies demonstrate rigor by enforcing standardized submission requirements. Journals demonstrate methodological sophistication by publishing increasingly complex models. None of these activities require demonstration that the numerical constructs involved satisfy representational measurement axioms. Measurement becomes irrelevant to quantitative practice.

The logit profile reveals the structural consequences of this closure. Statements asserting measurement axioms collapse to floor values because they do not operate as operational requirements. They are absent not because they are explicitly rejected, but because they are unnecessary within the memeplex. Quantitative practice proceeds without them. Arithmetic operations are performed because the evaluative framework requires numerical outputs, not because measurement validation has been established. The memeplex reproduces itself by requiring numerical conformity rather than empirical measurement.

The central constructs of HTA illustrate this process clearly. The QALY, derived from multiattribute utility instruments, combines time with preference-weighted scores representing multidimensional health state descriptions. These utility scores do not possess demonstrated ratio scale properties. They are constructed through valuation procedures rather than measurement transformation. Yet once embedded within HTA guidelines, they become operational decision

variables. Cost-per-QALY ratios derived from these constructs become determinants of reimbursement decisions, pricing negotiations, and therapy access. The numerical framework becomes institutionalized regardless of its measurement validity [8].

Simulation modeling further stabilizes the memeplex by producing precise numerical outputs derived from these constructs. These outputs appear quantitative. They can be expressed to multiple decimal places. They can be subjected to sensitivity analysis. They can be recalculated under alternative assumptions. This computational flexibility creates the appearance of scientific rigor. Yet the numerical precision of simulation outputs reflects internal model coherence rather than measurement validity. The models operate on numerical constructs whose scale properties have not been demonstrated. Their outputs are computationally precise but empirically undefined.

Institutional reproduction ensures the persistence of this framework. Academic programs train analysts to apply existing methods rather than question their measurement foundations. Journals publish studies that conform to established evaluative conventions. Agencies require submissions structured according to reference case guidelines. Each institutional component reinforces the memeplex. Alternative approaches grounded in representational measurement, linear ratio measurement for manifest attributes and Rasch logit ratio measurement for latent constructs, remain outside the operational framework. They are not rejected through empirical refutation. They are excluded through institutional irrelevance.

This exclusion is reflected directly in the logit profile. Statements asserting the necessity of Rasch transformation collapse to floor values because Rasch measurement does not function as an operational requirement within HTA. Statements asserting that arithmetic requires ratio measurement similarly collapse because arithmetic operations are performed independently of scale validation. The memeplex stabilizes itself by defining quantitative practice in terms of procedural conformity rather than measurement validity.

The result is a closed epistemic system. Within this system, numerical constructs acquire authority through institutional adoption rather than empirical correspondence. Their persistence does not depend on measurement validation. It depends on institutional reproduction. This condition explains the global convergence observed in the logit profiles. Different countries, agencies, and journals reproduce the same evaluative architecture not because they independently validate its measurement foundations, but because they adopt established quantitative conventions.

The transition from measurement to memeplex represents a fundamental shift in the epistemic status of quantitative claims within HTA. Measurement-based science derives authority from empirical evaluation. Memeplex-based systems derive authority from institutional stabilization with endless repetition. The global logit profile demonstrates that HTA operates within the latter condition. Its quantitative framework persists because it is institutionally required, not because it satisfies the axioms that make quantification possible.

SECTION 4: OPERATIONAL CONSEQUENCES: WHEN MEASUREMENT FAILURE GOVERNS PATIENT ACCESS AND RESOURCE ALLOCATION

The inversion of representational measurement axioms documented in the global logit profile is not a purely theoretical defect. It has direct operational consequences. Health technology assessment exists to inform decisions that affect patient access to therapy, physician treatment choices, and health system resource allocation. These decisions rely explicitly on quantitative claims. Therapies are approved, rejected, or restricted based on numerical estimates of incremental benefit and cost-effectiveness. Prices are negotiated based on cost-per-QALY ratios. Clinical access is conditioned on model projections of long-term outcomes. When the quantitative constructs underlying these decisions lack measurement validity, the consequences extend beyond methodological error. Measurement failure becomes a determinant of clinical reality.

