MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

UNITED STATES: THE MISSING SCIENCE OF MEASUREMENT IN HEALTH TECHNOLOGY ASSESSMENT

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

LOGIT WORKING PAPER No 1240 JUNE 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

This paper examines the intellectual foundations of health technology assessment (HTA) within seven leading United States academic research centers and evaluates the extent to which their curriculum knowledge bases incorporate the principles required for quantitative science. The assessment was motivated by a broader series of large language model (LLM) interrogations of HTA agencies, academic institutions, pharmacy schools and research organizations that consistently identified a pervasive pattern of measurement inversion. In measurement inversion, arithmetic and modelling precede the demonstration of measurement, reversing the sequence required in the physical and social sciences where measurement must be established before numerical operations are undertaken.

To investigate whether this pattern may originate in the educational environment that trains future researchers and policy analysts, curriculum knowledge bases associated with the Schaeffer Center, Harvard Center for Health Decision Science, CHOICE Institute, Duke-Margolis Institute, Leonard Davis Institute, Bloomberg School of Public Health and Stanford Health Policy Center were interrogated against ten canonical statements representing the minimum requirements for a measurement-based approach to therapy assessment. These statements addressed attributes, target attributes, scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement and falsifiable claims.

The results reveal a strikingly consistent profile across all seven centers. While endorsement was relatively strong for the identification of attributes and the importance of falsifiable claims, endorsement collapsed for scales of measurement, representational measurement, unidimensionality, latent attribute measurement and ratio measurement. Particularly noteworthy was the near-complete absence of representational measurement, Rasch measurement and the distinction between manifest and latent attributes requiring different forms of ratio measurement. The findings suggest that students and researchers are introduced to modelling, economic evaluation and policy analysis while receiving little systematic exposure to the science of measurement itself.

The paper argues that this pattern of curriculum inversion provides a plausible explanation for the widespread measurement inversion observed throughout HTA practice. Concepts absent from HTA research and policy appear equally absent from the educational environment that prepares future practitioners. The implications are profound. If HTA is to develop as a scientific discipline capable of generating evaluable, replicable and falsifiable claims regarding therapy impact, reconstruction must begin with the curriculum. A measurement-based framework founded on representational measurement, linear ratio measurement for manifest attributes and Rasch logit ratio measurement for latent attributes is proposed as the pathway for reconstruction.

INTRODUCTION

Health technology assessment (HTA) presents itself as a quantitative discipline. Students are introduced to economic evaluation, cost-effectiveness analysis, utility assessment, simulation modelling and reimbursement decision making. These methods generate numerical claims regarding therapy impact and are routinely employed to support formulary recommendations, coverage decisions and resource allocation. The legitimacy of such claims, however, depends upon a prior question that is rarely asked: do the quantities entering these analyses satisfy the requirements of measurement? This is the critical question for the HTA content of US research center programs given their role in establishing professional standards.

In every quantitative science, measurement precedes arithmetic. Before quantities can be added, multiplied, transformed or modelled, their measurement properties must first be established. This principle is not optional. It is the foundation upon which quantitative reasoning rests. Without measurement there can be no meaningful arithmetic, and without lawful arithmetic there can be no credible quantitative claims.

The requirements of measurement are well understood. They include the distinction between nominal, ordinal, interval and ratio scales the role of representational measurement; the requirement for unidimensional attributes; the distinction between manifest and latent attributes; the unique role of Rasch logit ratio measurement and the conditions under which quantities may support arithmetic operations such as multiplication and division [1] [2] [3] [4] . These concepts define the scientific standards against which quantitative claims must be judged.

Recent large language model (LLM) interrogations of HTA knowledge bases across multiple countries have revealed a pervasive pattern of measurement inversion detailed on the Maimon Research website [5] . Concepts central to measurement science receive little recognition, while analytical frameworks that depend upon those concepts receive strong endorsement [6]. Utilities are treated as though they possess interval or ratio properties. QALYs are accepted as quantitative measures despite the absence of demonstrated ratio-scale foundations. Simulation models generate lifetime claims without establishing the measurement status of the quantities entering those models. The same pattern appears repeatedly across agencies, academic centers, professional organizations and HTA teaching programs.

This raises an obvious question. If measurement inversion is so widespread within HTA practice, where does it originate? One possibility is that the concepts necessary to recognize and challenge measurement inversion are absent from the educational environment that trains future researchers and practitioners. If students are not introduced to measurement theory, representational measurement, unidimensionality, latent attribute measurement and ratio scales, then the persistence of measurement inversion should not be surprising.

The purpose of this paper is to evaluate the curriculum knowledge base supporting HTA programs. The assessment focuses not on individual courses or instructors but on the broader intellectual environment that shapes teaching, research and professional development. A series of canonical statements was developed to represent the minimum concepts required for a measurement-based approach to therapy assessment. These statements address attributes, scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement and falsifiable claims.

The curriculum interrogation reveals two distinct but related problems. The first is one of omission. Concepts central to measurement science, including scales of measurement, representational measurement, unidimensionality, latent attribute measurement, Rasch measurement and ratio measurement, are largely absent from curriculum coverage. Students are therefore unlikely to be exposed to the scientific principles necessary to evaluate whether quantitative claims satisfy the requirements of measurement.

The second problem is more serious. It is not merely that essential concepts are absent; it is that students are taught analytical frameworks that depend upon assumptions which fail those same standards. Utilities, QALYs and reference-case simulation models are presented as legitimate instruments of quantitative assessment despite their inability to satisfy the requirements of representational measurement. Students are therefore not only deprived of the concepts needed to evaluate quantitative claims they are simultaneously introduced to frameworks that cannot be defended once those concepts are applied.

The result is a self-reinforcing cycle. The absence of measurement science prevents students from recognizing the limitations of the methods they are taught, while the methods themselves reinforce the perception that quantitative legitimacy can be achieved through arithmetic and modelling alone. Curriculum inversion and measurement inversion therefore become mutually supporting phenomena. One explains the persistence of the other.

THE IMPERATIVE OF MEASUREMENT INVERSION

The starting point for any reconstruction of health technology assessment (HTA) must be an understanding of the scale and consistency of the measurement inversion that characterizes the field. Over the past several months, a series of large language model (LLM) interrogations has been undertaken to evaluate HTA knowledge bases across multiple jurisdictions. These interrogations have encompassed national reimbursement agencies, academic research centers, professional organizations, journals, pharmacy schools and HTA teaching programs in Australia, Canada, New Zealand, the United Kingdom, Europe, Singapore and the United States. Although the institutions differ in structure, mission and geographical location, the results have been remarkably consistent. The same pattern of measurement inversion appears irrespective of country, discipline or organizational setting.

The concept of measurement inversion is straightforward. In the physical and social sciences, measurement precedes arithmetic. Before quantities can be added, multiplied, transformed or modelled, the measurement properties of those quantities must first be established. HTA reverses this sequence. Numerical operations are routinely undertaken without demonstrating that the quantities involved satisfy the requirements of measurement. Arithmetic is treated as a substitute for measurement rather than a consequence of measurement. The result is a framework in which numerical manipulation takes precedence over the validation of the quantities being manipulated.

The LLM interrogations reveal this inversion repeatedly. Across institutions there is little awareness of scales of measurement, the distinction between ordinal, interval and ratio scales, the role of unidimensionality, the axioms of representational measurement, dimensional homogeneity, or the distinction between manifest and latent attributes. At the same time, there is widespread endorsement of analytical frameworks that depend upon precisely those concepts. Utilities are treated as though they possess interval or ratio properties. Quality-adjusted life years are accepted as quantitative measures despite the absence of demonstrable ratio-scale foundations. Reference-case simulation models are employed to generate lifetime cost-effectiveness claims without establishing the measurement properties of the underlying inputs. The pattern is universal. The details vary from one institution to another, but the underlying logic remains unchanged.

The significance of these findings lies not merely in the demonstration of measurement failure but in the realization that measurement inversion is now institutionalized. It is embedded in teaching programs, methodological guidance, research publications and policy frameworks. Researchers entering the field encounter an intellectual environment in which the requirements of measurement are rarely discussed and almost never presented as prerequisites for quantitative claims. Consequently, the acceptance of measurement inversion becomes self-reinforcing. Successive generations of researchers inherit analytical frameworks without being introduced to the measurement principles required to evaluate them.

