MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

UNITED STATES: PHARMACY EDUCATION, MEASUREMENT, DUTY OF CARE AND THE FAILURE TO TRAIN FOR EVIDENCE

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

LOGIT WORKING PAPER No 1235 8 APRIL 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

This paper examines the failure of pharmacy education to meet its duty of care in the preparation of graduates for roles in health technology assessment (HTA), pharmacoeconomics, and outcomes research. The central claim is that current training frameworks do not equip students to determine whether the claims they construct or interpret satisfy the axioms of representational measurement. As a consequence, graduates are proficient in the application of analytical techniques particularly cost-effectiveness modelling, utility-based evaluation, and reference case analysis but are not trained to assess whether the underlying constructs support admissible arithmetic operations.

The analysis draws on a structured interrogation of HTA-related knowledge bases associated with U.S. pharmacy schools using a 24-item canonical statement framework implemented through large language models (LLMs). Each statement reflects a necessary condition for measurement, including unidimensionality, dimensional homogeneity, and the requirement that measurement must precede arithmetic. Across more than 120 global assessments, a consistent pattern emerges. Statements aligned with measurement theory receive low endorsement, while constructs central to current HTA practice, utilities, QALYs, and simulation-based reference case outputs are treated as if they possess the properties required for arithmetic operations. This asymmetry is identified as measurement inversion.

The implications are direct. Claims derived from non-measurable constructs cannot be empirically evaluated, replicated, or falsified. Yet these claims are routinely used to support decisions affecting therapy access, pricing, and resource allocation. This constitutes a breach of duty of care. The paper argues that the reference case framework is not a competing analytical approach but a structure that excludes the possibility of measurement and must therefore be abandoned.

A transition is proposed based on single-attribute claims supported by linear ratio measures for manifest attributes and Rasch-based logit ratio scales for latent constructs. This reconstitution is presented as a necessary condition for restoring HTA to the standards of normal science and for ensuring that pharmacy education fulfills its professional obligations.

SECTION I: DUTY OF CARE AND THE OBLIGATION TO EVIDENCE

Pharmacy education cannot avoid the question of duty of care. It is not confined to dispensing accuracy or clinical competence at the bedside. It extends to the preparation of graduates who participate directly in decision-making frameworks that determine therapy access, pricing, and resource allocation. Increasingly, pharmacists and pharmacy-trained researchers operate within health technology assessment (HTA), pharmacoeconomics, and outcomes research environments where their analyses influence physicians, patients, and health systems. In this context, duty of care must include the obligation to ensure that the claims graduates are trained to construct and interpret are scientifically valid.

The issue is immediate. If a graduate is equipped to generate cost-effectiveness claims, utility-based evaluations, or modelled projections of therapy impact, then the institution that trained them

has an obligation to ensure that these claims meet the standards required for evidence. That obligation cannot be discharged by teaching technique alone. It requires that the underlying constructs satisfy the axioms of representational measurement. Without this, the appearance of analytical sophistication masks a fundamental absence: the inability to determine whether the outputs produced have meaning.

Across pharmacy education, this obligation has not been met. Students are trained in the mechanics of HTA involving how to build models, calculate incremental cost-effectiveness ratios, and interpret QALYs but they are not trained to interrogate whether these operations are admissible. They are not equipped with forensic measurement tools to determine whether the attributes being combined are measurable, whether the scales employed support arithmetic, or whether the resulting claims can be empirically evaluated. This is not a minor curricular gap. It is a structural failure in professional preparation.

The consequence is that pharmacy schools certify competence in methods that can generate numerical outputs but cannot guarantee that those outputs constitute evidence. Once this is recognized, the question of duty of care becomes unavoidable. If graduates are not trained to distinguish between measurable and non-measurable claims, then institutions are placing them and those affected by their decisions at risk. The failure is not theoretical. It is embedded in the current routine practice of HTA and pharmacoeconomics, where decisions with direct implications for patient care are supported by claims that have not met the minimum standards for measurement.

SECTION II: MEASUREMENT, ARITHMETIC AND THE CONDITIONS FOR SCIENCE

The requirements for representational measurement are not optional refinements. They are preconditions for the application of arithmetic and, by extension, for the construction of evaluable claims. Measurement must precede arithmetic. Attributes must be unidimensional. Dimensional homogeneity must be preserved. Ratio scales require a true zero. These conditions are not matters of convention; they are necessary for numerical operations to have meaning.

