

RESEARCH LLC

**ARTIFICIAL INTELLIGENCE LARGE LANGUAGE
MODEL INTERROGATION**



**REPRESENTATIONAL MEASUREMENT FAILURE IN
HEALTH TECHNOLOGY ASSESSMENT**

**UNITED KINGDOM: CURRICULUM INVERSION AND
PARADIGM FAILURE IN HEALTH TECHNOLOGY
ASSESSMENT**

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

LOGIT WORKING PAPER No 585 JUNE 2026

www.maimonresearch.com

Tucson AZ

INTRODUCTION

The starting point for any reconstruction of health technology assessment (HTA) in the UK must be an understanding of the scale and consistency of the false 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, 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. A recent assessment of measurement inversion in the same research centers that are assessed here concludes it is a major failure in UK HTA ¹

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 in Canada 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 large language model (LLM) interrogations reveal this inversion repeatedly ². 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 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. Canada is no exception. 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 possible for a professional career to pass without knowledge of the scales of measurement or the axioms of representational measurement ^{3 4} .

It is worth noting the complete absence of Rasch measurement from the curriculum profiles of the Canada HTA research centers ⁵. 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 faculty and 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.

The result is that one of the most important developments in measurement theory over the past 80 years is effectively invisible within Canada HTA education. Without exposure to Rasch measurement, faculty and 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 and the unique contribution of ratio measurement 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 ⁶

STANDARDS FOR MEASUREMENT

The starting point for any scientific discipline that seeks to make quantitative claims is measurement. Before quantities can be manipulated mathematically, it must first be demonstrated that they possess the properties necessary to support the proposed arithmetic operations. This principle is fundamental to both the physical and social sciences. Measurement precedes arithmetic. Quantitative claims are valid only when the quantities involved satisfy the requirements of measurement. If these requirements are absent, arithmetic operations may still be performed, but the resulting outputs have no scientific standing as measures.

The importance of this principle is reflected in the theory of measurement scales. Not all numerical assignments possess the same properties. Nominal scales classify. Ordinal scales rank. Interval scales support differences between values. Ratio scales alone support the full range of arithmetic operations because they possess a true zero and permit proportional comparisons. Consequently, the admissibility of arithmetic depends upon scale type. Addition and subtraction require at least interval properties. Multiplication and division require ratio properties. This is not a matter of convention. It is a requirement imposed by the structure of measurement itself.

The central importance of ratio measurement follows directly from these considerations. Any claim involving multiplication, division, proportional comparison, growth rates, averages of ratios, or cost-effectiveness ratios requires quantities that possess ratio properties. If ratio measurement has not been demonstrated, these operations are inadmissible. Numerical manipulation cannot create measurement properties that are absent from the underlying scale. Arithmetic cannot substitute for measurement.

These requirements are formalized in the axioms of representational measurement. Representational measurement provides the scientific framework that links empirical observations to numerical representations. Its purpose is to ensure that numerical assignments preserve the structure of the attribute being measured. Only when this correspondence is demonstrated can arithmetic operations be regarded as meaningful. The axioms of representational measurement therefore establish the conditions under which quantitative claims can be considered scientifically legitimate.

Among the most important of these requirements is unidimensionality. Measurement requires that an attribute represent a single dimension. If multiple attributes are combined into a composite score, numerical aggregation may be possible, but measurement has not necessarily occurred. Without unidimensionality there is no assurance that a numerical value represents a coherent quantity. The distinction between aggregation and measurement is therefore fundamental. Numbers can always be combined. Measures cannot be assumed.

Equally important is the distinction between manifest and latent attributes. Manifest attributes are directly observable and, where appropriately specified, support linear ratio measurement. Latent attributes are not directly observable and require a measurement model capable of estimating possession of the attribute. In the latter case, the required measure is the Rasch logit ratio scale. These two forms of ratio measurement, linear ratio measurement for manifest attributes and Rasch logit ratio measurement for latent attributes, provide the only scientifically defensible basis for quantitative claims regarding therapy impact.

Taken together, these principles establish a clear standard. Measurement must precede arithmetic. Scale properties determine admissible operations. Ratio measurement is required wherever proportional comparisons or multiplication are involved. Unidimensionality must be demonstrated before measurement can be claimed. Representational measurement provides the governing scientific framework. Any discipline seeking to generate quantitative claims must satisfy these requirements. Without them, numerical outputs remain constructions rather than measures, and quantitative claims become matters of assumption rather than science.

CURRICULUM EXPECTATIONS

A curriculum intended to prepare students for health technology assessment should provide more than instruction in economic evaluation, evidence synthesis, utility assessment, and simulation modelling. It should equip students with the conceptual tools necessary to evaluate the scientific legitimacy of quantitative claims.

At a minimum, students should be introduced to the role of attributes in therapy assessment and the requirement that the target attribute be specified before evaluation begins. They should understand the principal scales of measurement and the implications these have for admissible arithmetic. They should be familiar with the distinction between nominal, ordinal, interval, and ratio scales and appreciate that different scales support different forms of analysis.

A scientifically grounded curriculum should also introduce students to representational measurement and its role in determining whether numerical assignments constitute valid measures. This includes an understanding of the relationship between measurement and arithmetic, the importance of unidimensionality, and the conditions required for quantitative claims.

Students should further be expected to recognize the distinction between manifest and latent attributes. Observable outcomes such as survival, hospital admissions, adverse events, and treatment persistence present different measurement challenges from latent constructs such as pain, fatigue, quality of life, functioning, and patient satisfaction. A curriculum should therefore provide an understanding of the different measurement frameworks required for these two classes of attributes.

Finally, students should be exposed to the principles of empirical evaluation and falsification. Quantitative claims should not be treated as self-validating simply because they are numerical. Rather, students should understand that scientific claims must be capable of evaluation, replication, and potential refutation through observation and evidence.

These expectations do not represent advanced or specialist topics. They constitute the foundational knowledge required for any discipline that seeks to support quantitative claims regarding therapy impact. Their presence would indicate a curriculum grounded in measurement and scientific inquiry. Their absence would suggest that students are being trained to apply analytical techniques without first acquiring the conceptual framework necessary to evaluate whether the resulting claims are scientifically defensible.