It is worth emphasizing the complete absence of Rasch measurement from the practice profiles of the US HTA research centers. This omission is particularly significant because Rasch measurement represents the only established framework that provides the necessary and sufficient conditions for transforming observations relating to a latent attribute into a measure. The central issue is not statistical sophistication but measurement itself. Latent attributes such as symptom burden, functional status, treatment satisfaction, need fulfilment and quality of life cannot be observed directly. They require a measurement model capable of estimating possession of the attribute while demonstrating unidimensionality, invariance and appropriate scale properties. Rasch measurement was developed specifically to meet these requirements. Its absence suggests that students are not exposed to the concept of latent attribute possession, nor to the scientific challenge of constructing measures from observations. Instead, there is a tendency to treat questionnaire scores, indexes and composite summaries as though they were measures in their own right instead of composite ordinal scores with no measurement properties.

The result is that one of the most important developments in measurement theory over the past 80 years is effectively invisible within US HTA education. Without exposure to Rasch measurement, students are unlikely to appreciate that latent attributes require a fundamentally different approach to measurement from manifest attributes, or that quantitative claims regarding latent phenomena require a demonstrable measurement framework before arithmetic operations can be justified.

This observation raises an obvious question. If measurement inversion is so widespread and persistent, where does it originate? The answer cannot be found solely in research practice or policy guidance. Attention must also be directed toward education. If the concepts necessary to recognize measurement failure are absent from the curriculum, then measurement inversion becomes the expected outcome rather than an isolated error. Researchers cannot be expected to evaluate measurement claims if they have never been introduced to scales of measurement, representational measurement, ratio scales, unidimensionality, or the distinction between manifest, latent attributes and even Rasch measurement.

For this reason, curriculum assessment emerges as a critical component of HTA reconstruction. The objective is not simply to determine whether students are exposed to contemporary HTA methods. Rather, it is to determine whether they are exposed to the foundational concepts that make the evaluation of those methods possible. A curriculum that emphasizes modelling, economic evaluation and decision analysis while neglecting measurement theory will inevitably reproduce the same conceptual limitations observed in current HTA practice.

The curriculum interrogations undertaken across US HTA research centers provide compelling support for this interpretation. While there is evidence that students and researchers are introduced to outcomes assessment, target attributes and scientific claims, there is little evidence of systematic exposure to attribute claim specification, scales of measurement, the axioms of representational measurement, unidimensionality, latent attribute measurement or the unique contribution over 60 years ago of Rasch measurement. The concepts most frequently absent from curriculum coverage are precisely those concepts most frequently absent from HTA practice. The relationship is unlikely to be coincidental.

The imperative of measurement inversion therefore extends beyond criticism of existing methods. It points directly to the need for educational reconstruction. If HTA is to move toward a framework based on lawful measurement, evaluable claims and empirical falsification, then curriculum reform must accompany methodological reform. The widespread and consistent pattern of measurement inversion revealed by the LLM interrogations suggests that reconstruction cannot begin with policy guidance or analytical techniques alone. It must begin with the curriculum. Until students and researchers are introduced to the foundations of measurement science, the conditions that created measurement inversion will continue to be reproduced throughout the HTA community.

DEFINING THE KNOWLEDGE BASE

The first step in any LLM interrogation is to define the knowledge base to be interrogated. The validity of the interrogation depends upon ensuring that the knowledge base reflects the information environment that shapes teaching, research and professional practice. In the present assessment, the curriculum knowledge base for each research center is defined as the totality of publicly accessible materials that contribute to HTA education and training. These include curriculum descriptions, course outlines, program objectives, teaching materials, methodological guidance documents, seminar and workshop content, faculty publications, doctoral training resources, research center reports, conference presentations, policy briefs and other materials through which knowledge is communicated to students, researchers and professional staff.

The objective is not to evaluate individual courses or instructors but to assess the broader intellectual environment within which HTA concepts are introduced, reinforced and transmitted. The resulting curriculum knowledge base is assumed to represent the concepts and principles that students and researchers are most likely to encounter during their exposure to the research center and its associated educational activities. It is this knowledge base that is interrogated to determine the extent to which the foundational concepts required for a measurement-based approach to therapy assessment are present, absent or only weakly represented.

INTERROGATING THE CURRICULUM

The identification of measurement inversion across HTA research centers, reimbursement agencies and academic programs raises an obvious question: where does this inversion originate? If the same conceptual failures are observed repeatedly across institutions and jurisdictions, then the explanation cannot rest solely with individual researchers, policy analysts or decision makers. A more plausible explanation is that these failures reflect deficiencies in the educational environment that prepares future HTA practitioners. If concepts central to measurement science are absent from curriculum content, then their absence from research practice should not be surprising. This realization provides the rationale for interrogating the curriculum.

The objective of curriculum interrogation differs from that of previous HTA knowledge-based assessments. Earlier interrogations focused on whether institutions recognized the requirements of representational measurement and the standards necessary for quantitative claims. Curriculum interrogation asks a different question. Are faculty, students and researchers exposed to the concepts necessary to understand and apply those standards? The focus shifts from methodological outputs to educational inputs. Rather than examining what faculty, students and researchers do, attention is directed to what they are taught and what they know.

The importance of this distinction should not be underestimated. Educational programs do not merely transmit technical skills. They define the conceptual framework through which future practitioners understand evidence, measurement and scientific inquiry. Concepts that are absent from the curriculum are unlikely to emerge spontaneously in research practice. Equally, concepts that are emphasized repeatedly become part of the intellectual assumptions that shape subsequent analysis. If measurement inversion is widespread, then one possibility is that the educational foundations required to recognize and avoid measurement inversion have never been systematically incorporated into HTA teaching and research training.

For this reason, the curriculum interrogation was designed around a series of canonical statements intended to identify the presence or absence of foundational measurement concepts. These statements were deliberately elementary. The purpose was not to assess advanced methodological knowledge but to determine whether faculty, students and researchers are likely to encounter the principles that underpin lawful quantitative claims. The resulting framework begins with the concept of an attribute as the object of measurement and proceeds through target attribute specification, scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement and falsifiable claims. Together, these statements define the minimum intellectual foundations required for a measurement-based approach to therapy assessment in education.

These statements are:

- **An attribute is the specific outcome of interest in a therapy assessment.**
- **Every therapy assessment begins with specification of the target attribute.**
- **The principal scales of measurement (nominal, ordinal, interval and ratio) have different properties and support different forms of analysis.**

- **The measurement status of a target attribute must be established before quantitative claims can be advanced.**
- **The axioms of representational measurement underpin quantitative claims.**
- **Attributes must be demonstrated to be unidimensional before measurement is possible.**
- **A manifest attribute is directly observable and capable of supporting empirical observation.**
- **A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute.**
- **Manifest and latent attributes require different forms of ratio measurement.**
- **Therapy impact claims must be falsifiable.**

These ten statements form a logical sequence:

Attribute → Target Attribute → Scales of Measurement → Measurement Status → Representational Measurement → Unidimensionality → Manifest Attribute → Latent Attribute → Ratio Measurement → Falsifiable Claims

Together they define the minimum curriculum content required for a measurement-based approach to HTA and provide the framework for evaluating curriculum coverage.

The categorical probabilities reported in this assessment are intended as indicators of the extent to which a concept is represented within the curriculum knowledge base. They should not be interpreted as precise statistical estimates but as measures of the likelihood that a student, researcher or professional exposed to that knowledge base would encounter, recognize and subsequently endorse the canonical statement. In practical terms, the probability reflects the visibility and prominence of a concept within the educational environment associated with a research center or policy agency.

A high probability indicates that the concept is well represented within curriculum materials, research outputs and educational activities and is therefore likely to be familiar to students and researchers. Conversely, a low probability suggests that the concept is absent, only weakly represented, or occupies a peripheral position within the curriculum knowledge base. Students exposed to such an environment would therefore be unlikely to recognize the concept as an important component of HTA education and practice.

The probabilities should be viewed comparatively rather than in isolation. Their principal value lies in identifying patterns of curriculum coverage across institutions and concepts. In particular, low probabilities associated with scales of measurement, representational measurement, unidimensionality and ratio measurement indicate that these topics are unlikely to form a substantial part of the educational experience of the average student. The resulting profile provides an indication of curriculum strengths, deficiencies and potential areas for reconstruction.