Within pharmacy education, these requirements are largely absent from the methodological framework presented to students. Instead, arithmetic is introduced as if it were inherently valid once numerical values are assigned. Preference scores are treated as if they can be multiplied by time. Composite cost measures are treated as if they represent single attributes. Aggregations of disparate dimensions are assumed to produce meaningful totals. The prior question—whether the constructs involved satisfy the axioms required for these operations—is not addressed.

This inversion has direct implications. If a scale is ordinal, it cannot support multiplication or division. If an attribute is multidimensional, it cannot be collapsed into a single measure without violating dimensional homogeneity. If a latent construct has not been transformed through an invariant measurement model, it remains a classification, not a measure. Arithmetic applied under these conditions is inadmissible. The resulting outputs are not approximations. They are undefined.

The only defensible pathway for latent constructs is through Rasch-based conjoint measurement, where item difficulty and person ability are jointly estimated to produce an invariant logit scale.

For manifest attributes, measurement must be anchored in linear ratio scales with a true zero. These two forms—linear ratio and Rasch logit ratio—are sufficient. They are also necessary. Without them, there is no basis for arithmetic, no basis for comparison, and no basis for empirical evaluation.

The failure to embed these principles within pharmacy education has created a situation where graduates operate within a numerical framework that appears scientific but lacks the conditions required for science. The distinction is critical. Science requires falsifiability, replication, and the accumulation of objective knowledge. These depend on measurement. Where measurement is absent, these conditions cannot be met, regardless of the sophistication of the analytical tools employed.

SECTION III: LLM INTERROGATION AND THE EMPIRICAL IDENTIFICATION OF MEASUREMENT FAILURE

The failure of pharmacy education to address the requirements of representational measurement has long been present, but until recently it has remained diffuse, embedded within curricula, publications, and methodological conventions without a systematic means of interrogation. Critiques have been advanced, inconsistencies identified, and limitations acknowledged, but these have typically been treated as isolated concerns rather than as evidence of a coherent structural problem. It is only with the application of large language models (LLMs) to defined knowledge bases that the scale and consistency of this failure has become empirically visible.

The LLM interrogation framework applies a fixed set of canonical statements derived from the axioms of representational measurement. Each statement represents a necessary condition for admissible arithmetic and evaluable claims. By interrogating a defined knowledge base—comprising teaching materials, methodological texts, guidelines, and the broader corpus that informs professional practice—the model generates categorical endorsement probabilities for each statement. These probabilities are then transformed into normalized logits, producing a profile that reflects how the knowledge base behaves with respect to the requirements of measurement.

Across interrogations of pharmacy colleges and schools in the United States, a consistent pattern emerges. Statements that reflect the foundational requirements of measurement—measurement preceding arithmetic, the necessity of unidimensionality, dimensional homogeneity, and the requirement for ratio scales in multiplication—receive uniformly low endorsement. In contrast, statements aligned with the operational framework of HTA—particularly those associated with utilities, QALYs, and reference case modelling—are effectively supported in application, despite their incompatibility with those same axioms.

This pattern is not incidental. It is stable across institutions, reproducible across interrogations, and consistent with findings from HTA agencies, academic research centers, and methodological publications. The resulting logit profiles exhibit a clear polarity: strong negative values associated with measurement requirements and positive reinforcement of practices that presuppose those requirements without satisfying them. This is the empirical signature of measurement inversion.

The importance of this result lies not in the authority of the LLM, but in the structure of the interrogation. The model does not introduce new theory. It applies a fixed and transparent set of criteria to a defined corpus. The output is therefore not an opinion, but a structured representation of how the knowledge base responds to those criteria. Where the knowledge base is internally consistent with measurement theory, endorsement probabilities would align with the axioms. They do not. Instead, they reveal a systematic divergence between what is required for measurement and what is practiced.

This divergence has direct implications for pharmacy education. The knowledge base that informs curricula and training is not neutral. It embodies the same inversion identified in the broader HTA environment. Students are exposed to methods and constructs that assume the validity of arithmetic operations without demonstrating that the underlying attributes are measurable. The LLM interrogation makes this visible in a way that was previously difficult to demonstrate. It shows that the issue is not one of individual misunderstanding or isolated curricular gaps, but of a structured and pervasive absence of measurement standards.