The central mechanism through which this occurs is the substitution of composite scoring systems for measurement. Multiattribute utility instruments generate numerical scores based on preference-weighted descriptions of health states. These scores are treated as if they represent invariant quantities of health benefit. They are multiplied by time to generate QALYs. These QALYs are then combined with cost estimates to produce cost-effectiveness ratios. These ratios become the basis for reimbursement decisions. At each stage, arithmetic operations are performed on numerical constructs whose measurement properties have not been established. The resulting outputs appear quantitative. They produce precise numerical thresholds. They enable comparative ranking of therapies. Yet their numerical legitimacy rests on assumptions rather than meeting the axioms of representational measurement.

The logit profile demonstrates that this substitution is structural. Statements asserting the necessity of measurement prior to arithmetic collapse to floor values. Statements endorsing arithmetic manipulation of composite utility constructs register strong positive logits. This pattern confirms that the evaluative framework treats numerical scores as quantities without requiring demonstration that they represent measurable attributes. Arithmetic becomes an administrative procedure rather than an empirical operation.

This condition has direct implications for patient access. When reimbursement decisions are based on cost-per-QALY thresholds, therapies may be rejected because their modeled cost-effectiveness ratios exceed predefined limits. These ratios are derived from composite utility scores rather than measurement-valid quantities. The rejection of therapy is therefore determined by arithmetic operations performed on constructs whose empirical meaning has not been established. The numerical framework acquires operational authority independent of measurement validity. Patients may be denied access to therapy based on quantitative constructs that do not represent measured differences in therapeutic benefit.

Physicians face parallel constraints. Clinical decision-making is increasingly influenced by HTA recommendations derived from cost-effectiveness analysis. Treatment guidelines incorporate economic evaluation results. Access to certain therapies may require demonstration of cost-effectiveness within defined thresholds. Physicians are therefore required to operate within a

quantitative framework that assumes measurement validity without demonstrating it. Clinical judgment becomes subordinate to numerical constructs whose empirical basis remains undefined.

Resource allocation decisions within healthcare systems are similarly affected. HTA frameworks are designed to optimize resource use by comparing therapies based on quantitative estimates of benefit and cost. These comparisons assume that the underlying numerical constructs represent measurable quantities. When this assumption is false, resource allocation decisions become detached from empirical measurement. Resources are allocated based on numerical outputs whose quantitative meaning is assumed rather than demonstrated.

Simulation modeling amplifies these effects by projecting numerical estimates over extended time horizons. Models integrate clinical data, epidemiological assumptions, and utility scores to produce projected cost-effectiveness ratios. These projections are treated as if they represent future measurable outcomes. Yet their numerical outputs are functions of assumptions rather than measurements. They cannot be empirically falsified because they do not correspond to observable quantities. The logit profile confirms that simulation-based constructs are treated as legitimate evaluative objects despite lacking measurement validity.

The operational authority of these constructs is reinforced by institutional requirements. Agencies mandate their use. Submission guidelines specify their structure. Journal publications normalize their application. The memeplex reproduces itself operationally. Measurement validity is not required for institutional acceptance. Procedural conformity is sufficient.

This condition has profound implications for the duty of care owed to patients and physicians. Healthcare decisions must be grounded in empirical evidence. Quantitative claims used to justify access restrictions, pricing decisions, or resource allocation must represent measurable attributes. When arithmetic operations are performed on non-measures, the resulting numerical outputs lack empirical grounding. Decisions based on those outputs cannot be justified as measurement-based assessments of therapeutic benefit.

The logit evidence demonstrates that this condition is not exceptional. It is systemic. Measurement axioms do not function as operational constraints within HTA frameworks. Numerical constructs lacking ratio scale properties are treated as if they support arithmetic operations. Simulation outputs derived from these constructs are treated as empirical evidence. The quantitative framework of HTA governs clinical and economic decisions despite lacking measurement foundation.