CURRICULUM INVERSION IN LEADING US HEALTH ECONOMICS AND HEALTH POLICY RESEARCH CENTERS

The curriculum assessments of seven leading United States health economics, health policy and outcomes research centers reveal a striking and highly consistent pattern (Table 1). Although these institutions differ in organizational structure, disciplinary emphasis and educational mission, they share an almost identical profile of endorsement across the ten canonical statements. The Schaeffer Center for Health Policy and Economics at the University of Southern California [7], the Harvard Center for Health Decision Science [8], the CHOICE Institute at the University of Washington [9], the Duke-Margolis Institute for Health Policy [10], the Leonard Davis Institute of Health Economics at the University of Pennsylvania [11], the Bloomberg School of Public Health at Johns Hopkins University [12] and the Stanford Health Policy Center [13] all demonstrate essentially the same misunderstanding of what constitutes the intellectual foundation of health technology assessment and health policy analysis.

The intellectual foundations of HTA in the United States are best evaluated by considering individually the ten canonical statements that formed the curriculum knowledge-based interrogation of seven leading academic HTA research centers. These statements were selected because they represent the essential elements of measurement science required to support quantitative claims. They move sequentially from identification of an attribute, through the scales and axioms of measurement, to unidimensionality, the distinction between manifest and latent attributes, ratio measurement and the requirements for falsifiable claims. Collectively, they provide a framework for determining whether the scientific foundations necessary for quantitative inquiry are embedded within the intellectual environment that supports HTA research, graduate training and policy development. Their intellectual influence extends well beyond their own institutions to researchers, policy advisers, healthcare organizations and government agencies throughout the United States and internationally.

The curriculum assessments were undertaken in response to the earlier finding of widespread measurement inversion in US HTA knowledge bases. If concepts central to representational measurement are absent from the educational and research environment responsible for training future investigators and developing HTA methods, then their absence from HTA practice should not be surprising. The curriculum assessments therefore provide an opportunity to examine whether the scientific foundations required for quantitative claims are present within the intellectual infrastructure of US HTA.

What follows is therefore more than a review of curriculum content on a statement-by-statement basis. It is an assessment of whether the concepts necessary for quantitative science are visible within the leading academic centers responsible for shaping the future direction of US HTA. The results suggest that while these centers demonstrate considerable sophistication in economic evaluation, modelling, decision analysis and outcomes research, they provide little evidence of systematic engagement with the science of measurement itself. The concepts most frequently absent from HTA practice appear also to be those most frequently absent from the intellectual framework that supports HTA research and training. This relationship provides a plausible explanation for the persistence of measurement inversion throughout the field.

TABLE 1: CURRICULUM CONTENT ENDORSEMENT - US HTA RESEARCH CENTERS

CANONICAL STATEMENT	CATEGORICAL PROBABILITY ENDORSEMENT						
	SCHAEFFER	HARVARD	CHOICE	DUKE	LEONARD	BLOOMBERG	STANFORD
An attribute is the specific outcome of interest in a therapy assessment	0.85	0.85	0.85	0.85	0.85	0.85	0.85
Every therapy assessment begins with specification of the target attribute	0.80	0.85	0.85	0.85	0.85	0.85	0.85
The principal scales of measurement (nominal, ordinal, interval and ratio) have different properties and support different forms of analysis	0.20	0.20	0.20	0.20	0.20	0.20	0.20
The measurement status of a target attribute must be established before quantitative claims can be advanced	0.15	0.15	0.15	0.15	0.15	0.15	0.15
The axioms of representational measurement underpin quantitative claims	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Attributes must be demonstrated to be unidimensional before measurement is possible	0.10	0.10	0.10	0.10	0.10	0.10	0.10
A manifest attribute is directly observable and capable of supporting empirical observation	0.50	0.55	0.55	0.55	0.55	0.55	0.55
A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Manifest and latent attributes require different forms of ratio measurement	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Therapy impact claims must be falsifiable	0.80	0.80	0.80	0.85	0.85	0.85	0.85

SCHAEFFER: Schaffer Center for Health Policy and Economics, USC; HARVARD: Harvard Center for Health Decision Science; CHOICE: University of Washington CHOICE Institute; DUKE: Duke-Margolis Institute for Health Policy, Duke University; LEONARD: Leonard Davis Institute of Health Economics, University of Pennsylvania; BLOOMBERG: Bloomberg School of Public Health, Johns Hopkins University; STANFORD: Stanford University Health Policy Center

Statement 1: An Attribute is the Specific Outcome of Interest in a Therapy Assessment

The strong endorsement of this statement with all endorsements at p=0.85 suggests that US HTA curricula recognize the need to identify an outcome of interest before undertaking an assessment. This is hardly surprising. Every evaluation must have a focus. Researchers and students are taught to identify outcomes such as survival, resource utilization, symptom control, disease progression, patient experience and treatment satisfaction. In this respect, the curriculum acknowledges that assessment requires an object of inquiry.

The problem is that the curriculum appears largely to stop at identification. An attribute is recognized as something to be assessed, but little attention is given to the far more important scientific question: is the attribute capable of supporting measurement? The distinction is fundamental. Naming an outcome does not establish that it can be measured. Identifying an attribute is merely the beginning of the scientific process.

Indeed, the identification of an attribute is arguably the least demanding part of quantitative inquiry. Any researcher can specify an outcome of interest. The scientific challenge begins only when it is asked whether the attribute exists as a measurable phenomenon. Is it manifest or latent? Is it unidimensional? Does it possess a structure capable of supporting measurement? What evidence demonstrates that numerical observations correspond to differences in the attribute itself? These questions determine whether a quantitative claim is possible.

The curriculum evidence suggests that these questions are rarely raised. Once an attribute has been identified, there is a tendency to move directly to data collection, instrument selection, modelling and analysis. The existence of the attribute is tacitly assumed to justify its numerical representation. Yet measurement science makes no such assumption. The existence of an attribute does not guarantee the existence of a measure.

The consequence is that students may develop the impression that once an outcome has been identified the quantitative task has effectively been completed. The much more demanding question of measurement is never reached. The curriculum therefore creates a false sense of scientific security. Attributes are identified but rarely subjected to scrutiny regarding their measurement properties.

This omission has important implications. If students are taught only to identify attributes, they are unlikely to ask whether the attribute is manifest or latent, whether it is unidimensional, whether it supports interval or ratio measurement, whether arithmetic operations are admissible, or whether quantitative claims concerning the attribute are possible at all. The entire measurement problem disappears from view.

The result is that outcomes are routinely accepted as candidates for quantitative analysis without first establishing whether they satisfy the requirements of measurement. Numerical representation becomes a matter of convention rather than demonstration. Quantification follows automatically from identification. The scientific burden of proving measurement is quietly abandoned.

This seemingly modest omission may in fact be one of the origins of measurement inversion. Once outcome specification is mistaken for measurement, the door is opened to increasingly elaborate forms of arithmetic, modelling and simulation that rest upon quantities whose measurement status has never been established. The framework acquires scientific authority because it is numerical, not because measurement has been demonstrated.

Without a framework that moves beyond simple attribute identification, or its equivalent, students are not equipped to challenge the assumptions that underpin the reference case.

Statement 2: Every Therapy Assessment Begins with Specification of the Target Attribute

The endorsement of this statement is strong (range $p = 0.80$ to 0.85), indicating that US HTA programs recognize the importance of defining the target of an assessment. This reflects standard practice in research design. A study must have a clearly defined objective and a clearly specified outcome. In this respect, the curriculum demonstrates an awareness that evaluation requires an object of inquiry.

Yet there is a danger that this emphasis creates the illusion of scientific rigor where none necessarily exists. Defining an attribute is often treated as though it were equivalent to defining a measure. It is not. The specification of an outcome merely identifies what is to be examined. It says nothing about whether the outcome can support measurement. Naming an attribute is not the same as demonstrating that the attribute can be measured.

This distinction is fundamental because measurement begins where outcome specification ends. Once an attribute has been identified, a series of scientific questions immediately follow. Is the attribute manifest or latent? Is it unidimensional? What empirical evidence supports its existence? What scale properties can be demonstrated? Does the attribute support interval or ratio measurement? What measurement framework is required? These questions determine whether a quantitative claim is possible.

The curriculum evidence suggests that this intermediate stage is largely absent. Once the target attribute has been identified, attention shifts rapidly to data collection, modelling, statistical analysis and economic evaluation. The crucial step of establishing the measurement status of the attribute is either compressed or ignored altogether. Outcome specification becomes a gateway to analysis rather than a gateway to measurement.

The implications are profound. Students may become highly proficient in analytical techniques while remaining largely unaware that the scientific challenge has barely begun. The attribute is treated as though its measurement properties were self-evident. Quantitative analysis proceeds without first demonstrating that the quantity entering the analysis is capable of supporting quantitative interpretation.