CURRICULUM INVERSION

Curriculum inversion occurs when a curriculum teaches the application of quantitative methods while failing to teach the measurement principles that determine whether those methods are scientifically legitimate. In a scientifically coherent curriculum, measurement precedes arithmetic. Students first learn the nature of attributes, the requirements of representational measurement, the distinctions among nominal, ordinal, interval, and ratio scales, and the conditions necessary for valid quantitative claims. Only then are they introduced to the arithmetic, statistical, and modelling procedures that depend upon those measurement properties. Curriculum inversion reverses this sequence. Students learn how to calculate, model, and analyze before they learn how to determine whether the quantities entering those analyses are measures. Arithmetic becomes detached from measurement, and numerical manipulation is treated as though it were equivalent to quantitative science.

The consequences are profound. A curriculum affected by inversion reproduces a professional culture in which measurement is assumed rather than demonstrated. Concepts such as

unidimensionality, dimensional homogeneity, admissible arithmetic, manifest and latent attributes, ratio measurement, and Rasch measurement either disappear entirely or are treated as peripheral concerns. Students become proficient in the techniques of economic evaluation, utility assessment, QALY construction, and simulation modelling without acquiring the conceptual tools necessary to evaluate the scientific legitimacy of those methods. The result is that the curriculum not only fails to identify measurement errors but actively reproduces them across successive generations of researchers, analysts, reviewers, and decision makers. Curriculum inversion therefore becomes the educational mechanism through which measurement inversion is institutionalized within a discipline.

The importance of curriculum inversion extends beyond education. Universities and research centers are responsible for training the individuals who subsequently populate HTA agencies, reimbursement committees, consulting organizations, professional societies, and academic departments. If the concepts necessary to evaluate quantitative claims are absent from the curriculum, they will also be absent from research, policy, and practice. The omissions identified in educational programs therefore become embedded throughout the wider HTA environment. What begins as a curriculum deficiency ultimately becomes a disciplinary norm.

For this reason, curriculum assessment is a critical component of any evaluation of the scientific standing of HTA. 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 evaluation of those methods possible. A curriculum that emphasizes modelling, economic evaluation, utility assessment, and decision analysis while neglecting measurement theory will inevitably reproduce the same conceptual limitations observed in current HTA practice.

This is why curriculum inversion provides an important explanation for the persistence of the reference-case paradigm. Utilities, QALYs, cost-effectiveness ratios, and simulation models have survived not because their measurement foundations were demonstrated, but because the concepts required to challenge those foundations were largely absent from the educational environment. Curriculum inversion reproduces measurement inversion. Together they explain how a framework built on arithmetic before measurement could become established, institutionalized, and defended for more than four decades.

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. A knowledge base is more than a collection of documents. It represents the accumulated concepts, assumptions, methods, definitions, and analytical frameworks that are communicated through educational and research activities. It is the intellectual environment within which students, researchers, and professional staff acquire their understanding of a discipline.

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 and reinforced. Taken together, these materials provide a representation of the concepts and analytical frameworks that students and researchers are most likely to encounter during their exposure to the institution.

The importance of defining the curriculum knowledge base extends beyond simple course content. Curricula do more than transfer technical skills. They establish what is regarded as important knowledge, what questions are considered relevant, what assumptions are accepted, and what analytical approaches are treated as legitimate. Over time, these priorities shape the intellectual culture of a discipline. Concepts that are consistently emphasized become embedded in professional practice. Concepts that are neglected gradually disappear from the educational environment and, ultimately, from the research and policy communities that the curriculum supports.

A curriculum designed to support scientific evaluation of therapeutic interventions would be expected to expose students to the foundational principles required for quantitative reasoning. These include the specification of attributes, the theory of measurement scales, representational measurement, admissible arithmetic, unidimensionality, dimensional homogeneity, the distinction between manifest and latent attributes, ratio measurement, Rasch measurement, and the requirement that quantitative claims be capable of empirical evaluation and potential falsification. These concepts are not specialist additions to HTA. They are the intellectual foundations upon which all quantitative claims ultimately depend.

The objective of the present interrogation is therefore not to evaluate individual courses, faculty members, or teaching styles. Rather, it is to assess the broader intellectual environment represented by the curriculum knowledge base. The question is straightforward: to what extent are the concepts required for a measurement-based approach to therapy assessment present, absent, or weakly represented within the educational and research environment of the institution? The answer provides an indication of whether future researchers, analysts, and decision makers are likely to encounter the conceptual foundations necessary to evaluate quantitative claims or whether they are being trained primarily in the application of analytical techniques. In this respect, the curriculum knowledge base provides a window into the intellectual foundations of the discipline itself.

INTERROGATING THE CURRICULUM KNOWLEDGE BASE

The purpose of the present interrogation is not to determine what a large language model "believes." Large language models possess no beliefs, opinions, or independent understanding. Their value lies in their ability to interrogate and summarize the conceptual content of a defined knowledge base. The objective is therefore to examine the intellectual environment represented within a curriculum and to determine which concepts are present, absent, weakly represented, or strongly reinforced.

In the context of curriculum assessment, the knowledge base is defined as the totality of publicly accessible materials through which a research center communicates educational and methodological content. These materials include curriculum descriptions, course outlines, learning objectives, training programs, methodological guidance documents, workshop and seminar content, faculty publications, doctoral training resources, technical reports, conference presentations, and other educational materials associated with the institution. Taken together, these sources provide a representation of the concepts and analytical frameworks that students, researchers, and professional staff are most likely to encounter.

The interrogation therefore focuses not on individual instructors or specific courses but on the broader intellectual environment that the curriculum creates and sustains. A curriculum functions as a mechanism for transmitting knowledge across successive generations of students and researchers. Concepts that are consistently emphasized become embedded within professional practice. Concepts that are absent or weakly represented are unlikely to become part of the analytical framework employed by future practitioners. The curriculum knowledge base therefore provides an important indicator of the intellectual foundations of a discipline.

