MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

UNITED STATES: MEASUREMENT FAILURE IN HEALTH TECHNOLOGY ASSESSMENT—EVIDENCE FROM 72 COLLEGES AND SCHOOLS OF PHARMACY

The Institutionalization of Measurement Inversion in PharmD Education

Paul C Langley PhD Adjunct Professor, College of Pharmacy, University of Minnesota, Minneapolis, MN

Logit Working Paper No 777 May 2026

www.maimonresearch.com

Tucson AZ

UNITED STATES: MEASUREMENT FAILURE IN HEALTH TECHNOLOGY ASSESSMENT—EVIDENCE FROM 72 COLLEGES AND SCHOOLS OF PHARMACY

The Institutionalization of Measurement Inversion in PharmD Education

ABSTRACT

Health Technology Assessment (HTA) in the United States is widely presented as a quantitative, evidence-based discipline. This paper challenges that characterization through a structured interrogation of HTA-related knowledge bases across 72 colleges and schools of pharmacy. The findings demonstrate a consistent absence of adherence to the axioms of representational measurement. In particular, there is no evidence that attributes are defined and measured according to scale requirements, nor that the conditions necessary for arithmetic operations are recognized.

The consequence is the systematic presence of measurement inversion, where numerical manipulation precedes the establishment of measurement. Ordinal and multi-attribute constructs are routinely treated as if they were interval or ratio measures, and combined through arithmetic operations that lack dimensional coherence. Outputs derived from these procedures cannot be interpreted as quantities and do not support falsifiable claims.

This pattern is not localized but institutionalized, reflected in teaching, research, and decision frameworks. The result is a closed analytical system that produces numerical outputs without a basis in measurement, limiting the capacity for replication and the evolution of objective knowledge.

A transition to a measurement-based HTA framework is therefore required. This paper outlines a simplified alternative based on two admissible forms of measurement: linear ratio measures for manifest attributes and Rasch logit-based measures for latent attributes. The implications for curriculum design and faculty training are discussed.

INTRODUCTION

Health Technology Assessment (HTA) in the United States is taught, practiced, and applied as if it were a quantitative science. Students are trained to construct models, estimate costs and outcomes, and interpret results expressed in numerical form. These activities carry the implicit assumption that the values being manipulated are measures**;** that they represent quantities defined on scales that support arithmetic operations. This assumption is not examined. It is simply taken as given.

The evidence presented here indicates that this assumption cannot be sustained. Across 72 colleges and schools of pharmacy, there is no indication that the foundational requirements of representational measurement are recognized or applied. Attributes are not defined in a manner consistent with measurement theory; scale properties are not established; and arithmetic operations are applied without regard to dimensional homogeneity or invariance. The result is not a partial failure or an isolated oversight, but a systematic inversion of the measurement process. Numbers are generated and manipulated in advance of measurement, and the outputs are interpreted as if measurement had already been achieved. This is measurement inversion. In any other scientific field, this would signal a halt. In HTA, it has been normalized.

CURRENT HTA FRAMEWORK IN US PHARMD PROGRAMS IN COLLEGES AND SCHOOLS

In colleges and schools of pharmacy in the United States, Health Technology Assessment (HTA) is not typically presented as a standalone discipline but is embedded within pharmacoeconomics and outcomes research teaching. Across accredited PharmD programs, pharmacoeconomic content is widely incorporated, with studies showing that almost all programs include at least some formal instruction, although the depth and structure vary considerably . Typically, this content is delivered through required courses in the professional years, often combined with pharmacoepidemiology and healthcare systems.

The core of the HTA framework taught to students centers on economic evaluation methods. These include cost-minimization, cost-effectiveness, cost-utility, and cost-benefit analyses, alongside instruction in decision-analytic modeling and the interpretation of published economic evaluations . Students are trained to assess comparative value, estimate costs and outcomes, and interpret measures such as incremental cost-effectiveness ratios. A dominant outcome metric is **the** quality-adjusted life year (QALY), which is widely used in cost-utility analysis and policy decision-making .

In addition, students are taught to evaluate health-related quality of life measures, review literature, and apply economic reasoning to formulary and reimbursement decisions . These topics are often supported by case studies, project work, and experiential learning components, where students apply methods to simulated or real-world decision contexts. The broader PharmD curriculum integrates these topics with clinical pharmacotherapy, biostatistics, and healthcare policy, reinforcing their role in decision-making within healthcare systems .

However, while the framework is presented as quantitative and evidence-based, it is primarily method-driven rather than measurement-driven. Instruction focuses on the application of analytical techniques and model construction, with less emphasis on the underlying requirements for measurement, such as scale properties, dimensional homogeneity, and invariance. As a result, students are trained to manipulate numerical inputs and interpret outputs without a formal foundation in measurement theory.