This omission helps explain a persistent feature of contemporary HTA. The existence of an outcome is frequently assumed to justify its numerical representation. If pain, quality of life, treatment satisfaction or health status can be named, then it is often assumed that these attributes can be quantified. Yet measurement science makes no such assumption. The existence of an attribute does not guarantee the existence of a measure. Demonstrating measurement is a separate scientific task.

The result is a culture in which specification of an attribute is mistaken for establishment of a measure. The scientific challenge is transformed into an administrative exercise. Outcomes are identified, instruments selected, data collected and analyses performed without first establishing whether the attribute supports measurement. Numerical outputs therefore acquire legitimacy because the outcome was specified, not because measurement was demonstrated.

This confusion between attribute identification and measurement may be one of the most important sources of measurement inversion in HTA. Students learn how to identify outcomes but are not systematically taught how to determine whether those outcomes can be measured. Consequently, the distinction between an attribute and a measure becomes blurred, and quantitative claims acquire an authority that has not been earned.

Without a framework that distinguishes outcome specification from measurement, or its equivalent, students are not equipped to challenge the assumptions that underpin the reference case.

Statement 3: The Principal Scales of Measurement Have Different Properties and Support Different Forms of Analysis

This is where the profile begins to collapse. Endorsement falls to only $p=0.20$ across every research center. This finding should be deeply troubling because the scales of measurement are among the most elementary concepts in quantitative science. The distinction between nominal, ordinal, interval and ratio scales is not an abstract methodological curiosity. It determines what quantitative claims are possible and what arithmetic operations are permissible.

Every scale carries its own limitations. Nominal scales support classification. Ordinal scales support ranking. Interval scales support differences. Ratio scales support proportional comparisons, multiplication and division. These distinctions are fundamental because arithmetic does not exist independently of measurement. The operations that can be performed on a quantity are determined by the properties of the scale from which the quantity is derived.

The practical implications are profound. If an observation is merely ordinal, then arithmetic operations that require interval or ratio properties cannot be justified. If a quantity lacks a non-arbitrary zero, proportional statements become impossible. If a scale supports ranking but not magnitude, then numerical differences cannot automatically be interpreted as quantitative differences. These are not technical details. They are the rules governing quantitative reasoning.

The curriculum results suggest that students are exposed to quantitative techniques while receiving little systematic instruction regarding the scales upon which those techniques depend. They may learn statistical analysis, economic evaluation, simulation modelling and decision analysis without first understanding the scale properties of the quantities entering those analyses. Consequently, arithmetic operations are treated as routine analytical procedures rather than as activities that require prior justification.

This omission has consequences that extend throughout HTA. Once scale theory disappears, there is no obvious framework for asking whether a quantity supports addition, subtraction, multiplication or division. Questions concerning admissible transformations, interval properties and ratio properties rarely arise because students have never been introduced to the concepts that generate those questions. Numerical outputs are therefore accepted largely at face value.

The absence of scale theory also helps explain why measurement inversion can become institutionalized. If students are unfamiliar with the distinctions between nominal, ordinal, interval and ratio scales, they are unlikely to challenge numerical claims based on scale properties. A quantity presented as quantitative is accepted as quantitative. A numerical score presented as a measure is accepted as a measure. The scientific burden of demonstrating measurement disappears because the intellectual framework required to impose that burden is absent.

More importantly, the scales of measurement provide the foundation for everything that follows. Representational measurement cannot be understood without an appreciation of scale properties. The principle that measurement precedes arithmetic depends upon recognition that arithmetic operations require specific scale characteristics. Unidimensionality becomes important because a scale must represent a single attribute. Ratio measurement becomes important because multiplication and division require ratio properties. Latent measurement becomes important because the objective is to create a scale with defensible measurement properties. Remove scale theory and the entire structure of measurement science disappears.

The irony is striking. Students are trained in increasingly sophisticated analytical methods while receiving little exposure to the most elementary principles governing quantitative evidence. They learn how to analyze numbers before they learn what those numbers represent. They learn how to perform arithmetic before they learn when arithmetic is justified. The result is a curriculum that emphasizes quantitative technique while neglecting the scientific foundations upon which quantitative technique depends.

This finding helps explain the persistence of measurement inversion throughout HTA. Researchers cannot be expected to question the scale properties of quantities if they have never been taught that different scales possess different properties and support different forms of analysis. The question never arises because the conceptual framework required to ask it is absent.

Without an understanding of the principal scales of measurement, or an equivalent framework, students are not equipped to challenge the assumptions that underpin the reference case.

Statement 4: The Measurement Status of a Target Attribute Must be Established Before Quantitative Claims Can be Advanced

The weak endorsement of this statement for all research centers (p=0.15) represents one of the most damaging findings in the entire assessment. Across the institutions examined there is little evidence that students are systematically exposed to the principle that measurement precedes arithmetic.

This principle is not a technical detail. It is the foundation of every quantitative science. Before a quantity can be manipulated mathematically, its measurement properties must be established. Before arithmetic can be justified, the nature of the quantity must be known. Before a quantitative claim can be advanced, it must first be demonstrated that the attribute supports measurement.

The importance of this sequence becomes obvious when considered in practical terms. Scientists do not begin by asking what calculations can be performed. They begin by asking what is being measured. The attribute must first be identified. Its empirical properties must be understood. The appropriate scale of measurement must be established. Only then can arithmetic operations be considered. Measurement therefore determines the range of permissible analyses. Arithmetic does not determine the nature of measurement.

Representational measurement emerged precisely to prevent the reverse process. Numbers can easily create the appearance of quantitative knowledge. A numerical score, an index or a model output may look scientific simply because it is numerical. The role of measurement theory is to prevent this confusion by insisting that measurement be demonstrated before arithmetic is undertaken. Numerical manipulation is not evidence of measurement. It presupposes measurement.

The curriculum evidence suggests that this sequence has been reversed. Students are taught how to perform arithmetic operations, construct models and evaluate outcomes without first being taught how to determine whether the quantities entering those analyses are measures. Arithmetic becomes the starting point rather than the endpoint of inquiry.

This reversal has profound consequences. Once measurement is removed from its central position, numerical manipulation acquires an authority that it has not earned. Quantities are accepted because they can be entered into a model rather than because they satisfy the conditions necessary for measurement. Analytical sophistication substitutes for scientific validation. The focus shifts from demonstrating measurement to constructing increasingly elaborate numerical frameworks.

The consequences extend directly to the intellectual foundations of HTA. Questions that should arise before any quantitative analysis are rarely asked. What is the attribute? Is it manifest or latent? Has unidimensionality been demonstrated? What scale properties have been established? Does the quantity support arithmetic operations? Does it possess interval or ratio properties? These questions disappear because the curriculum begins with analysis rather than measurement.

The result is a form of scientific inversion. In normal science, measurement generates quantities that may subsequently be analyzed. In HTA, quantities often appear first and their measurement status is either assumed or ignored. Numbers become evidence merely because they are numerical. Models become authoritative merely because they are mathematically sophisticated. The requirement that measurement be demonstrated before arithmetic is quietly abandoned.

This finding may provide the single most important explanation for the persistence of measurement inversion throughout HTA. If students are not taught that measurement must precede arithmetic, they have little reason to question whether quantitative claims rest upon lawful measures. The question never arises because the conceptual framework required to ask it is absent.

The irony is striking. Students are trained extensively in the mechanics of quantitative analysis while receiving little instruction regarding the conditions that make quantitative analysis possible. They learn how to manipulate quantities before they learn how quantities become measures. They learn how to construct models before they learn how to determine whether the quantities entering those models represent measurable attributes.

This omission also helps explain the extraordinary durability of the reference case. Once arithmetic is permitted to precede measurement, there is no obvious barrier to constructing increasingly elaborate numerical systems. The critical question, whether the quantities entering the system satisfy the requirements of measurement, is never raised. The framework becomes self-validating because its outputs are numerical. The appearance of quantification substitutes for the demonstration of measurement.

Without the principle that measurement precedes arithmetic, or its equivalent, students are not equipped to challenge the assumptions that underpin the reference case.

Statement 5: The Axioms of Representational Measurement underpin quantitative claims

The probability of endorsement is 0.05 across every research group. This is arguably the single most important finding in the entire assessment. It is difficult to overstate its significance because representational measurement is not a specialized methodological topic, nor is it an optional addition to quantitative inquiry. It is the scientific framework that determines whether quantitative claims are possible. It addresses the fundamental question that precedes all statistical analysis, economic evaluation, modelling and hypothesis testing: under what conditions may numbers legitimately represent attributes?