The CHEERS 2022 guidance for acceptable reference case claims makes this structural priority unmistakable, not through what it requires, but through what it excludes. Nowhere does CHEERS 2022 require that outcome measures satisfy the axioms of representational measurement. Nowhere does it require demonstration that the constructs entering arithmetic operations possess unidimensionality, invariance, or ratio scale properties. Nowhere does it require that latent constructs be transformed through Rasch measurement to establish invariant interval or logit ratio scales prior to arithmetic manipulation. These omissions are not incidental. They define the epistemic boundary of the framework. CHEERS 2022 specifies how to report modeled cost-effectiveness claims, but it does not require demonstration that the numerical quantities being

reported constitute measures in the scientific sense. Measurement validity is treated as irrelevant to reporting validity. The result is a framework in which numerical claims can achieve formal legitimacy without satisfying the preconditions that make quantitative claims meaningful.

This exclusion has direct implications for duty of care. In any scientific discipline concerned with real-world consequences, quantitative claims must be grounded in measurement structures capable of falsification, replication, and empirical verification. This is not an optional refinement; it is the foundation of scientific accountability. Health technology assessment directly influences patient access to therapy, physician treatment options, and health system resource allocation. The acceptance of numerical claims that lack measurement validity therefore carries clinical and ethical consequences. Yet CHEERS 2022 does not require that submitted claims be empirically evaluable in a measurement-theoretic sense. Instead, it requires only that modeling assumptions, analytic structure, and sensitivity analyses be transparently reported. Transparency of procedure substitutes for validity of measurement. The framework ensures that claims can be documented, reviewed, and reproduced computationally, but not that they can be verified empirically as measurements of therapy impact.

Equally important, CHEERS 2022 interrupts the evolution of objective knowledge by privileging closure over falsification. In normal science, quantitative claims remain provisional. They are exposed to the risk of refutation through replication, measurement refinement, and empirical testing. Measurement-invalid constructs are replaced by constructs that satisfy representational criteria. Knowledge evolves through the elimination of error. The CHEERS 2022 framework, by contrast, is designed to produce definitive cost-effectiveness claims within a specified modeling architecture. Once constructed and reported according to the checklist, the claim achieves administrative finality. It becomes a decision object. There is no requirement for subsequent empirical evaluation of whether the modeled QALY gains correspond to measurable changes in patient status. There is no protocol-driven replication requirement grounded in measurement-valid endpoints. Closure replaces falsification as the operational goal.

This shift from falsification to closure defines the operational logic of the reference case memeplex. The purpose of the framework is not to discover measurable quantities but to produce administratively actionable numbers. Simulation outputs achieve legitimacy by conforming to reporting standards rather than by satisfying measurement axioms. Numerical precision replaces empirical validity as the criterion of acceptance. The checklist becomes the guarantor of legitimacy. In this environment, the absence of measurement validity does not function as a barrier to acceptance because the framework does not recognize measurement validity as a requirement.

The consequence is a self-stabilizing evaluative system. By defining acceptable reporting without requiring measurement validity, CHEERS 2022 ensures the continued reproduction of reference case simulation claims. Each accepted claim reinforces the legitimacy of the framework. Each published model contributes to the accumulation of internally coherent but measurement-ungrounded numerical outputs. The system evolves administratively while remaining epistemically static. It produces numbers capable of supporting decisions but incapable of supporting scientific knowledge.

In this sense, CHEERS 2022 is best understood not as a guide to scientific reporting but as a procedural mechanism for achieving epistemic closure with false measurement. Its function is to standardize the production of administratively acceptable claims, not to ensure that those claims satisfy the conditions required for measurement-based science. Duty of care, empirical falsifiability, and the evolution of objective knowledge fall outside its operational scope. What matters is that a cost-effectiveness claim can be constructed, documented, and accepted. Once closure is achieved, the system moves forward; not by correcting measurement error, but by reproducing it.

This condition transforms quantitative evaluation into administrative computation. Numbers retain their authority, but their empirical meaning is lost. The appearance of measurement persists, while measurement itself is absent. The consequences are operational, not merely theoretical. Measurement failure becomes embedded in the decision architecture governing patient access, physician choice, and health system resource allocation.

SECTION 5: DUTY OF CARE AND THE ARRESTED EVOLUTION OF OBJECTIVE KNOWLEDGE

The global inversion of representational measurement axioms documented by the canonical logit diagnostic has implications that extend beyond methodological correctness. It reaches the core ethical and scientific responsibilities of health technology assessment. HTA exists to inform decisions that directly affect patient welfare, physician practice, and the allocation of finite healthcare resources. These decisions are justified by reference to quantitative claims. The legitimacy of those claims depends entirely on whether they represent measurable quantities. When measurement axioms are absent, the quantitative framework ceases to function as an instrument of empirical knowledge. It becomes an administrative belief system. The resulting condition constitutes not merely a technical failure but a failure of duty of care.