For the purposes of the present assessment, the objective is straightforward: to determine the extent to which the curriculum knowledge base recognizes and reinforces the concepts required for a measurement-based approach to health technology assessment. Particular attention is given to attributes, measurement scales, representational measurement, admissible arithmetic, unidimensionality, the distinction between manifest and latent attributes, ratio measurement, Rasch measurement, and the requirement that scientific claims be evaluable and potentially falsifiable. These concepts constitute the foundational framework necessary for evaluating the validity of quantitative claims.

The interrogation proceeds through a series of canonical statements representing concepts that would be expected to appear within a curriculum grounded in measurement science. The resulting endorsement probabilities provide an indication of the extent to which these concepts are embedded within the curriculum knowledge base. High endorsement probabilities suggest that the concept is visible and reinforced within the educational environment. Low endorsement probabilities suggest that the concept is absent, weakly represented, or largely ignored.

The significance of this approach is that it allows the curriculum to be evaluated as a coherent intellectual system rather than as a collection of individual courses. The question is not whether students are taught contemporary HTA methods. The question is whether they are taught the concepts necessary to evaluate the scientific legitimacy of those methods. The resulting profile provides a measure of the extent to which the curriculum supports, neglects, or potentially undermines the foundations of measurement-based scientific inquiry.

TOOLS FOR INTERROGATION

The identification of measurement inversion across HTA research centers, reimbursement agencies and academic programs raises an obvious question: where does this inversion originate?

The objective of curriculum interrogation differs from that of previous HTA knowledge-based practice 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 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 in these UK HTA research centers.

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.

ENDORSEMENT OF CURRICULUM INVERSION

The results of the curriculum interrogation are presented in Table 1 as endorsement probabilities for each of the canonical statements. These probabilities provide an estimate of the extent to which a particular concept is represented within the curriculum knowledge base of each research center. High endorsement probabilities indicate that a concept is visible, reinforced, and likely to be encountered by students and researchers within the educational environment. Low endorsement probabilities indicate that a concept is weakly represented, largely absent, or rarely articulated within the curriculum. The significance of the results lies not in any single statement but in the overall pattern of endorsement. If a curriculum is aligned with the principles of measurement science, concepts such as representational measurement, scale theory, unidimensionality, ratio measurement, admissible arithmetic, manifest and latent attributes, and falsifiable claims would be expected to receive strong endorsement. Conversely, propositions that conflict with these principles would be expected to receive weak endorsement. The resulting profile therefore provides an indication of the extent to which the curriculum knowledge base equips students and researchers with the conceptual foundations necessary to evaluate the scientific legitimacy of quantitative claims in health technology assessment.

TABLE 1: CURRICULUM CONTENT ENDORSEMENT - UK HTA RESEARCH CENTERS

CANONICAL STATEMENT	CANONICAL PROBABILITIES				
	CHE	CRD	HERC	SchHARR	IHS
An attribute is the specific outcome of interest in a therapy assessment	0.85	0.80	0.85	0.85	0.80
Every therapy assessment begins with specification of the target attribute	0.80	0.75	0.80	0.80	0.75
The principal scales of measurement (nominal, ordinal, interval and ratio) have different properties and support different forms of analysis	0.25	0.25	0.25	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	0.10	0.10
The axioms of representational measurement underpin quantitative claims	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
A manifest attribute is directly observable and capable of supporting empirical observation	0.55	0.60	0.65	0.60	0.60
A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute	0.10	0.15	0.15	0.10	0.10
Manifest and latent attributes require different forms of ratio measurement	0.05	0.05	0.05	0.05	0.05
Therapy impact claims must be falsifiable	0.70	0.75	0.80	0.75	0.70

Note: CHE Centre for Health Economics, York; CRD Centre for Reviews and Dissemination; Health Economics Research Centre/Oxford Population Health HTA Group (HERC); School of Health and Related Research (SchHARR), University of Sheffield; Institute of Health and Society, Newcastle HTA Group (IHS)

REVIEW: CENTRE FOR HEALTH ECONOMICS

The profile is striking because it reveals a curriculum strongly oriented toward evaluation but weakly oriented toward measurement. CHE is arguably the intellectual center of the NICE reference-case paradigm. Its teaching and research programs focus heavily on economic evaluation, cost-effectiveness analysis, decision modelling, value-of-information analysis, evidence synthesis, and reimbursement decision-making. CHE staff have played central roles in NICE technology appraisals, methodological development, and HTA training internationally.

As a consequence, the strongest endorsement probabilities are attached to outcome identification and target specification. Students are likely to be taught that therapies should be evaluated against clearly defined outcomes and that assessment requires explicit comparators and decision frameworks. These are standard components of modern HTA. The curriculum therefore performs strongly at the level of defining what is to be evaluated.

The difficulty emerges when the analysis moves from outcomes to measurement. Recognition that different scales possess different analytical properties collapses to a probability of only 0.25. More importantly, the proposition that the measurement status of an attribute must be established before quantitative claims are advanced falls to only 0.15. This is a critical finding because it reverses the normal scientific sequence. In measurement science, the first question is whether a quantity has been measured. In HTA, the first question becomes how the quantity will be incorporated into an economic evaluation.

The lowest endorsement probabilities concern representational measurement, unidimensionality, latent attribute measurement, and the distinction between manifest and latent forms of ratio measurement. These concepts lie at the heart of measurement science. Yet they are largely invisible within the CHE knowledge environment. The probability assigned to representational measurement is only 0.05. The distinction between manifest and latent ratio measures also receives 0.05. The implication is that students can complete advanced training in HTA without being systematically exposed to the conceptual foundations required to evaluate whether the quantities entering economic models possess the properties necessary to support arithmetic operations.

Particularly revealing is the treatment of latent attributes. CHE has been central to the development and dissemination of utility-based evaluation, cost-effectiveness analysis, and QALY methodology. Yet there is little evidence that latent attributes are treated as a distinct measurement problem requiring a dedicated measurement framework. Instead, latent outcomes such as quality of life are incorporated into utility systems and economic models without first addressing the question of measurement. This helps explain the near absence of Rasch measurement within the CHE knowledge base. Once latent attributes are not recognized as a unique measurement challenge, there is little reason to introduce the framework required to construct invariant measures of latent attribute possession.