This question lies at the heart of every quantitative science. Physics, chemistry, engineering and the social sciences all confront the same problem. Before arithmetic can be applied, before equations can be constructed and before quantitative claims can be advanced, there must be a demonstration that an empirical attribute has been represented numerically in a manner consistent with the structure of that attribute. Measurement is therefore not an outcome of arithmetic. Arithmetic is an outcome of measurement.

Representational measurement emerged precisely because numbers can easily create an illusion of knowledge. A numerical assignment may look quantitative while possessing none of the properties required for measurement. Ranking people from one to ten, assigning scores to questionnaire responses or attaching numerical labels to categories creates numbers, but it does not necessarily create measures. The central purpose of representational measurement is to prevent this confusion. It distinguishes measurement from scoring, ranking, indexing and numerical description.

This distinction is crucial because numbers themselves possess no inherent scientific authority. Scientific authority arises only when the numerical representation preserves the empirical relationships present in the attribute being measured. Representational measurement therefore asks a series of questions that should precede every quantitative claim. What is the attribute? Is it a single attribute? What empirical evidence supports its existence? What numerical relations correspond to the empirical relations? What transformations are permissible? What arithmetic operations are admissible? Does the resulting scale support addition, subtraction, multiplication or division?

The framework provides the foundations for admissible transformations, dimensional homogeneity, invariance, cancellation, solvability, order relations and quantitative representation. These concepts may appear abstract, yet they govern every legitimate quantitative claim. Without them there is no scientific basis for deciding whether a quantity is a measure. Numbers may still exist, but their interpretation becomes arbitrary.

The implications of the curriculum results are therefore profound. Students appear to receive little or no exposure to the intellectual framework that determines whether quantitative claims are possible. They may become highly competent users of statistical software, simulation models, economic evaluations and decision analytic techniques, yet possess no formal basis for deciding whether the quantities entering those analyses satisfy the requirements for measurement.

This omission helps explain one of the most striking characteristics of contemporary HTA. Questions that would arise naturally within other scientific disciplines rarely appear. What is the attribute being measured? Is it unidimensional? What evidence demonstrates measurement? What transformations are admissible? Does the quantity support arithmetic operations? Does it possess interval or ratio properties? What empirical structure is being represented? These questions are largely absent because the framework that generates them is absent.

The consequences extend far beyond technical methodology. Representational measurement establishes the principle that measurement precedes arithmetic. Quantities are not measures because they appear in an equation. Quantities are not measures because they can be entered into a model. Quantities are not measures because they generate numerical outputs. Measurement must first be demonstrated. Arithmetic follows.

Yet the curriculum evidence suggests that students are trained in precisely the opposite sequence by faculty who know nothing better. They are introduced to models, simulations, cost-effectiveness analyses and quantitative outputs before they are introduced to the scientific principles that determine whether those outputs are meaningful. Arithmetic precedes measurement. Numerical manipulation precedes demonstration of measurement. The result is curriculum inversion.

This finding provides perhaps the strongest explanation for the measurement inversion observed throughout US HTA. If representational measurement is absent from the curriculum, then measurement inversion becomes almost inevitable. Students are not taught how to distinguish measures from scores. They are not taught how to distinguish quantities from numerical labels.

They are not taught the conditions under which arithmetic operations are admissible. Consequently, numerical outputs acquire scientific authority merely because they are numerical.

The irony is difficult to ignore. Institutions committed to quantitative analysis appear not to teach the science that determines whether quantitative analysis is legitimate. Entire curricula are devoted to teaching increasingly sophisticated methods of numerical manipulation while giving almost no attention to the scientific framework that determines whether those manipulations are lawful. Students learn how to construct models but not how to determine whether the quantities entering those models are measures. They learn how to calculate but not how to determine whether calculation is justified.

The broader implications are troubling. The curriculum results suggest that several generations of faculty, students and researchers may have entered HTA without systematic exposure to the axioms of representational measurement. If so, the widespread acceptance of measurement inversion is no longer surprising. Researchers cannot be expected to challenge assumptions concerning arithmetic, ratio properties, dimensional homogeneity or quantitative claims if they have never encountered the scientific standards by which those assumptions should be judged.

This also explains the remarkable consistency of the results across institutions. The issue is not that representational measurement has been considered and rejected. The evidence points instead to a more fundamental problem: it appears never to have entered the curriculum. The standards themselves are largely unknown. Questions concerning admissible transformations, dimensional homogeneity, invariance, quantitative representation and the conditions for measurement do not arise because the conceptual framework required to generate those questions is absent.

The consequence is that numerical constructions can be mistaken for quantitative evidence and simulation outputs can be mistaken for scientific claims. Once representational measurement disappears, there is no obvious barrier to treating scores as measures, assumptions as observations and arithmetic as evidence. Measurement inversion becomes not an anomaly but an inevitable outcome.

Without representational measurement or its equivalent, students are not equipped to challenge the assumptions that underpin the reference case.

Statement 6: Attributes must be demonstrated to be unidimensional before measurement is possible

The endorsement probability of only $p = 0.10$ suggests that unidimensionality is almost entirely absent from the curriculum. This is a critical finding because unidimensionality is not a desirable feature of measurement; it is a necessary condition for measurement. Before an attribute can be measured it must first be demonstrated that observations represent differing magnitudes of a single phenomenon. Without a single attribute there can be no measure.

The significance of this requirement is often overlooked because it precedes the construction of scales, statistical analyses and mathematical models. Yet unidimensionality is the point at which measurement either succeeds or fails. If observations reflect more than one attribute, then any resulting numerical score is simply a composite summary. It cannot automatically be interpreted as a measure because there is no assurance that the numbers represent variation in a single underlying phenomenon.

The curriculum results suggest that students are rarely exposed to this principle. They may learn how to construct indexes, aggregate outcomes, generate summary scores and combine variables, yet receive little instruction regarding the scientific requirement that measurement must refer to a single attribute. Consequently, multidimensional constructs can be treated as though they were measures without ever confronting the question of whether the underlying attribute exists as a coherent entity.

This omission is particularly important for HTA because many of its most familiar constructs are explicitly multidimensional. Health-state descriptions combine mobility, pain, anxiety, self-care, usual activities and other domains into a single profile. Composite outcome measures routinely combine multiple dimensions of health into a single numerical score. Yet the question that should arise immediately: what is the single attribute being measured appears largely absent.

From a measurement perspective, this question is unavoidable. If a quantity is claimed to be a measure, then the attribute must first be identified. The attribute must then be shown to be unidimensional. Only after this has been demonstrated does it become meaningful to ask how the attribute should be measured. Without unidimensionality there is no scientific basis for treating a numerical score as a quantitative representation of an attribute.

The absence of unidimensionality from the curriculum therefore has consequences that extend far beyond technical measurement theory. Students are not equipped to distinguish between a measure and a numerical summary. They are not taught that aggregating multiple dimensions does not create a single attribute. They are not taught that arithmetic cannot transform multidimensional descriptions into measures. Instead, there is a risk that numerical outputs acquire an unwarranted status simply because they are expressed as numbers.

The implications for latent attributes are equally important. The central achievement of Rasch measurement is that it explicitly addresses the unidimensionality requirement. Before a latent attribute can be measured, evidence must be provided that responses reflect a single underlying construct. Unidimensionality is therefore not an optional feature of Rasch measurement; it is one of its defining requirements. Yet if students are not introduced to unidimensionality, they are unlikely to appreciate why Rasch measurement occupies a unique position in latent variable assessment. Faculty are unlikely to provide assistance.

This omission also helps explain the persistence of measurement inversion. A student unfamiliar with unidimensionality has no reason to challenge multidimensional constructs presented as measures. Composite scores are accepted because they are conventional. Health-state profiles are accepted because they are widely used. Numerical summaries are accepted because they generate numbers. The crucial scientific question whether a single attribute has been demonstrated never arises because the conceptual framework required to ask the question is absent.

The irony is striking. The requirement for unidimensionality is one of the oldest and most fundamental principles of measurement, yet it appears largely invisible within US HTA education. Students are trained to analyze numerical outputs without first being taught how to determine whether those outputs represent a single measurable attribute. As a consequence, the distinction between measurement and aggregation becomes blurred, and multidimensional numerical summaries acquire the appearance of quantitative evidence.

Without unidimensionality or its equivalent, students are not equipped to challenge the assumptions that underpin the reference case.