This creates a curriculum in which HTA is defined by tools and techniques, rather than by the conditions required for valid quantitative claims.

INTERROGATING LARGE LANGUAGE MODELS

Large language models (LLMs) are computational systems trained on extensive text corpora to generate and evaluate language-based responses. Their relevance for this work lies not in their ability to provide definitive answers, but in their capacity to reflect the structure and reinforcement patterns of a defined knowledge base. When appropriately constrained, LLMs can be used as instruments to interrogate how concepts are represented, prioritized, or neglected within a body of literature.

The approach adopted here is based on a diagnostic statement interrogation. A fixed set of canonical statements covering core principles of measurement theory, scale properties, arithmetic conditions, and HTA constructs is presented to the model. Each statement requires a binary response (true/false), with the model then assigning a categorical probability of endorsement. These probabilities are not treated as precise estimates, but as structured indicators of the degree to which a statement is supported within the knowledge base.

To avoid spurious precision, probabilities are restricted to a defined set of categories (e.g., 0.05, 0.10, 0.20, 0.50, 0.80, 0.90, 0.95). These values are then transformed into normalized logits, calculated as the natural logarithm of the odds of endorsement. The logit transformation provides a continuous scale—bounded within a specified range (typically ±2.50)—that allows comparison across statements while preserving the ordinal structure of the categorical probabilities.

The purpose of this process is not to validate individual responses, but to identify patterns of reinforcement and absence. Statements that are true but receive low endorsement probabilities indicate concepts that are largely absent from the knowledge base. Conversely, false statements that receive high endorsement probabilities indicate the normalization of incorrect or non-measurement-based assumptions. The resulting logit profile provides a structured representation of how the knowledge base aligns—or fails to align—with the requirements of measurement.

Applied across multiple institutions, this method reveals consistent patterns. In particular, statements relating to representational measurement, dimensional homogeneity, and Rasch-based measurement of latent attributes tend to show low endorsement, while statements supporting composite constructs and non-measurement-based arithmetic often show high endorsement. These patterns are not isolated anomalies but reflect a systematic configuration of the knowledge base.

In this sense, LLM interrogation serves as a tool for mapping the conceptual foundations of HTA. It does not replace analysis, but provides a reproducible method for demonstrating the extent to which measurement principles are embedded or excluded from current teaching and research frameworks.

WHAT IS MEASUREMENT INVERSION?

Measurement inversion describes a fundamental error in the construction and use of quantitative claims. In any scientific discipline, the correct sequence is clear: an attribute must be defined, a scale with appropriate properties must be established, and only then may arithmetic operations be applied. Measurement precedes arithmetic. Measurement inversion occurs when this sequence is reversed—when numerical operations are applied before the conditions required for measurement have been satisfied.

At its core, measurement inversion treats numbers as if they were measures. Numerical values are assigned, combined, averaged, or multiplied without first demonstrating that they represent a single, well-defined attribute on a scale that supports those operations. The presence of numbers is taken as sufficient justification for arithmetic. This assumption is false. Numbers may represent labels, ranks, or counts, but only under specific conditions do they represent **quantities**.

The consequences follow immediately. Arithmetic operations impose structure. Addition assumes that differences are meaningful and comparable across the scale. Averaging assumes that values can be combined without loss of meaning. Multiplication and division require a **ratio scale**, including a meaningful zero and dimensional coherence between the quantities involved. Where these conditions are not met, the result of the calculation may be a number, but it does not represent a measurable quantity.

Measurement inversion is therefore not simply a technical oversight; it is a category error. It substitutes numerical manipulation for measurement and presents the result as if it were quantitative. The outputs appear precise, often to several decimal places, but their precision is illusory. Without established scale properties, there is no basis for interpreting differences, ratios, or changes.

In practice, measurement inversion leads to the construction of claims that cannot be evaluated in a scientific sense. If the underlying values are not measures, then observed differences cannot be interpreted as changes in an attribute, and ratios cannot be interpreted as proportional effects. Such claims are not falsifiable because there is no defined quantity against which observation can be compared.

The significance of measurement inversion lies in its implications for scientific inquiry. A discipline that applies arithmetic without establishing measurement cannot generate credible, evaluable, and replicable knowledge. It produces numbers, but not quantities; outputs, but not evidence. Recognizing and correcting measurement inversion is therefore a prerequisite for restoring the conditions under which quantitative claims can be meaningfully interpreted and tested.