It is important to emphasize what this does and does not establish. The interrogation does not claim that faculty are unaware of measurement theory, nor does it suggest that all aspects of pharmacy education are flawed. What it demonstrates is that, at the level of the knowledge base that supports HTA and pharmacoeconomics, the conditions required for evaluable claims are not met. This is sufficient to raise a question of professional responsibility. If the methods being taught cannot produce claims that satisfy the axioms of measurement, then the training provided does not equip graduates to generate evidence in the scientific sense.

The application of LLM interrogation therefore marks a shift in the discussion. The issue is no longer whether measurement concerns can be raised in principle. It is whether a consistent, reproducible pattern of non-compliance with measurement standards can be ignored once it has been demonstrated. The answer, in the context of duty of care, is no. Once the failure is visible, it becomes actionable. Institutions can no longer rely on the assumption that established practice is sufficient. They must either demonstrate that their methods satisfy the required axioms or accept that the claims produced under current training frameworks are not evaluable.

This places pharmacy education at a point of decision. The LLM results do not compel a particular solution, but they remove the possibility of inaction. They show that the problem is not marginal, not isolated, and not confined to a subset of institutions. It is systemic. As such, it demands a systemic response. The subsequent sections address the nature of that response and the implications for the continued use of the reference case framework within pharmacy education and HTA practice.

SECTION IV: THE REFERENCE CASE AND THE INSTITUTIONALIZATION OF MEASUREMENT FAILURE

The contemporary HTA framework, centered on utilities, QALYs, and reference case modelling, represents the institutionalization of this measurement failure. It is not a competing methodological approach that can be evaluated alongside alternatives. It is a framework that excludes the possibility of measurement at each stage of its construction.

Utility instruments are built from composite health state descriptions that lack unidimensionality. Preference algorithms transform ordinal responses into numerical scores without establishing interval or ratio properties. These scores are then multiplied by time to create QALYs, assuming ratio-scale behavior that has not been demonstrated. Costs are aggregated from heterogeneous resource inputs, violating dimensional homogeneity. Simulation models combine these elements to produce projections that cannot be empirically tested or falsified. Each step denies the axioms required for measurement. The final output—cost-effectiveness ratios—rests on arithmetic operations that are inadmissible.

The persistence of this framework within pharmacy education is not due to its validity, but to its institutional acceptance. It provides a standardized template that aligns with regulatory expectations and policy processes. It generates numbers that can be compared, thresholds that can be applied, and recommendations that can be justified within existing decision structures. But this coherence is internal. It does not extend to the requirements of measurement.

This is the defining contradiction. Pharmacy schools teach evidence-based practice while endorsing a framework that cannot generate evidence. They promote scientific rigor while training students in methods that cannot satisfy the conditions for scientific evaluation. The result is a professional culture in which numerical outputs are accepted as evidence by convention, not because they meet the standards required for evidence.

The implications for duty of care are direct. When graduates use these frameworks to support decisions on therapy access, pricing, or clinical guidelines, they do so with tools that cannot produce evaluable claims. Patients and physicians are therefore affected by decisions grounded in constructs that lack measurement validity. This is not a technical limitation. It is a systemic failure with real-world consequences.

SECTION V: RECONSTITUTION AND THE RESTORATION OF PROFESSIONAL RESPONSIBILITY

The response to this failure cannot be incremental adjustment. It requires a reconstitution of pharmacy education around the principles of representational measurement. The objective is not to refine existing models, but to replace them with a framework that can support evaluable claims.

This begins with the identification of single attributes for therapy impact. Each claim must be anchored in a clearly defined, unidimensional attribute, linked to a target population and a specified timeframe. For manifest attributes, measurement must be in linear ratio terms, ensuring dimensional homogeneity and admissible arithmetic operations. For latent attributes, measurement must be constructed through Rasch models to produce invariant logit scales. These measures provide the only defensible basis for comparison, aggregation, and inference.

Claims must be supported by protocols that allow empirical evaluation. This includes specification of measurement instruments, data sources, and analytical methods prior to assessment. Outcomes must be replicable and subject to refutation. Simulation models, where used, must be subordinated to empirical validation rather than presented as evidence in their own right.