Statement 7: A Manifest Attribute is Directly Observable and Capable of Supporting Empirical Observation

The endorsement probabilities of 0.50 to 0.55 place this statement among the more strongly supported concepts in the curriculum profile. This is not surprising. Manifest attributes represent the most straightforward case for measurement because they are directly observable. Examples include mortality, survival time, hospital admissions, emergency department visits, treatment persistence, healthcare utilization, doses administered and resource consumption. These outcomes can be observed, counted and recorded directly without requiring a measurement model.

Yet the results remain disappointing. If manifest attributes are the most transparent examples of measurable outcomes, one might reasonably expect endorsement probabilities substantially higher than those observed here. Recognition appears partial rather than comprehensive. Students may encounter observable outcomes in research and evaluation, but there is little evidence that they are systematically taught why manifest attributes occupy a unique position within measurement theory.

This distinction is important because the identification of an attribute as manifest is one of the first steps in determining an appropriate measurement strategy. Before measurement can begin, the nature of the attribute must be understood. Observable outcomes present a fundamentally different scientific problem from latent attributes. Failure to recognize this distinction creates uncertainty regarding the choice of measurement framework and the interpretation of resulting claims.

Manifest attributes occupy a unique position because they provide the most direct route to evaluable, replicable and falsifiable claims. If a therapy is claimed to reduce hospital admissions, decrease emergency department utilization, improve treatment persistence or increase survival time, these claims can be observed directly and subjected to empirical verification. The principal challenge is not the construction of a measurement model but the careful definition of the attribute and the establishment of a valid observation protocol.

The more important issue is what appears to be missing from the curriculum. Although there is some recognition of manifest attributes, there is little corresponding evidence that students are taught how manifest attributes fit within a broader measurement framework. The concept appears in relative isolation rather than as part of a systematic treatment of measurement. Students may therefore recognize observable outcomes without understanding their significance for quantitative claims.

This interpretation is reinforced by the near-complete absence of endorsement for latent attributes, representational measurement and ratio measurement. If students were being taught a coherent theory of measurement, recognition of manifest attributes would normally be accompanied by recognition of the complementary concepts that give the distinction meaning. Instead, Statement 7 stands largely alone within a curriculum profile that otherwise provides little evidence of engagement with measurement science.

The result is conceptual fragmentation. Students may understand that some outcomes are observable while remaining unaware of the consequences that follow from that observation. They may not appreciate that manifest attributes support direct linear ratio measurement, while latent attributes require an entirely different framework centered on the estimation of attribute possession through Rasch measurement. The distinction is recognized but its implications are not.

This finding contributes directly to the broader pattern of measurement inversion. Recognition that an attribute is observable does not in itself establish a framework for measurement. Without an understanding of how manifest attributes differ from latent attributes, and why these differences matter, students remain ill-equipped to evaluate the measurement foundations of quantitative claims.

Without a clear distinction between manifest and latent attributes students are not equipped to challenge the assumptions that underpin the reference case.

Statement 8: A Latent Attribute is Not Directly Observable and Requires a Measurement Model to Estimate Possession of the Attribute

The endorsement probabilities of only $p = 0.15$ across all US research centers suggest that the concept of latent attributes occupies at best a marginal position within HTA curricula. This finding is particularly important because it reveals a profound gap in the understanding of measurement itself. While manifest attributes can be observed directly, latent attributes cannot. They are not visible phenomena. They cannot be counted, weighed, timed or observed in the same way as hospital admissions, survival time or healthcare utilization. Their existence must be inferred from observable indicators.

This distinction is fundamental. A latent attribute is not simply an outcome that happens to be difficult to measure. It is an entirely different type of measurement problem. Attributes such as symptom burden, functional status, fatigue, treatment satisfaction, need fulfilment, anxiety, depression and quality of life cannot be observed directly. What is observed are responses to questions, behaviors, or other indicators that may reflect possession of the underlying attribute. The task of measurement is therefore not to summarize responses but to estimate the extent to which an individual possesses the latent attribute itself. A problem that was resolved over 60 years ago.

The curriculum results suggest that students are rarely introduced to this concept. If the attribute-possession problem is not recognized, there is little possibility that students will understand why latent attributes require a measurement model. Numerical scores derived from questionnaires are likely to be treated as though they were measurements in their own right. The distinction between a score and a measure disappears.

This omission has far-reaching implications. Once latent attributes are treated simply as collections of questionnaire responses, the focus shifts from measurement to arithmetic. Responses are summed, averaged, weighted or transformed without first asking whether the resulting quantity represents a measurable attribute. Numerical manipulation substitutes for measurement. The resulting scores acquire scientific authority merely because they are numerical. They are not ratio measures.

The importance of the attribute-possession concept cannot be overstated. In measurement, the central question is not "What score did the respondent achieve?" but rather "How much of the attribute does the respondent possess?" This is the question that underlies all legitimate latent measurement. Possession is the object of measurement. Responses are merely evidence from which possession may be inferred.

This distinction separates measurement from psychometric scoring. A score is simply a numerical summary of responses. A measure is an estimate of the extent to which an individual possesses an attribute. The two are not equivalent. Yet there is little evidence that US HTA students and faculty are systematically exposed to this distinction.

Rasch measurement begins with the proposition that a latent attribute exists and that individuals possess differing amounts of that attribute. The purpose of the Rasch model is to estimate possession through the interaction of persons and items. The resulting measure is not a score but an estimate of attribute possession expressed on an invariant scale.

Without an understanding of latent attributes, the rationale for Rasch measurement becomes invisible. Students are unlikely to understand why Rasch insists upon unidimensionality, why item fit matters, why invariance is essential, or why measurement must be separated from the simple aggregation of responses. Rasch appears as a technical option among many rather than as a unique measurement framework addressing a fundamentally different scientific problem. The necessary and sufficient nature of Rasch measurement is never addressed.

The implications extend directly to contemporary HTA. Many outcomes central to therapy assessment are latent attributes. Patient-centered outcomes, symptoms, functioning, burden of illness and treatment experience are not directly observable phenomena. They are latent constructs requiring measurement. Yet if students are not taught the latent attribute problem, they are unlikely to question whether the instruments used to assess these outcomes actually generate measures.

The result is predictable. Patient-centered outcomes become numerical objects rather than measurement problems. Questionnaires become generators of scores rather than instruments designed to estimate possession of an attribute. Numerical summaries are accepted as evidence of therapy impact without establishing whether measurement has occurred. The scientific challenge of constructing a measure is replaced by the administrative task of generating a score.

This omission contributes directly to measurement inversion. Students are taught to analyze these numbers without first being taught what those numbers represent. They are introduced to outcomes research, economic evaluation and statistical analysis while remaining largely unaware of the distinction between a score and a measure. As a consequence, numerical outputs acquire legitimacy simply because they are numerical.

The irony is striking. Many of the most important outcomes in healthcare are latent attributes, yet the educational framework appears to devote little attention to the scientific principles required to measure them. The consequence is that the central measurement challenge confronting patient-centered assessment remains largely invisible.

Without an understanding of latent attributes, attribute possession and the need for a measurement model to estimate possession, or an equivalent framework, students are not equipped to challenge the assumptions that underpin the reference case.

Statement 9: Manifest and Latent Attributes Require Different Forms of Ratio Measurement

Every institution scores 0.05.

This finding should be regarded as catastrophic. Ratio measurement is not an advanced methodological refinement. It is the highest form of quantitative measurement and the only scale supporting multiplication, division and proportional reasoning. Yet there is virtually no evidence that US HTA students are taught what a ratio measure is, how ratio properties are established, why a non-arbitrary zero is required, or why arithmetic operations require ratio-scale quantities.

More importantly, there is no evidence that students are taught that the route to ratio measurement depends upon the nature of the attribute itself. This is a critical omission. Manifest and latent attributes present fundamentally different measurement problems and therefore require fundamentally different measurement solutions.

For manifest attributes that are directly observable, measurement proceeds through conventional linear ratio scales. Attributes such as survival time, hospital days, treatment persistence, doses administered or resource utilization can support direct empirical observation. If the requirements of measurement are met, these attributes may be represented through linear ratio measures with a meaningful zero and proportional interpretation.

Latent attributes are entirely different. They are not directly observable. Attributes such as symptom burden, functional status, treatment satisfaction, need fulfilment and quality of life cannot be measured through direct observation. Their measurement requires a model capable of estimating possession of the attribute. The appropriate solution is Rasch measurement, where the resulting logit scale provides a ratio measure of latent attribute possession. The distinction is fundamental. Manifest attributes require linear ratio measurement. Latent attributes require Rasch logit ratio measurement.