Duty of care requires that decisions affecting patient access and treatment availability be grounded in empirically valid evidence. Quantitative claims used to support such decisions must correspond to measurable attributes. Measurement provides the mechanism through which claims can be tested, falsified, and refined. Without measurement, quantitative outputs cannot be empirically evaluated. They cannot be proven wrong because they do not represent quantities. Their persistence reflects institutional adoption rather than empirical validation. When HTA frameworks rely on such constructs to determine therapy access, pricing, or prioritization, the resulting decisions lack empirical justification.

The logit profile demonstrates that this condition is structural. Statements asserting the axioms required for measurement collapse to floor values, indicating that these principles do not function as operational constraints. Arithmetic operations are performed on composite utility scores whose scale properties have not been established. Cost-per-QALY ratios are constructed from numerators and denominators lacking dimensional homogeneity. Simulation outputs derived from these constructs are treated as evidence despite being functions of assumptions rather than measurements. The quantitative framework governing healthcare decisions operates independently of measurement validity.

This condition arrests the evolution of objective knowledge. Scientific progress depends on falsification [9]. Measurement-valid constructs produce numerical outputs that correspond to empirical quantities. These outputs can be tested against observation. Errors can be identified. Constructs can be refined or replaced. Knowledge evolves through this process. Measurement-invalid constructs do not support falsification. Their numerical outputs cannot be empirically tested because they do not represent measurable attributes. They can be recalculated under alternative assumptions, but recalculation is not empirical testing. It is variation within a closed computational system. The absence of measurement prevents the self-correcting mechanism that defines scientific progress.

The institutionalization of measurement-invalid constructs therefore creates epistemic closure. Quantitative claims acquire authority through procedural adoption rather than empirical validation. Their persistence reflects institutional reproduction rather than scientific testing. The memeplex stabilizes itself by defining quantitative practice in terms of procedural conformity. Measurement axioms become irrelevant because they are not required for participation within the evaluative

framework. This condition explains the global convergence observed in the logit profiles. Countries, agencies, and journals reproduce the same evaluative architecture not because they independently validate its measurement foundations, but because they adopt established conventions. The memeplex is entrenched.

The consequences extend directly to patient care. Decisions governing therapy access and reimbursement are justified by reference to numerical constructs lacking measurement validity. Patients may be denied access to therapies based on cost-effectiveness ratios derived from composite utility scores rather than measurement-valid quantities. Physicians must operate within treatment frameworks shaped by these constructs. Resource allocation decisions reflect arithmetic operations performed on non-measures. The quantitative framework acquires operational authority despite lacking empirical grounding.

This condition is incompatible with the scientific and ethical responsibilities of healthcare decision-making. Duty of care requires that quantitative claims reflect measurable attributes. Decisions affecting patient welfare must be grounded in empirical reality. Measurement provides the only mechanism through which quantitative claims can achieve this grounding. Without measurement, numerical outputs become administrative artifacts rather than empirical discoveries.

The logit evidence demonstrates that HTA, as currently practiced, operates within a quantitative framework detached from representational measurement axioms. The resulting system retains procedural sophistication but lacks measurement foundation. It produces numbers without measurement. It performs arithmetic without quantities. It generates quantitative claims that cannot be empirically falsified. This condition arrests the evolution of objective knowledge within HTA.

Recovery requires structural reconstruction. Quantitative evaluation must be grounded in measurement-valid constructs. Manifest attributes must be measured using linear ratio scales. Latent constructs must be measured using Rasch transformation to establish invariant logit ratio scales. Arithmetic operations must be restricted to constructs satisfying representational measurement axioms. Only within such a framework can quantitative claims support falsification, replication, and cumulative knowledge development.

Until this reconstruction occurs, the global HTA memeplex will continue to reproduce pseudo-quantitative outputs detached from empirical measurement. Its numerical authority will remain institutional rather than scientific. Its decisions will reflect administrative computation rather than measurement-based evaluation. And its quantitative framework will remain incompatible with the duty of care owed to patients, physicians, and healthcare systems.