Pharmacy education must therefore shift from teaching how to apply existing false HTA frameworks to teaching how to determine whether a claim is measurable. This requires the integration of measurement theory into curricula, the development of tools for forensic assessment of claims, and the rejection of constructs that do not meet the required standards. Graduates must be able to identify when arithmetic is inadmissible, when a scale lacks the necessary properties, and when a claim cannot be evaluated.

This is not an optional enhancement. It is a requirement for meeting duty of care. Without it, pharmacy schools will continue to produce graduates capable of generating numerical outputs but not of determining whether those outputs have meaning. In a field where decisions directly affect patient outcomes and healthcare resource allocation, this is untenable.

The transition to measurement-based HTA is therefore not a matter of preference. It is a professional obligation. Once the axioms of representational measurement are recognized, the reference case framework cannot be sustained. The only path forward is the adoption of claims that are measurable, testable, and replicable. Anything less perpetuates a system that produces the appearance of evidence while denying its substance.

CONCLUSION

The argument presented is not a critique of method preference, but a demonstration of structural failure. Pharmacy education, in its current form, does not meet its duty of care in the domain of health technology assessment. Graduates are trained to generate and interpret claims that are presented as evidence, yet they are not equipped to determine whether those claims satisfy the axioms required for measurement. The consequence is unavoidable. Where measurement is absent, arithmetic is inadmissible; where arithmetic is inadmissible, claims cannot be evaluated, replicated, or falsified. They do not constitute evidence.

The application of LLM-based interrogation has made this failure visible. Across knowledge bases associated with U.S. pharmacy schools, the same pattern emerges: low endorsement of measurement requirements and implicit acceptance of constructs that presuppose those requirements without satisfying them. This is measurement inversion. It is not a marginal deficiency but a defining characteristic of the HTA framework as currently taught and applied.

The implications extend beyond academic practice. Decisions on therapy access, pricing, and resource allocation are influenced by claims derived from this framework. If those claims are not measurable, then those decisions are not evidence-based. This represents a breach of professional responsibility with direct consequences for patients, physicians, and health systems.

There is no pathway to recovery within the existing reference case framework. It cannot be refined to meet the axioms of representational measurement because it is constructed in violation of them. The only viable course is reconstitution. Pharmacy education must be aligned with a measurement-based framework in which claims are defined by single attributes, measured on admissible scales, and subject to empirical evaluation.

This is not optional. It is the minimum condition for restoring the integrity of HTA and for ensuring that pharmacy education fulfills its duty of care.

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.

MEASUREMENT: EVIDENCE AND VALUE

Maimon Research has developed three distance education programs to support the transition to a new paradigm in HTA. These comprise 12 module senior level program that details the standards for measurement, the failure of current HTA standards and the basis for protocol supported claims assessment for ratio measures of manifest attributes and Rasch logic ratio logit measures for latent attributes. The two other programs are only 5 modules but are designed to complement the 12-module program, for measurement axioms and Rasch attribute possession.

MAIMON RESEARCH LLC

DISTANCE EDUCATION PROGRAMS IN THE THEORY OF MEASUREMENT

Three programs are available: two short 5-module programs and a 12-module program that is structured as a senior level course on the transition from the current HTA belief system to a new paradigm for HTA

The two short programs are (i) **NUMERICAL STORYTELLING: SYSTEMATIC MEASUREMENT FAILURE IN HEALTH TECHNOLOGY ASSESSMENT** and (ii) A **NEW START IN MEASUREMENT FOR HEALTH TECHNOLOGY ASSESSMENT.** They are designed to complement the 12-module course program. They can be accessed through the **DISTANCE EDUCATION** section of the website with URL https://maimonresearch.com/distance-education-programs/

The senior level course **HEALTH TECHNOLOGY ASSESSMENT REBUILT: EVIDENCE AND VALUE** is accessed through the EVIDENCE AND VALUE section of the website or URL link https://maimonresearch.com/evidence-and-value/.

Together, these programs equip health systems, committees, and analysts with the competence required to enforce measurement standards consistently. Training does not replace judgment; it enables it. Without such preparation, the transition to meaningful measurement cannot be sustained. With it, formulary decision making can finally rest on claims that are not merely numerical, but measurable