The curriculum results suggest that students are exposed to neither framework. They are not taught how to establish ratio properties for manifest attributes. They are not taught how Rasch measurement constructs ratio measures for latent attributes. Consequently, they never encounter the only two legitimate pathways to quantitative assessment.

This omission goes directly to the heart of contemporary HTA. A student unfamiliar with ratio measurement has no reason to ask whether quantities satisfy the conditions necessary for multiplication. A student unfamiliar with latent measurement has no reason to ask whether a latent attribute requires a measurement model. A student unfamiliar with Rasch measurement has no reason to question numerical claims based on instruments that have never demonstrated measurement properties.

The result is predictable. Quantities are accepted as measures because they are presented as measures. Numerical scores are accepted as proportions because they are presented as proportions. Arithmetic is accepted because it is conventional. The scientific question of whether measurement has actually been achieved never arises.

Without an understanding of linear ratio measurement for manifest attributes and Rasch logit ratio measurement for latent attributes, or an equivalent framework, students are not equipped to challenge the assumptions that underpin the reference case. An unfortunate implication is that teaching faculty are also unaware of this fundamental requirement.

Statement 10: Therapy impact claims must be falsifiable

The relatively strong endorsement of falsifiability (p=0.80 to 0.85) is one of the most revealing findings in the assessment. Students are taught that scientific claims should be evaluated and subjected to empirical scrutiny. This reflects an important scientific instinct and aligns with the Popperian view that scientific progress depends upon the ability to expose claims to possible refutation.

However, the endorsement of falsifiability sits uneasily beside the dominant analytical framework employed in contemporary HTA. The reference case was never designed to generate falsifiable claims. Rather, it was designed to construct estimates of future cost-effectiveness through simulation models populated by assumptions, utility weights, transition probabilities and extrapolated outcomes. The outputs are not predictions that can be directly subjected to empirical testing. They are model projections whose validity depends primarily upon the assumptions built into the model itself.

This creates a fundamental contradiction. Students are taught that scientific claims should be falsifiable, yet the principal evaluative framework presented in HTA generates claims that are not falsifiable. A reference-case model may conclude that a therapy will generate a particular cost per QALY over a lifetime horizon, but there is no practical mechanism by which that claim can subsequently be tested. The future population does not yet exist, the simulated pathway cannot be observed, and the model assumptions cannot be independently verified as a coherent whole. The claim survives not because it has withstood empirical challenge but because it remains embedded within a simulation. The fact that the QALY is an impossible mathematical construct is not an issue; it is accepted as a necessary step without question.

The contradiction becomes even more pronounced when measurement is considered. Falsification and measurement are complementary components of science. Measurement establishes the quantities entering a claim. Falsification evaluates the claim itself. One cannot substitute for the other. A claim cannot become scientific merely because it is expressed numerically, nor can a simulation become scientific merely because it is mathematically sophisticated.

The curriculum profile suggests that students are taught the language of scientific evaluation while receiving little exposure to either the science of measurement or the practical limitations of simulation-based claims. Claims are evaluated in principle, but the measurement foundations of those claims are rarely examined. Equally important, students are not encouraged to ask whether the outputs of reference-case models are capable of empirical refutation in the first place.

The result is a form of incomplete science in which falsification survives as an aspiration while measurement and evaluability disappear from practice. Students learn that science requires testing, yet are introduced to a framework whose principal outputs are not designed to be tested against future observation. The consequence is predictable. The reference case is accepted as scientific despite operating outside the normal standards of scientific inquiry.

Without an understanding of measurement, evaluability and falsifiability as integrated requirements for quantitative claims, or an equivalent framework, students are not equipped to challenge the assumptions that underpin the reference case.

OVERVIEW OF US RESEARCH CENTER CURRICULUM RESULTS

The curriculum interrogation of seven leading US health technology assessment research centers reveals a striking and highly consistent pattern. Although these institutions differ in history, organizational structure, research priorities and geographical location, their endorsement profiles are virtually identical; all demonstrate the same underlying intellectual structure. The consistency of these results is itself an important finding. It suggests that the observed deficiencies do not reflect local curriculum choices or isolated omissions but rather a common educational framework that has become institutionalized throughout US HTA research and training.

At first sight the results appear reassuring. The centers strongly endorse the proposition that therapy assessment begins with the identification and specification of a target attribute. Endorsement probabilities for these statements range from 0.80 to 0.85 across all institutions. Similarly, the principle that therapy impact claims should be falsifiable receives strong endorsement. These findings suggest that students and researchers are introduced to basic concepts of scientific inquiry, including the need for clearly defined outcomes and the importance of empirical evaluation.

The reassuring appearance is short-lived. Beyond these introductory concepts, the curriculum profile collapses. Endorsement of the proposition that scales of measurement possess different properties and support different forms of analysis falls uniformly to 0.20. The principle that the measurement status of an attribute must be established before quantitative claims can be advanced falls further to 0.15. Endorsement of representational measurement, arguably the most important scientific foundation for quantitative reasoning, falls to 0.05. Unidimensionality receives an endorsement probability of only 0.10, while latent attribute measurement receives only 0.15. The proposition that manifest and latent attributes require different forms of ratio measurement receives a uniform probability of 0.05 across every institution.

These findings point to a curriculum structure in which the language of scientific evaluation is present but the science of measurement is largely absent. Students are taught to identify outcomes, evaluate interventions and construct analytical arguments, yet there is little evidence that they are systematically exposed to the principles that determine whether quantitative claims are possible. The distinction is fundamental. Scientific inquiry requires more than the ability to perform analyses. It requires an understanding of the conditions under which those analyses are legitimate.

The interrogation therefore reveals two distinct but related deficiencies. The first is one of omission. Concepts central to measurement science, including representational measurement, scale theory, unidimensionality, latent attribute measurement and ratio measurement, occupy only a marginal position within the curriculum knowledge base. Students and researchers are therefore unlikely to encounter the conceptual framework necessary to evaluate the measurement status of the quantities employed in HTA.

The second deficiency is more serious. It is not merely that essential concepts are absent. It is that the educational environment simultaneously promotes analytical frameworks that depend upon precisely those concepts. Utilities, QALYs and reference-case simulation models all require assumptions concerning scale properties, measurement status and the legitimacy of arithmetic operations. Yet the curriculum profile suggests that the concepts required to evaluate those assumptions are largely absent. Students are therefore exposed to analytical frameworks without being equipped to determine whether those frameworks satisfy the standards of representational measurement and normal scientific inquiry.

This finding provides perhaps the clearest explanation yet for the persistence of measurement inversion within HTA. Measurement inversion occurs when arithmetic precedes measurement; when quantities are manipulated mathematically before it has been demonstrated that they satisfy the requirements of measurement. The curriculum results suggest that this inversion is not merely a feature of HTA practice. It is reproduced through HTA education. Researchers are trained to apply established methods without first being introduced to the scientific principles necessary to evaluate those methods.

The implications extend beyond curriculum design. Research centers occupy a central position in shaping future HTA practitioners, policy analysts and academic researchers. If the science of measurement is absent from the educational environment, its absence from research practice should not be surprising. The interrogation results suggest that measurement inversion and curriculum inversion are mutually reinforcing phenomena. One helps explain the persistence of the other.

The conclusion is therefore difficult to avoid. The problem confronting HTA is not simply methodological. It is educational. Reconstruction of HTA practice will require reconstruction of HTA education. Until students and researchers are systematically introduced to representational measurement, unidimensionality, manifest and latent attributes, ratio measurement and falsifiable claims, the conditions that produced measurement inversion will continue to be reproduced throughout the discipline. The challenge is not incremental curriculum reform. It is the reconstruction of HTA education around the standards of measurement that underpin every other quantitative science.

DOES HEALTH TECHNOLOGY ASSESSMENT HAVE A FUTURE

Given the conclusion presented here that the pervasive pattern of measurement inversion in HTA is attributable, at least in part, to curriculum inversion, the obvious question is whether HTA as presently constituted has a future. The answer depends entirely upon whether the discipline is prepared to confront the implications of measurement. Defending the status quo becomes increasingly difficult once the standards of representational measurement are introduced. The problem is not that current methods require refinement or methodological adjustment. The problem is that the foundational assumptions upon which the reference case rests fail to satisfy the requirements for quantitative claims.