The result is a classic example of curriculum inversion. Students are introduced to economic evaluation, decision modelling, value-of-information analysis, cost-effectiveness thresholds, NICE appraisal methods, and advanced HTA techniques before they are introduced to the principles necessary to determine whether the quantities entering those analyses qualify as measures. Arithmetic therefore precedes measurement in education exactly as it precedes measurement in HTA practice.

This finding mirrors the pattern observed previously in Australia, Canada, and New Zealand. In each case, HTA centers demonstrate strong recognition of evaluation frameworks but weak recognition of measurement science. The consistency is particularly noteworthy because CHE is not a peripheral institution. It is one of the most influential HTA centers in the world and has played a major role in shaping NICE methodology and international HTA practice.

Consequently, the significance of the CHE interrogation extends beyond a single university. If curriculum inversion is present at the intellectual centers of the UK reference-case paradigm, then it provides a plausible explanation for the persistence of measurement inversion throughout UK HTA. Students are trained to build models, estimate cost-effectiveness ratios, and support reimbursement decisions. They are not trained to determine whether utilities, QALYs, and model outputs satisfy the standards required for measurement. The concepts necessary to challenge the paradigm are therefore largely absent from the educational environment responsible for reproducing it.

The conclusion is straightforward. CHE York does not merely exhibit isolated omissions in curriculum content. Rather, it displays the characteristic profile of curriculum inversion observed throughout contemporary HTA. Measurement science occupies a peripheral position, while economic evaluation occupies the center. The result is a curriculum that prepares students to operate within the reference-case framework but provides limited opportunity to evaluate its scientific legitimacy. In this sense, curriculum inversion becomes the educational mechanism through which measurement inversion is reproduced across successive generations of HTA researchers, analysts, and policy advisers.

REVIEW: CENTRE FOR REVIEWS AND DISSEMINATION

The Centre for Reviews and Dissemination (CRD) occupies a distinctive position within the United Kingdom HTA environment. Unlike centers whose primary focus is economic evaluation and decision modelling, CRD has built its reputation around evidence synthesis, systematic reviews, technology assessment reviews, and the dissemination of evidence to support healthcare decision making. Its role within UK HTA has therefore been to determine what the evidence shows and how that evidence should be assembled to inform policy and reimbursement decisions.

The interrogation results reveal a curriculum profile that is strongly aligned with this mission. High endorsement probabilities are observed for outcome identification, target specification, manifest attributes, and falsifiable claims. The proposition that an attribute is the outcome of interest in a therapy assessment receives a probability of 0.80, while specification of the target attribute receives 0.75. Recognition of manifest attributes receives 0.60 and falsifiable claims 0.75. These findings indicate a curriculum environment that places considerable emphasis on identifying outcomes, defining review questions, evaluating evidence, and drawing conclusions from empirical studies.

However, the profile changes dramatically when attention shifts from evidence to measurement. Recognition of scale theory falls to 0.25. The proposition that the measurement status of an attribute must be established before quantitative claims can be advanced falls to 0.15. Most strikingly, representational measurement receives a probability of only 0.05 across the knowledge base. Unidimensionality receives 0.10, latent attribute measurement 0.15, and the proposition that manifest and latent attributes require different forms of ratio measurement only 0.05.

These results reveal a curriculum environment in which evidence synthesis is emphasized while measurement science remains largely absent. This distinction is critical. A systematic review can identify studies, assess methodological quality, summarize findings, and combine evidence across

investigations. What it cannot do is establish the measurement status of the quantities being reviewed. Measurement is logically prior to evidence synthesis. Before studies can be synthesized, the quantities entering those studies must first satisfy the requirements of measurement.

This is where the interrogation findings become particularly significant. The low endorsement of representational measurement suggests little recognition that quantitative claims depend upon the axioms of measurement. The similarly low endorsement of scale theory suggests limited attention to whether arithmetic operations are admissible. The near absence of latent attribute measurement and the distinction between manifest and latent ratio measures indicates little appreciation that different classes of outcomes require different measurement frameworks.

The consequence is that the curriculum prepares students and researchers to evaluate evidence without first equipping them to evaluate the measurement status of that evidence. Evidence synthesis becomes detached from measurement. Studies reporting utility scores, preference-based measures, composite indices, quality-of-life scores, QALYs, and cost-effectiveness ratios can be systematically reviewed and synthesized without first establishing whether the quantities involved qualify as measures. The question shifts from “Is this a measure?” to “What do the studies report?”

This is the defining feature of curriculum inversion within the CRD environment. Students are trained to synthesize evidence before they are trained to determine whether the quantities being synthesized satisfy the requirements of quantitative science. The curriculum therefore reproduces the same inversion identified elsewhere in HTA. Measurement no longer governs arithmetic. Instead, arithmetic and numerical outputs are accepted as evidence while measurement is assumed.

The implications extend well beyond curriculum content. CRD occupies a central position within the UK HTA evidence infrastructure. If the concepts required to evaluate measurement are absent from the educational and research environment, then the same omission is likely to influence the reviews and assessments produced by that environment. Under these circumstances, systematic review risks becoming the aggregation of numerical constructions rather than the synthesis of valid measures.

The significance of the interrogation therefore lies not merely in the concepts omitted from the curriculum, but in the relationship between those omissions and the wider role of CRD within UK HTA. Evidence synthesis is a powerful scientific tool when applied to valid measures. It cannot compensate for their absence. A systematic review of quantities that fail the requirements of measurement does not remedy the defect. It merely reproduces it on a larger scale.

The CRD profile therefore provides a distinctive example of curriculum inversion. Whereas economic evaluation centers such as CHE reproduce the assumptions of the reference case through modelling and cost-effectiveness analysis, CRD reproduces the same assumptions through evidence synthesis. Different institutional missions lead to the same result. The concepts necessary to evaluate the scientific legitimacy of quantitative claims remain largely absent. The curriculum prepares students to review evidence, but not to determine whether the quantities presented as evidence possess the measurement properties required to support the claims being made.