SECTION 6: RECONSTRUCTION: REPLACING THE MEMEPLEX WITH MEASUREMENT-BASED EVALUATION

The global logit evidence establishes a structural conclusion: the quantitative architecture of health technology assessment does not satisfy the axioms required for measurement. Arithmetic operations are performed on constructs that do not possess ratio scale properties. Composite utility scores are treated as quantities despite lacking invariant unit structure. Simulation outputs are treated as evidence despite being functions of assumptions rather than measurements. These findings do not identify isolated methodological defects. They identify a structural condition. The evaluative framework itself operates independently of representational measurement constraints. Under such conditions, reform cannot consist of incremental refinement. It requires reconstruction. The objective is not to improve the existing architecture but to replace it with one grounded in measurement.

Measurement-based evaluation begins with a fundamental principle: arithmetic operations are admissible only when performed on quantities that satisfy representational measurement axioms. This requirement applies equally to manifest and latent attributes. Manifest attributes, such as survival time, hospitalizations, exacerbations, and resource utilization events, can be measured directly on linear ratio scales. These attributes possess natural zero points and invariant units. Arithmetic operations performed on these measures preserve empirical meaning. Differences and ratios correspond to observable differences and ratios in the attribute itself. Quantitative claims derived from such measures can be empirically tested, falsified, and replicated.

Latent attributes, such as symptom burden, functional limitation, or need fulfillment, require transformation before arithmetic operations become admissible. Ordinal observations derived from questionnaires or clinical ratings do not possess interval or ratio scale properties. They represent ordered categories, not quantities. Rasch measurement provides the lawful transformation mechanism. By converting ordinal responses into invariant logit measures, Rasch transformation establishes linear interval structure grounded in empirical observation. These logit measures support arithmetic operations that preserve empirical meaning. They allow latent constructs to be quantified without violating representational measurement axioms.

The distinction between measurement and scoring is decisive. Measurement produces quantities that exist independently of the scoring system. Scoring produces numbers defined entirely by the scoring algorithm. Measurement-valid constructs can be empirically evaluated. Scoring-derived constructs cannot. Reconstruction of HTA requires replacing scoring-based constructs with measurement-based constructs. Composite utility indices must be abandoned because they aggregate multidimensional attributes into single numerical scores without establishing unidimensional measurement structure. Cost-per-QALY ratios must be abandoned because their denominators lack ratio scale properties. Simulation outputs based on these constructs must be abandoned because they do not correspond to measured quantities.

The alternative is a portfolio of single-attribute value claims grounded in measurement. Each claim must correspond to a specific attribute measured on a valid scale. Manifest attributes must be measured on linear ratio scales. Latent attributes must be measured using Rasch transformation to establish invariant logit ratio scaling. Each claim must be accompanied by a protocol specifying

the target population, measurement instrument, observation period, and empirical evaluation procedure. Claims must be structured to allow falsification. They must produce quantitative outputs that can be empirically tested and replicated.

This reconstruction restores the scientific foundation of HTA. Measurement-valid claims support cumulative knowledge development. Empirical testing allows incorrect claims to be identified and replaced. Quantitative evaluation becomes a process of discovery rather than administrative computation. Numerical outputs regain empirical meaning. Arithmetic operations correspond to measurable differences in therapeutic impact.

The transition from memeplex to measurement-based evaluation also restores the ethical foundation of HTA. Duty of care requires that decisions affecting patient access and treatment availability be grounded in empirically valid evidence. Measurement-valid constructs provide this grounding. They ensure that quantitative claims reflect measurable differences in patient outcomes rather than numerical artifacts of scoring systems. Physicians can rely on measurement-based claims to inform treatment decisions. Patients can trust that access decisions reflect measurable therapeutic benefit rather than administrative convention.

The global logit evidence demonstrates that this reconstruction has not yet occurred. The memeplex persists because it is institutionally stabilized. Its constructs remain embedded in guidelines, journals, and decision frameworks. But institutional stabilization cannot substitute for measurement validity. The scientific status of HTA depends on its adherence to representational measurement axioms. Without measurement, quantitative evaluation lacks empirical foundation. With measurement, quantitative evaluation regains its role as an instrument of scientific discovery.