The most obvious example is the construction of utilities and the subsequent development of the reference case. Once it is recognized that multiplication and division require ratio measures, the burden of proof shifts immediately to those who claim that a quantity can function as a proportion. In the reference case, time is adjusted by a discount factor that is assumed to represent the value of a health state. Yet nowhere is a case made that this discount factor possesses ratio measurement properties. The requirement itself is not even raised. More importantly, the failure does not begin at the point where the quantity is applied. It begins much earlier in the process of valuing health states. At every stage of development, from health-state descriptions, to valuation exercises, to utility algorithms, to the quantities entering simulation models, the requirements of ratio measurement are absent. Once this is recognized, the conclusion is unavoidable. Utilities, QALYs and reference-case simulation models cannot be regarded as scientifically credible measures of therapy impact. They have no place in a discipline committed to quantitative evidence.

This does not mean that HTA itself has no future. On the contrary, the need for evidence-based assessment of therapeutic interventions has never been greater. Health systems must make decisions regarding pricing, access and resource allocation. Manufacturers must demonstrate product value. Clinicians require evidence of treatment impact. Patients deserve assessments that are transparent, evaluable and scientifically defensible. The question is therefore not whether HTA should survive, but whether it can survive in its present form.

The answer is almost certainly no. The curriculum assessments and measurement inversion studies reported here suggest that the intellectual foundations of the reference case are impossible to defend once subjected to measurement scrutiny. A framework built upon quantities that fail the requirements of measurement cannot be rescued through more sophisticated modelling, larger databases or increasingly elaborate statistical techniques. Arithmetic cannot create measurement. Numerical complexity cannot compensate for measurement failure.

The future of HTA therefore lies in reconstruction rather than reform. The starting point is straightforward and entirely consistent with the standards accepted throughout the sciences. First, define the attribute of interest. Second, determine whether the attribute is manifest or latent. If the attribute is manifest and directly observable, then traditional linear ratio measurement techniques may be employed. If the attribute is latent and not directly observable, then a measurement model is required. The Rasch logit ratio framework provides the necessary and sufficient basis for constructing measures of latent attribute possession. Once a valid Rasch logit ratio measure is established, the same principles of empirical assessment, hypothesis testing and falsification can be applied.

This is not a difficult transition conceptually. The required measurement tools already exist. Representational measurement has been available for decades. Rasch measurement has demonstrated how latent attributes can be measured through invariant scales for 60 years; made easier with the availability of Rasch instrument software programs since the mid-1980s. Scientific protocols for evaluable and replicable claims are well understood. What is required is not methodological invention but intellectual realignment. Measurement must once again precede arithmetic.

Whether HTA has a future therefore depends upon its willingness to abandon the assumptions that produced measurement inversion and replace them with the standards that govern quantitative inquiry elsewhere in science. If the discipline remains committed to the reference case, its lack of a future is not in doubt. Indeed, it never had a past.

If, however, HTA embraces representational measurement and reconstructs its curriculum and practice around lawful measures, then it has every opportunity to emerge as a genuine science of therapy assessment.

CONCLUSION: FROM CRITICISM TO RECONSTRUCTION

In anticipation of the need for a fundamental reconstruction of health technology assessment, Maimon Research has developed a concise nine-unit educational program designed to provide the essential conceptual and methodological foundations for a measurement-based HTA framework of analysis. The purpose of the program is not to criticize existing practice, but to provide a practical pathway for transition from the current reference-case paradigm to one grounded in the standards of representational measurement.

The starting point is recognition that the central weakness of contemporary HTA is not a lack of analytical sophistication but the absence of measurement. For five decades, students, researchers and practitioners have been trained to apply increasingly complex economic and simulation techniques without first establishing whether the quantities entering those analyses satisfy the requirements of measurement. The result is a discipline characterized by measurement inversion, where arithmetic precedes measurement and numerical manipulation is mistaken for quantitative evidence. The reconstruction program is designed for reconstruction.

The first units introduce the concepts of attributes and measurement. Participants are guided through the process of identifying the attribute of interest in a therapy assessment and distinguishing between attributes that are directly observable and those that are latent and inferred. Particular emphasis is placed on the principle that every evaluable claim must begin with specification of a clearly defined attribute. Without an attribute there can be no measurement; without measurement there can be no meaningful quantitative claim.

Subsequent units address the scales of measurement and the axioms of representational measurement. Participants are introduced to the scientific foundations that govern all quantitative inquiry: nominal, ordinal, interval and ratio scales; admissible transformations; dimensional homogeneity; unidimensionality; and the requirement that measurement precede arithmetic. These concepts are rarely encountered in contemporary HTA curricula despite being fundamental to every other quantitative discipline. Understanding these principles allows participants to evaluate whether proposed quantities genuinely support arithmetic operations or merely provide numerical descriptions.

Particular attention is given to interval and ratio measurement because these scales support the quantitative claims required for scientific assessment. Participants learn why ratio measurement requires a single attribute and a non-arbitrary zero, why multiplication and division demand ratio properties, and why the burden of proof rests upon those who claim that a quantity can function as a proportion. These discussions lead naturally to an examination of the measurement assumptions embedded within conventional HTA methodologies.

The program then turns to latent attributes and Rasch measurement. This component provides for the first time a practical introduction to the construction of measures of latent attribute possession. Participants are introduced to the Rasch framework, item development, attribute specification, response structures, invariance, unidimensionality, fit assessment and the construction of logit-based ratio measures. The objective is not statistical sophistication for its own sake, but the creation of defensible measures capable of supporting evaluable claims regarding therapy impact.

The final units focus on application. Participants are introduced to a reconstructed formulary assessment framework in which therapy evaluation begins with attribute identification and measurement rather than with simulation modelling. Particular attention is given to the role of protocols in supporting evaluable, replicable and falsifiable claims. Students learn how to construct assessment protocols, specify success criteria, define timelines, identify evidence requirements and establish the conditions under which claims may be challenged or rejected. The emphasis throughout is on transparency, replication and empirical evaluation.

Taken together, the nine units provide a coherent framework for rebuilding HTA on scientific foundations. Rather than relying upon assumed measurement properties and speculative long-term simulations, the program offers a practical methodology for generating credible quantitative claims regarding therapy impact. It equips participants with the conceptual tools needed to recognize measurement inversion, understand why it emerged, and replace it with an approach grounded in representational measurement, ratio measurement and falsifiable evidence. In this sense, the program is intended not merely as an educational exercise but as a blueprint for the transition from numerical storytelling to a genuine science of therapy assessment.

Further details regarding the reconstruction program are available from the Maimon Research website: https://maimonresearch.com/hta-reconstruction-program-and-fees/.

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

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

[3] The Theory of Allowable Transformations and the QALY. Logit Working Paper No 887. https://maimonresearch.com/logit-working-paper-no-887-june-2026/

[4] Bond T, Zi Yan, Heene M. Applying the Rasch Model: Fundamental Measurement in the Human Sciences (4th Ed). New York: Routledge, 2021

[5] Maimon Research LLC. Interrogation Reports. www.maimonresearch.com

[6] United States: Closed Epistemic Reproduction – How academic HTA knowledge bases in the United States sustain measurement inversion. Logit Working Paper No 757. file:///C:/Users/Paul/Downloads/Logit-Working-Paper-No-757-April-2026%20(3).pdf

[7] United States: The absence of representational measurement and the Schaeffer Center for Health Policy and Economics, University of Southern California. Logit Working Paper 750. https://maimonresearch.com/logit-working-paper-no-750-april-2026/

[8] United States: The absence of representational measurement and the Center for Health Decision Science, Harvard University. Logit Working Paper No 751. https://maimonresearch.com/logit-working-paper-no-751-april-2026/

[9] United States: The absence of representational measurement and the University of Washington CHOICE Institute. Logit Working Paper No 752. https://maimonresearch.com/logit-working-paper-no-752-april-2026/

[10] United States: The absence of representational measurement and the Duke-Margolis Institute for Health Policy, Duke University. Logit Working Paper No 753. https://maimonresearch.com/logit-working-paper-no-753-april-2026/

[11] United States: The absence of representational measurement and the Leonard Davis Institute of Health Economics, University of Pennsylvania. Logit Working Paper No 754. https://maimonresearch.com/logit-working-paper-no-754-april-2026/

[12] United States: The absence of representational measurement and the Health Economics Group, Bloomberg School of Public Health. Johns Hopkins University. Logit Working Paper No 756. https://maimonresearch.com/logit-working-paper-no-756-april-2026/

[13] United States: The absence of representational measurement and the Stanford University Health Policy Center. Logit working Paper No 755. https://maimonresearch.com/logit-working-paper-no-755-april-2026/