The conclusion is unavoidable. CRD's curriculum environment demonstrates strong recognition of evidence synthesis and policy evaluation but weak recognition of measurement science. The result is a form of curriculum inversion in which evidence is synthesized before measurement is established. In this respect, CRD reproduces the same foundational weakness identified throughout contemporary HTA: the assumption that quantitative claims can be evaluated without first demonstrating that the quantities involved are measures.

REVIEW: OXFORD HTA AND HEALTH ECONOMICS

The Oxford HTA and Health Economics program occupies a prominent position within UK health technology assessment. Through its work in economic evaluation, outcomes research, trial-based health economics, health policy, epidemiology, and healthcare decision-making, Oxford has contributed substantially to the evidence base supporting reimbursement and resource-allocation decisions. Its influence extends beyond the United Kingdom through international collaborations, graduate training programs, methodological research, and policy engagement.

The interrogation results reveal a curriculum environment that places considerable emphasis on outcomes, evidence, evaluation, and empirical assessment. High endorsement probabilities are observed for outcome identification, target specification, manifest attributes, and falsifiable claims. Students are likely to encounter extensive training in study design, clinical outcomes, economic evaluation, comparative effectiveness research, and healthcare policy. The curriculum therefore appears strongly aligned with the practical requirements of contemporary HTA.

However, as with other leading HTA centers, a striking discontinuity appears when attention shifts from outcomes and evidence to measurement itself.

The proposition that different scales of measurement possess different analytical properties receives a probability of only 0.25. The proposition that measurement status must be established before quantitative claims can be advanced receives only 0.15. Most revealingly, representational measurement receives a probability of only 0.05. These findings suggest that measurement theory occupies only a marginal position within the intellectual framework supporting Oxford HTA.

This distinction is important because Oxford's strength lies in empirical evaluation. Yet empirical evaluation and measurement are not identical. An outcome can be observed without being measured. A variable can be analyzed without satisfying the requirements of measurement. The central question is not whether an outcome exists but whether the numerical representation assigned to that outcome possesses the properties necessary to support arithmetic operations and quantitative claims.

The low endorsement of representational measurement indicates that this question receives little explicit attention. Instead, the curriculum appears to assume that once an outcome has been identified it can be incorporated into economic analyses, policy evaluations, comparative effectiveness studies, and decision frameworks. The issue is not the quality of the subsequent analysis but the absence of a prior measurement step.

The treatment of manifest attributes illustrates this pattern. Oxford demonstrates stronger recognition of observable outcomes than many HTA centers. This is unsurprising given its extensive involvement in clinical and population-health research. Survival, hospitalization, resource utilization, adverse events, and treatment persistence are all outcomes capable of direct observation. Yet even here the curriculum provides little indication that these outcomes are recognized as linear ratio measures. The practical use of ratio measures appears to occur without explicit recognition of ratio measurement itself.

The problem becomes substantially more serious when latent attributes are considered. Outcomes such as quality of life, functioning, fatigue, pain, patient satisfaction, and psychological well-being are central to contemporary HTA. These are not directly observable attributes. They require a measurement model capable of estimating possession of the latent trait.

Yet the proposition that latent attributes require a measurement model receives a probability of only 0.15. The distinction between manifest and latent forms of ratio measurement receives only 0.05. These findings suggest that one of the most important distinctions in measurement science remains largely absent from the Oxford HTA curriculum.

The implications are considerable. Once latent attributes are not recognized as a unique measurement problem, there is little reason to introduce the framework required to solve that problem. This helps explain the near absence of Rasch measurement. Rasch theory provides the only established framework for constructing invariant measures of latent attribute possession. It is the mechanism through which ordinal observations can be transformed into Rasch logit ratio measures. Yet the interrogation suggests little recognition of its relevance within the Oxford HTA environment.

This omission has direct implications for the treatment of utility instruments, quality-of-life measures, and patient-reported outcomes. Students may encounter EQ-5D, HUI, SF-6D, and related instruments as standard components of HTA without first being introduced to the measurement questions these instruments raise. The result is that utility generation, QALY construction, and economic evaluation become accepted analytical procedures rather than subjects of measurement scrutiny.

The consequence is curriculum inversion. Students are trained to apply analytical frameworks before they are introduced to the measurement principles necessary to evaluate those frameworks. Economic evaluation precedes measurement. Cost-effectiveness analysis precedes measurement. Utility construction precedes measurement. Arithmetic acquires authority simply because it is numerical.

The significance of this finding extends beyond Oxford itself. As one of the United Kingdom's most influential research environments, Oxford contributes to the training of future HTA researchers, analysts, policy advisers, and academic leaders. The concepts emphasized within its curriculum are therefore likely to influence future HTA practice. Conversely, concepts omitted from the curriculum are unlikely to appear in the research and policy environments that graduates subsequently enter.

The Oxford interrogation therefore provides further evidence that curriculum inversion is not confined to a single institution or methodological tradition. Whether the focus is economic evaluation at CHE, evidence synthesis at CRD, or outcomes and policy research at Oxford, the same pattern emerges. Students are taught how to evaluate therapies but not how to determine whether the quantities entering those evaluations satisfy the requirements of measurement science.

The conclusion is therefore clear. Oxford HTA and Health Economics demonstrates a strong commitment to evidence, outcomes research, empirical evaluation, and healthcare decision-making. What remains largely absent is a comparable commitment to the scientific foundations of measurement itself. The result is a curriculum that prepares students to participate effectively within the reference-case paradigm while providing limited opportunity to evaluate its measurement foundations. In this respect, Oxford reproduces the same curriculum inversion that characterizes contemporary HTA more generally: measurement becomes assumed, while analysis becomes the primary object of instruction.