IMPLICATIONS FOR TEACHING AND RESEARCH PROGRAMS

The category error of measurement inversion requires abandoning the analytical baggage that has accumulated around the valuation of health states. That decision, taking descriptive, multi-attribute health states as the starting point, set the discipline on a path that bypassed the requirements of representational measurement**.** What followed was not measurement, but the construction of numerical narratives: values assigned, combined, and manipulated without establishing the properties required for quantitative interpretation. The apparent sophistication of these approaches has obscured a simple fact: without measurement, such constructions remain descriptive artifacts rather than measurable quantities**.**

The implication for teaching and research is therefore not incremental revision, but conceptual reset. A curriculum that begins with the valuation of health states inherits these limitations from the outset. It directs students toward techniques that generate numbers, but not measures, and toward claims that cannot be evaluated in a scientific sense. Abandoning this starting point is not a matter of preference, but a requirement if HTA is to support credible, testable claims of therapy impact.

For teaching, the first implication is the need to reorder the curriculum. Measurement must precede arithmetic. Students should be introduced to attributes, scale properties, unidimensionality, invariance, and dimensional homogeneity before encountering economic evaluation methods or decision models. The distinction between manifest and latent attributes must be explicit. Manifest attributes can be measured directly on linear ratio scales; latent attributes require transformation to a valid measurement scale. Without this foundation, subsequent instruction in modeling and evaluation lacks interpretive validity.

A second implication is the need to remove or reposition content that relies on non-measured constructs. Methods based on composite indices or ordinal preference scores should not be presented as quantitative tools. They may be discussed as historical or descriptive approaches, but not as a basis for arithmetic operations or claims about magnitude. This is not a marginal adjustment; it requires a clear statement that certain widely taught techniques do not meet the conditions for measurement.

Third, teaching must shift toward protocol-driven, falsifiable claims. Students should learn to define a single attribute, specify the measurement method, identify the target population and timeframe, and state expected effects in advance. This redefines HTA from the interpretation of model outputs to the construction and testing of measurable claims.

For research programs, the implications are equally significant. Study design must begin with attribute specification and measurement selection, not with model construction. Data collection should be aligned with valid measurement instruments, including the use of the Rasch model for latent attributes. Analyses should be limited to operations supported by the scale properties of the data. This constrains what can be claimed, but it also ensures that claims are interpretable and testable.

Finally, both teaching and research must align with the requirements for the evolution of objective knowledge. Findings should be presented as provisional and subject to refutation. Replication and comparison across studies depend on invariant measurement, not on the consistency of model structures.

In summary, addressing measurement inversion requires a shift from method-centered to measurement-centered programs, replacing numerical manipulation with the disciplined construction of measurable, falsifiable claims.

IMPLICATIONS FOR FACULTY

The implications of measurement inversion are not confined to curriculum design or research structures; they extend directly to faculty responsibilities. Faculty are the point at which a discipline is defined, transmitted, and reproduced. What is taught in the classroom and embedded in research programs reflects faculty understanding of what constitutes valid knowledge.

At present, faculty are placed in a difficult position. They are required to teach and apply frameworks that are widely accepted within HTA, yet which do not meet the conditions required for measurement. In most cases, this is not a matter of oversight, but of inheritance. The prevailing framework has been adopted as standard practice, and its assumptions are rarely examined at the level of first principles.

However, the interrogation results make it increasingly difficult to sustain this position. If the values used in HTA do not represent measurable quantities, then the claims derived from them cannot be interpreted as evidence in a scientific sense. This places faculty in a position where they are effectively teaching numerical procedures without a measurement foundation.

The implication is not that faculty have failed, but that they must now reassess the basis on which HTA is taught and applied. This requires a shift from accepting established methods to evaluating whether those methods meet the conditions required for measurement. It also requires recognition that the distinction between manifest and latent attributes, and the appropriate measurement frameworks for each, is central to the construction of valid claims.

At the same time, faculty operate within institutional and policy constraints. Alignment with national agencies, existing curricula, and research expectations limits the scope for immediate change. As a result, transition must be structured and incremental, beginning with the development of internal capability.

The role of faculty is therefore pivotal. They are not only recipients of methodological frameworks, but the agents through which those frameworks are sustained or transformed. A transition to a measurement-based HTA discipline depends on faculty engagement with the principles of measurement and their application in teaching and research.

The question is not whether change is required, but whether faculty will lead that change or continue to operate within a framework that does not meet the standards of scientific inquiry.

A PRACTICAL PATHWAY FOR FACULTY TRANSITION

If the limitations of current HTA frameworks are accepted, the immediate question for faculty is not whether change is required, but how that change can be implemented in practice. The transition to a measurement-based discipline does not require the construction of a new theoretical edifice. The principles are established. What is required is the development of faculty capability to apply those principles in teaching and research.