As far as HTA is concerned, failure to reconstruct its measurement foundations beckons a future that may extend for decades, sustained by institutional inertia and reinforced by global diffusion. In that future, one closed reference case simulation will follow another in an unbroken sequence. Each will be constructed with increasing technical sophistication, incorporating more parameters, more sensitivity analyses, and more elaborate probabilistic structures. Yet none will resolve the foundational defect that defines them all: the absence of measurement. The result will be a vast and expanding archive of internally coherent but empirically untestable numerical constructions. Their authority will derive not from correspondence with measurable reality, but from conformity with procedural expectations. This is not the evolution of knowledge. It is the perpetuation of a closed system.

For those who subscribe to the memeplex, this trajectory offers a peculiar form of professional stability. It promises a career devoted to refinement without discovery, elaboration without falsification, and publication without the risk of being wrong. Each successive model will differ from its predecessor only in structure, assumptions, or data inputs, never in epistemic status. The core construct, the cost-effectiveness ratio built upon composite utility scores, will remain untouched, insulated from challenge by the very framework that requires its use. The intellectual horizon becomes bounded. There are no decisive experiments to conduct, no measurement transformations to develop, and no empirical contradictions to resolve. The field becomes self-referential, its outputs judged by internal standards rather than by their capacity to measure therapeutic impact.

This future is, in a profound sense, intellectually uninteresting. Scientific disciplines derive their vitality from the possibility of error and the prospect of correction. Progress occurs when invalid constructs are identified and replaced by measurement-valid successors. Within the memeplex, this mechanism is absent. There is nothing to overturn because nothing is exposed to falsification. Numerical claims may be recalculated, updated, or refined, but they cannot be empirically refuted. As a result, the central questions that animate scientific inquiry—What is being measured? How accurately? Under what conditions can the claim be proven wrong?—never arise. The field becomes administratively productive but scientifically static.

The consequences extend beyond intellectual stagnation. A framework that does not measure cannot generate cumulative knowledge. Each model exists as an isolated exercise, dependent on its own assumptions and valuation conventions. There is no invariant unit structure to support comparison across studies or over time. The apparent accumulation of evidence is an illusion created by the multiplication of simulations rather than the replication of measurements. The literature grows, but knowledge does not. The appearance of quantitative sophistication conceals an absence of empirical advancement.

For many observers, this prospect will hold little appeal. The commitment to pseudo-measurement forecloses the very possibility that makes scientific work meaningful: the opportunity to discover something that was previously unknown. Instead, it offers only the assurance that one may participate indefinitely in a system whose outputs are predetermined by its assumptions. The intellectual challenge shifts from understanding reality to maintaining coherence within the framework. This is not the future of a scientific discipline. It is the future of an administrative belief system sustained by numerical ritual rather than empirical inquiry with the memeplex offering the intellectual challenge of peeling potatoes.

The alternative is reconstruction. Reconstruction restores the conditions under which measurement becomes possible, arithmetic becomes admissible, and claims become empirically evaluable. Without that reconstruction, HTA faces not collapse, but something more insidious: indefinite continuation in a state of epistemic suspension, producing numbers that command authority without possessing meaning, and sustaining careers devoted not to the evolution of objective knowledge, but to its indefinite postponement.

CONCLUSION

The global logit evidence leaves no ambiguity. Health technology assessment, as presently constituted, operates within a quantitative framework detached from the axioms that make measurement possible. Across countries, agencies, journals, and evaluative instruments, the same structural pattern emerges. Statements expressing the foundational requirements of representational measurement unidimensionality, invariance, admissible scale transformations, and the necessity that measurement precede arithmetic collapse to floor logit values or adjacent negative ranges. These results demonstrate effective non-possession. The axioms do not function as binding constraints. They do not govern quantitative practice. Instead, arithmetic operations are performed on constructs whose scale properties do not support arithmetic. Composite utility scores are multiplied, aggregated, and compared. Simulation models produce precise numerical outputs. Cost-effectiveness ratios are constructed and used to justify decisions affecting therapy access and pricing. The numerical apparatus is elaborate, internally coherent, and globally diffused. But it is not grounded in measurement.