REVIEW: SCHOOL OF HEALTH AND RELATED RESEARCH (SCHARR)

The School of Health and Related Research (SchARR) at the University of Sheffield occupies a central position within the United Kingdom HTA environment. Through its extensive involvement in NICE technology appraisals, evidence review programs, economic evaluation, decision modelling, evidence synthesis, and healthcare policy research, SchARR has become one of the principal institutions responsible for training HTA practitioners and supporting reimbursement decision-making. Unlike centers whose primary focus is methodological research, SchARR is closely associated with the practical application of HTA methods within real-world policy environments. This makes the interrogation particularly important because it provides insight into the educational foundations of the operational HTA paradigm.

The results reveal a curriculum strongly oriented toward assessment, evaluation, and decision support. High endorsement probabilities are observed for outcome identification and target specification, while moderate probabilities are observed for manifest attributes and falsifiable claims. Students are therefore likely to receive extensive training in the practical processes of technology assessment: defining outcomes, evaluating evidence, comparing interventions, constructing economic models, and supporting reimbursement recommendations.

However, the interrogation reveals a dramatic decline in endorsement when attention shifts from evaluation to measurement. Scale theory receives a probability of only 0.20. The proposition that measurement status must be established before quantitative claims can be advanced receives only 0.10. Representational measurement receives the lowest probability, 0.05. Similar profiles are observed for unidimensionality, latent attribute measurement, and the distinction between manifest and latent forms of ratio measurement.

These findings indicate that the curriculum gives little attention to the scientific foundations that determine whether quantitative claims are admissible. The practical question of how to undertake an HTA receives considerable attention. The prior scientific question—whether the quantities entering the assessment qualify as measures—receives very little.

This distinction is particularly important because ScHARR operates at the point where HTA methods are translated into policy advice. The curriculum therefore appears designed to prepare students to function effectively within the reference-case framework. Economic evaluation, cost-effectiveness analysis, simulation modelling, evidence synthesis, and reimbursement assessment occupy a central position. Yet there is little indication that students are systematically exposed to representational measurement, dimensional homogeneity, ratio measurement, or the admissibility of arithmetic operations.

The consequence is that measurement becomes assumed rather than demonstrated. Utility values enter economic evaluations without first establishing whether they possess ratio properties. QALYs enter cost-effectiveness calculations without establishing whether multiplication is admissible. Simulation models extend these assumptions across hypothetical populations and time horizons. The educational framework therefore mirrors the analytical framework. Arithmetic precedes measurement because arithmetic precedes measurement in education.

The low endorsement of unidimensionality is particularly revealing. Measurement requires that an attribute represent a single dimension. Without unidimensionality there may be aggregation, scoring, or numerical assignment, but there is no measurement. Yet the interrogation suggests little recognition of this requirement. Students are therefore likely to encounter composite indices, utility systems, and multidimensional outcome measures without being equipped to determine whether the resulting quantities satisfy the conditions required for measurement.

The treatment of latent attributes is equally important. Contemporary HTA routinely evaluates outcomes such as pain, functioning, fatigue, quality of life, and patient experience. These are latent attributes. They are not directly observable and require a measurement model to estimate possession of the attribute. Yet the proposition that latent attributes require such a framework receives a probability of only 0.10. The proposition that manifest and latent attributes require different forms of ratio measurement receives only 0.05.

This finding has direct implications for Rasch measurement. Once latent attributes are not recognized as a distinct measurement challenge, there is little reason to introduce the framework required to address that challenge. The result is that Rasch measurement occupies little or no visible role within the curriculum despite providing the only established route to invariant latent-attribute measurement. Students therefore encounter utility instruments and quality-of-life measures without first understanding the measurement requirements that such instruments should satisfy.

The significance of the ScHARR interrogation extends beyond a single institution. ScHARR represents one of the principal training environments for future HTA analysts, reviewers, and policy advisers. The curriculum therefore helps shape the intellectual assumptions carried into NICE appraisals, reimbursement reviews, and technology assessments. The concepts emphasized in training become the concepts emphasized in practice. Conversely, concepts absent from training are unlikely to appear in the policy environment.

The result is curriculum inversion in its most developed form. Students are taught how to operate the machinery of HTA without first being taught how to evaluate the measurement foundations

upon which that machinery depends. Cost-effectiveness analysis, utility assessment, evidence synthesis, and decision modelling become normal professional activities. The scientific question of whether the quantities involved are measures largely disappears from view.

The importance of the interrogation lies precisely in this observation. ScHARR demonstrates strong endorsement of the activities required to implement the reference case while displaying minimal recognition of the concepts necessary to evaluate its scientific legitimacy. The curriculum therefore does not merely omit measurement science; it normalizes a framework in which measurement is assumed. This is the defining characteristic of curriculum inversion.

The conclusion is unavoidable. ScHARR provides further evidence that curriculum inversion is embedded within the UK HTA environment. Students are prepared to apply the reference-case paradigm, but they are given little opportunity to examine the measurement assumptions upon which that paradigm rests. As a result, the educational environment reproduces the same measurement inversion observed throughout contemporary HTA practice. The framework survives not because its measurement foundations have been demonstrated, but because the concepts required to challenge those foundations remain largely absent from the curriculum itself.

REVIEW: NEWCASTLE HTA GROUP

The Newcastle HTA Group occupies an established position within the United Kingdom health technology assessment environment through its contributions to evidence reviews, economic evaluation, health services research, guideline development, and policy support. Like other major UK HTA centers, Newcastle participates in the wider intellectual framework that underpins technology assessment, reimbursement evaluation, and healthcare decision-making. The interrogation therefore seeks to determine whether the concepts necessary to support valid quantitative claims occupy a meaningful place within the educational and research environment associated with Newcastle HTA.

The results reveal a familiar pattern. Strong endorsement is attached to outcome identification, target specification, manifest outcomes, and falsifiable claims. Students and researchers are likely to encounter extensive discussion of clinical outcomes, comparative effectiveness, healthcare interventions, policy evaluation, evidence assessment, and decision-making frameworks. In this respect the curriculum is strongly aligned with the operational requirements of contemporary HTA. However, the profile changes dramatically when attention shifts from outcomes to measurement.