The pathway is therefore straightforward. It begins with a clear understanding that measurement must precede arithmetic, followed by the ability to distinguish between manifest and latent attributes, and the selection of appropriate measurement frameworks for each. For manifest attributes, this involves the use of linear ratio measures. For latent attributes, it requires transformation of observations into measures through established approaches such as the Rasch model. From this foundation, faculty can construct single-attribute, protocol-driven claims that are capable of empirical evaluation.

The challenge is not conceptual, but operational. Faculty must be able to translate these principles into curriculum design, research protocols, and the evaluation of therapy impact. This requires a structured approach to training that is focused, practical, and directly aligned with existing academic responsibilities.

To support this transition, a six-unit, faculty-focused "train-the-trainer" program has been developed. The program is designed to provide a concise and implementable framework for introducing measurement-based HTA within academic settings. It is delivered by ZOOM in small groups to ensure interaction and application, and is structured to allow participants to integrate the material directly into their teaching and research activities.

The objective is not incremental adjustment, but capability transfer. Faculty completing the program should be able to redesign course content, reframe research questions, and evaluate claims according to the requirements of measurement. In this sense, the transition is not dependent on external mandates, but on internal academic leadership.

As the first step in this disciplinary transition, Maimon Research has developed a faculty-focused "train-the-trainer" program that introduces the analytical framework required to support true measurement in HTA. The objective is not to add another layer to existing teaching, but to replace numerical manipulation with measurement-based practice, equipping faculty to redesign courses, research, and formulary evaluation around valid, testable claims.

The program is structured in six modules.

1. **Understanding Measurement: Why measurement must precede arithmetic**
   Introduces the role of measurement in supporting credible HTA claims and establishes a foundation for subsequent discussion.
2. **Attributes and Claims in Health Technology Assessment**
   Explores how value claims are framed and the importance of clearly defined attributes in their interpretation.
3. **Latent Attributes: Transforming Observations to Measurement**
   Details how observations can be transformed to measurement with the necessary and sufficient Rasch rules.
4. **Measurement, Possession and the Structure of Claims**
   Examines the requirements for claims to be evaluated within a structured and testable framework.

5. **Claims, Protocols and the Evolution of Objective Knowledge**
   Applies these principles to the assessment of claims within formulary and decision-making contexts.
6. **From Numerical Manipulation to Measurement: A new health technology assessment curriculum.** A detailed review of possible curricula and modules for course development with a draft 10-week course and submission documentation

The program extends over 3-weeks with 2 one-hour ZOOM sessions each week with Dr Paul Langley. Each session is supported by a 5,000-word handout with questions and answers. The enrollment, because of the need for meaningful interaction is restricted to 8 participants per program. Participant professional fee is $1,250.00

For program scheduling and enrollment**[Insert Program Link]**

The transition to a measurement-based HTA discipline will not occur through policy statements alone. It will occur where faculty have the tools to implement it.

CONCLUSION

Applying the large language model interrogation framework to 72 colleges and schools of pharmacy in the United States, and to a further 160 HTA-related knowledge bases globally, provides, for the first time, a structured basis for concluding that the field has experienced **four** decades of measurement failure in teaching, research and publications. This is not a marginal shortcoming but a category error. The discipline has relied on assumption-driven constructs that bypass the requirements of representational measurement and employ procedures that cannot support quantitative interpretation.

The implication is unavoidable: HTA, as currently constituted, cannot sustain claims to scientific status. A framework built on non-measured constructs cannot be repaired through incremental refinement. It must abandon the analytical legacy that treats descriptive health state valuations as if they were measurable quantities. Until this step is taken, numerical outputs will continue to lack interpretive meaning, and claims for therapy impact will remain unevaluable.

There is, however, a clear and positive pathway forward. Once the requirements of measurement are applied, the analytical framework becomes simpler, not more complex. Attention is confined to two admissible forms of measurement: linear ratio measures for manifest attributes and Rasch logit–based ratio measures for latent attributes. Within this structure, claims can be specified as single-attribute, protocol-driven propositions, expressed on scales that support valid arithmetic and interpretation.

The gain is substantial. HTA is repositioned as a discipline capable of supporting falsifiable claims, where outcomes can be tested, replicated, and refined. The focus shifts from the production of numerical outputs to the evaluation of measurable effects, restoring the conditions necessary for the evolution of objective knowledge in therapy impact.

In this sense, the recognition of measurement failure is not an endpoint, but a transition. By replacing numerical constructs with valid measures, HTA can move from a closed, self-referential system to a scientifically grounded enterprise, capable of generating credible, evaluable, and replicable evidence.

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.