This condition defines the global HTA memeplex. Its authority derives from institutional reproduction rather than empirical validation. Its constructs persist because they are required, not because they measure. Its numerical outputs acquire administrative legitimacy independent of empirical grounding. Within this system, numbers function as instruments of procedural justification rather than empirical discovery. They create the appearance of quantification while bypassing the structural conditions required for quantification itself. The memeplex is stabilized through guidelines, journals, training programs, and operational agencies. Each institutional component reinforces the others. Measurement axioms are not refuted. They are simply excluded from operational relevance.

The consequences extend beyond epistemology. HTA governs decisions that affect patient access, physician choice, and health system resource allocation. These decisions are justified by quantitative claims. When those claims are not grounded in measurement, their numerical authority lacks empirical legitimacy. Duty of care requires that decisions affecting patient welfare be grounded in empirically valid evidence. Measurement provides the mechanism through which quantitative claims acquire empirical meaning. Without measurement, arithmetic operations lose their interpretability. Numerical outputs become administrative artifacts. Decisions based on those outputs cannot claim scientific foundation.

Equally significant is the impact on the evolution of objective knowledge. Scientific progress depends on falsification. Measurement-valid constructs produce quantitative claims that can be empirically tested and, if incorrect, replaced. This process allows knowledge to evolve. Measurement-invalid constructs do not support falsification. Their outputs cannot be empirically tested because they do not represent measurable quantities. They can be recalculated but not refuted. The memeplex therefore arrests the evolution of knowledge. It produces numerical stability without empirical progress. It refines procedures while leaving measurement absence untouched.

The canonical logit diagnostic provides empirical confirmation of this condition. It demonstrates that the inversion of representational measurement axioms is not local or accidental. It is global. It

is structural. It is reproduced across national HTA systems, operational agencies, academic journals, and utility instrument frameworks. This convergence identifies the memeplex as an institutionalized evaluative architecture rather than a collection of isolated methodological choices.

Recovery requires reconstruction, not refinement. HTA must abandon constructs that do not satisfy measurement axioms. Manifest attributes must be measured on linear ratio scales. Latent attributes must be measured using Rasch transformation to establish invariant logit ratio scales. Quantitative claims must correspond to measurable quantities. Arithmetic operations must be restricted to constructs satisfying representational measurement requirements. Only within such a framework can quantitative evaluation regain empirical meaning.

The choice facing HTA is therefore not between alternative modeling techniques or methodological refinements. It is between two fundamentally different epistemic foundations. One treats numbers as administrative instruments, stabilized through institutional reproduction. The other treats numbers as representations of measurable quantities, grounded in representational measurement axioms and exposed to empirical testing. Only the latter satisfies the conditions of science. Only the latter fulfills the duty of care owed to patients, physicians, and healthcare systems. Only the latter allows the evolution of objective knowledge.

ACKNOWLEDGEMENT

I acknowledge that I have used OpenAI technologies, including the large language model, to assist in the development of this work. All final decisions, interpretations, and responsibilities for the content rest solely with me.

REFERENCES

[1] Krantz D, Luce R, Suppes P, Tversky A. Foundations pf Measurement Vol 1: Additive and Polynomial Representations. New York: Academic Press, 1971

[2] Rasch G, Probabilistic Models for some Intelligence and Attainment Tests. Chicago: University of Chicago Press, 1980 [An edited version of the original 1960 publication]

[3] Dawkins R. The Selfish Gene. 30th Anniversary Edition. Oxford: Oxford University Press, 2006

[4] Husereau D, Drummond M, Augustovski F et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. *Value in Health* 25 (1): 3–9

[5] Stevens S. On the Theory of Scales of Measurement. *Science*. 1946;103(2684):677-80

[6] Popper K. Objective Knowledge: An Evolutionary Approach. Revised edition. Oxford: Oxford University Press, 1979

[7] Bond T, Z Yan, Heene M. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. 4th ed. New York: Routledge, 2021

[8] Drummond M, Sculpher M, Claxton K et al. Methods for the Economic Evaluation of Health Care Programmes (4th Ed). New York, Oxford University Press, 2015

[9] Pigliucci M. Nonsense on Stilts: How to tell science from bunk. Chicago: University of Chicago Press, 2010