Recognition of scale theory receives a probability of only 0.20. The proposition that the measurement status of an attribute must be established before quantitative claims can be advanced receives only 0.10. Representational measurement receives the lowest probability, 0.05. Equally low endorsement is observed for unidimensionality, latent attribute measurement, and the distinction between manifest and latent forms of ratio measurement.

These findings indicate that the Newcastle curriculum, like those of the other leading UK HTA centers, gives little attention to the scientific foundations upon which quantitative claims ultimately depend. Students are taught how to identify outcomes, assess evidence, and undertake evaluations.

They are not systematically taught how to determine whether the quantities entering those evaluations satisfy the requirements of measurement science.

This distinction is critical. Identifying an outcome does not establish that the outcome has been measured. Observing a phenomenon does not demonstrate that the numerical representation assigned to that phenomenon possesses ratio properties. Yet the interrogation suggests that measurement is largely assumed rather than demonstrated. The curriculum appears to move directly from outcome identification to analysis, bypassing the prior question of whether the quantities involved qualify as measures.

The low endorsement of representational measurement is particularly revealing. Representational measurement provides the framework that determines whether a numerical assignment corresponds to an empirical attribute in a manner that supports lawful arithmetic. Without this framework there is no basis for deciding whether averaging, multiplication, division, or ratio construction are admissible. Yet the interrogation suggests that students can progress through HTA training with little exposure to these concepts.

The same problem is evident in the treatment of unidimensionality. Measurement requires that an attribute represent a single dimension. Without unidimensionality, aggregation may occur but measurement does not. Nevertheless, the curriculum appears to provide little recognition of this requirement. Composite indices, utility systems, patient-reported outcomes, and multidimensional instruments may therefore be encountered without explicit discussion of whether the resulting numerical outputs satisfy the conditions necessary for measurement.

The treatment of latent attributes is equally significant. Contemporary HTA increasingly relies upon outcomes such as pain, functioning, quality of life, fatigue, mental health, and patient satisfaction. These are latent attributes. They cannot be directly observed and require a measurement model capable of estimating possession of the attribute. Yet the probability attached to recognition of latent attributes is only 0.10, while the probability attached to the proposition that manifest and latent attributes require different forms of ratio measurement is only 0.05.

The consequence is that one of the most important distinctions in measurement science effectively disappears from the curriculum. Students may encounter both survival and quality of life, hospital admissions and patient satisfaction, resource utilization and functional status, without recognizing that these outcomes belong to fundamentally different measurement domains. The distinction between observable and latent phenomena becomes blurred, and with it the need for different measurement frameworks.

This finding has direct implications for Rasch measurement. Once latent attributes are not recognized as requiring a unique measurement framework, there is little reason to introduce the Rasch model. Consequently, the curriculum reproduces the broader HTA pattern in which utility instruments, quality-of-life measures, and patient-reported outcomes occupy a central place while Rasch measurement remains largely absent. Students therefore encounter numerical summaries of latent attributes without first being introduced to the framework required to construct invariant measures of latent attribute possession.

The significance of the Newcastle interrogation lies not in any unique deficiency but in its similarity to the other UK centers. CHE, CRD, Oxford, ScHARR, and Newcastle all display remarkably similar profiles. The same concepts receive strong endorsement. The same concepts receive weak endorsement. The same omissions recur regardless of institutional mission. Economic evaluation centers evidence-synthesis centers, policy-oriented centers, and health-services research centers all exhibit the same curriculum structure.

This consistency is perhaps the most important finding. Had the deficiencies been confined to a single institution, they might reasonably be interpreted as local curriculum choices. Instead, the same pattern appears repeatedly across the UK HTA landscape. The findings therefore point not to institutional variation but to a shared intellectual framework.

The result is curriculum inversion. Students are introduced to the methods, tools, and outputs of HTA before they are introduced to the principles necessary to determine whether those outputs possess scientific legitimacy. Economic evaluations, utility measures, QALYs, evidence syntheses, and policy models become accepted analytical constructs while the concepts required to evaluate their measurement foundations remain largely absent.

The conclusion is clear. Newcastle HTA Group demonstrates strong recognition of outcomes, evaluation, and evidence assessment but weak recognition of measurement science. The curriculum prepares students to operate within the contemporary HTA paradigm while providing little opportunity to evaluate its measurement assumptions. In doing so it reproduces the same curriculum inversion identified across the wider UK HTA environment. Measurement is assumed, arithmetic is accepted, and the scientific foundations necessary to support quantitative claims remain largely invisible.

OVERVIEW: CURRICULUM INVERSION

The curriculum inversion interrogation of five leading United Kingdom HTA research centers reveals a remarkably consistent pattern. Despite differences in institutional mission, research focus, and organizational history, the endorsement profiles are almost identical.

This consistency is perhaps the most important finding. The centers represent different components of the UK HTA infrastructure. CHE is closely associated with economic evaluation and the intellectual foundations of the NICE reference case. CRD is internationally recognized for evidence synthesis and systematic reviews. HERC has a strong focus on health economics and outcomes research. ScHARR is deeply involved in technology appraisal and reimbursement assessments. Newcastle's Institute of Health and Society contributes to HTA reviews, health services research, and policy evaluation. Yet despite these differences, all five centers demonstrate almost identical endorsement profiles.

The strongest probabilities are attached to outcome identification and target specification. Across all centers there is clear recognition that therapy assessment requires the identification of an outcome of interest and specification of a target attribute. Probabilities range from 0.75 to 0.85 for these propositions. Similarly, there is moderate recognition of manifest attributes and relatively strong recognition that therapy impact claims should be falsifiable. These findings indicate that

the UK HTA curriculum places considerable emphasis on defining outcomes, evaluating interventions, and assessing evidence.

The difficulty emerges when attention shifts from outcomes to measurement. Recognition that the principal scales of measurement possess different analytical properties receives probabilities of only 0.20 to 0.25. The proposition that measurement status must be established before quantitative claims can be advanced receives probabilities of only 0.10 to 0.15. More striking still, the proposition that the axioms of representational measurement underpin quantitative claims receives a probability of only 0.05 across every center. This complete uniformity is highly significant. Regardless of institutional setting, representational measurement occupies virtually no visible position within the curriculum environment.

The same pattern is evident for unidimensionality, latent attribute measurement, and the distinction between manifest and latent forms of ratio measurement. Unidimensionality receives a probability of only 0.10 across all five centers. Recognition that latent attributes require a measurement model receives probabilities of only 0.10 to 0.15. The proposition that manifest and latent attributes require different forms of ratio measurement receives a probability of only 0.05 across every institution. These findings point to the near absence of the concepts that underpin modern measurement theory.

The implications are profound. A scientific curriculum should equip students with the tools required to evaluate the validity of quantitative claims. In HTA this would require explicit instruction in representational measurement, scale theory, admissible arithmetic, unidimensionality, ratio measurement, dimensional homogeneity, and the distinction between manifest and latent attributes. Yet these concepts receive consistently low endorsement across all five UK centers. Students are therefore likely to encounter economic evaluation, cost-effectiveness analysis, utility assessment, QALY construction, evidence synthesis, and simulation modelling without first being exposed to the measurement principles necessary to evaluate those activities.

This pattern is the defining characteristic of curriculum inversion. The curriculum begins with analytical techniques and assumes that measurement has already been established. Students learn how to generate and interpret quantitative outputs before they learn how to determine whether the quantities entering those analyses satisfy the requirements of measurement. Arithmetic precedes measurement in education just as it precedes measurement in the reference-case framework itself.

The findings are particularly important because these centers collectively represent the intellectual foundations of UK HTA. They train researchers, analysts, policy advisers, reviewers, and academics who subsequently populate NICE appraisal groups, research organizations, universities, and consulting firms. The concepts emphasized within the curriculum therefore shape the future development of HTA. Conversely, concepts absent from the curriculum are unlikely to appear in research practice or policy analysis. Curriculum inversion thus becomes the mechanism through which measurement inversion is reproduced across successive generations of HTA practitioners.

The UK results also mirror findings reported previously in Australia, Canada, and New Zealand. In each country the same concepts receive strong endorsement and the same concepts receive weak endorsement. Outcome identification, evaluation, and evidence assessment are emphasized.

Representational measurement, ratio measurement, unidimensionality, latent attribute measurement, and Rasch measurement remain largely absent. The consistency of these findings across countries suggests that curriculum inversion is not a local institutional problem but a defining characteristic of the contemporary international HTA paradigm.

The conclusion is therefore unavoidable. The leading UK HTA research centers do not simply exhibit isolated gaps in curriculum content. Rather, they share a common educational framework in which the concepts necessary to evaluate the scientific legitimacy of quantitative claims occupy a marginal position. The result is a self-reinforcing intellectual environment in which students are trained to operate within the reference-case paradigm while lacking the conceptual tools necessary to evaluate its measurement foundations. Curriculum inversion therefore provides a compelling explanation for the persistence of measurement inversion throughout UK HTA and points directly to the need for a fundamental reconstruction of both HTA methodology and HTA education.

After more than forty years of dominance by the reference-case paradigm, the critical question is whether these research centers will respond to the evidence. Will they acknowledge the implications of measurement inversion and curriculum inversion and undertake the substantial task of reconstructing their educational and research programs? Or will they attempt to preserve the *status quo* through increasingly elaborate defenses of utilities, QALYs, and simulation models despite their failure to satisfy the standards of measurement science? The choice is unavoidable. Once the absence of representational measurement, ratio measurement, unidimensionality, and Rasch measurement has been demonstrated, there can be no return to business as usual. The false intellectual foundations of the current paradigm have been exposed.

HTA RECONSTRUCTION PROGRAM

The reference-case paradigm has reached the end of its scientific life. For more than four decades, health technology assessment (HTA) has relied upon utilities, quality-adjusted life years (QALYs), cost-effectiveness analysis, and reference-case simulation models to support claims regarding therapy impact. Recent assessments of HTA knowledge bases across government agencies, research centers, professional organizations, journals, and university programs demonstrate that this paradigm was established without satisfying the requirements of representational measurement. Arithmetic has consistently been placed before measurement. The result is measurement inversion, reinforced by curriculum inversion, in which generations of researchers and practitioners have been trained in analytical techniques without first acquiring the scientific principles required to determine whether those techniques can support lawful quantitative claims.

The HTA Reconstruction Program has been developed by Maimon Research LLC in response to this paradigm failure⁷. Its objective is not to reform the existing framework but to replace it with one grounded in representational measurement and the standards of normal science. Beginning with the principle that measurement must precede arithmetic, the program develops the scientific foundations required for credible therapy assessment: attributes, the principal scales of measurement, lawful ratio measurement, manifest and latent attributes, Rasch logit ratio measurement, and protocol-driven claims that are evaluable, replicable, and capable of falsification.

Designed for faculty, graduate students, researchers, manufacturers, formulary committees, and HTA agencies, the program provides a structured pathway from assumption-driven modelling to scientific HTA. It is intended to establish the competencies required for a new generation of health technology assessment in which evidence is grounded in lawful measurement rather than numerical construction, and where therapy impact claims meet the same scientific standards expected throughout the quantitative sciences

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 United Kingdom. Measurement inversion and paradigm failure in health technology assessment. Logit Working Paper No 586 <https://maimonresearch.com/logit-working-paper-no-586-june-2026/>

² Maimon Research LLC. Interrogations in HTA. www.maimonresearch.com

³ Stevens S. On the Theory of Scales of Measurement. *Science*. 1946;103(2684):677-80

⁴ Krantz D, Luce R, Suppes P, Tversky A. Foundations of Measurement Vol 1: Additive and Polynomial Representations. New York: Academic Press, 1971

⁵ Bond T, Zi Yan, Heene M. Applying the Rasch Model: Fundamental Measurement in the Human Sciences (4th Ed). New York: Routledge, 2021

⁶ The Theory of Allowable Transformations and the QALY. Logit Working Paper No 887. <https://maimonresearch.com/logit-working-paper-no-887-june-2026/>

⁷ Maimon Research Transition Program <https://maimonresearch.com/hta-reconstruction-program-and-fees/>