MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

NETHERLANDS: THE ENDORSEMENT OF CURRICULUM INVERSION

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

LOGIT WORKING PAPER No 1511 JULY 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

Previous large language model (LLM) interrogations of the Dutch health technology assessment (HTA) knowledge base demonstrated that contemporary Dutch HTA is characterized by measurement inversion. The accepted principles of representational measurement receive little recognition, while utilities, QALYs, cost-effectiveness analysis, and reference-case simulation modelling are accepted as though they provide scientifically valid quantitative measures of therapy impact. The present study addresses the complementary question of how this framework has become established and sustained. It argues that the principal mechanism is curriculum inversion.

The Dutch HTA curriculum knowledge base was interrogated using ten canonical statements representing the scientific foundations of representational measurement. These encompass specification of the target attribute, the principal scales of measurement, the axioms of representational measurement, unidimensionality, the distinction between manifest and latent attributes, the corresponding requirements for linear ratio and Rasch logit ratio measurement, and the requirement that therapy-impact claims be prospectively evaluable, independently replicable, and capable of empirical falsification. Together these statements define the minimum educational content required for a scientifically defensible approach to HTA.

The interrogation demonstrates a consistent pattern of curriculum inversion. Concepts fundamental to quantitative science receive uniformly weak endorsement, while educational emphasis is directed toward comparative effectiveness assessment, economic evaluation, utilities, QALYs, decision modelling, and reference-case methodology. Students therefore learn how to construct and apply the reference-case paradigm before acquiring the scientific principles necessary to determine whether its constituent quantities satisfy the requirements for lawful measurement. Particularly notable is the absence of any explicit distinction between manifest and latent attributes and the corresponding absence of Rasch measurement as the only established framework capable of constructing quantitative measures of latent attribute possession.

These findings carry particular significance because the Netherlands has been one of the principal architects of the European reference-case tradition. Curriculum inversion within the Dutch educational framework therefore has implications extending well beyond national HTA practice, providing one mechanism through which measurement inversion has been disseminated throughout Europe. The paper concludes that the reference-case paradigm cannot be rescued through incremental methodological refinement because its central constructs fail the accepted axioms of representational measurement. Reconstruction must instead begin with education. HTA curricula should be reorganized around representational measurement, explicit specification of target attributes, the distinction between manifest and latent attributes, lawful linear ratio and Rasch logit ratio measurement, and prospectively specified, evaluable, replicable, and falsifiable claims regarding therapy impact. Only such a reconstruction can provide a scientifically defensible foundation for the future development of HTA.

INTRODUCTION

A series of large language model (LLM) interrogations of the Dutch health technology assessment (HTA) knowledge base has demonstrated strong endorsement of measurement inversion [1] [2] [3], Across national methodological guidance, university teaching, academic research, professional training, and policy practice, the same pattern emerges: the concepts required for lawful quantitative measurement are either weakly represented or absent, while utilities, QALYs, cost-effectiveness analysis, and reference-case simulation modelling are accepted as though they provide scientifically valid measures of therapy impact. These findings are especially significant because the Netherlands has not merely adopted the contemporary European reference-case tradition; it has been one of its principal architects.

The present study examines whether this Dutch endorsement of measurement inversion is accompanied by curriculum inversion. This is a critical question because evidence from Australia, Canada, France, Germany, Italy, New Zealand, Spain, the United Kingdom, and the United States indicates that measurement inversion is sustained through the educational framework that prepares successive generations of HTA practitioners. The Dutch case is particularly important because the Netherlands has played a leading role in promoting health economic evaluation, cost-utility analysis, QALYs, modelling, and standardized decision methods across Europe. If curriculum inversion is present in the Dutch HTA curriculum knowledge base, then one of the principal sources of European HTA methodology is itself reproducing the educational conditions that sustain measurement inversion and the measurem3nt failures of the reference case framework [4].

The interrogation reveals a consistent pattern of curriculum inversion within the Dutch curriculum knowledge base. Across the ten canonical statements, the scientific concepts that should form the foundation of quantitative HTA receive uniformly weak endorsement. Fundamental principles including specification of the target attribute, scales of measurement, representational measurement, measurement before arithmetic, unidimensionality, manifest and latent attributes, Rasch measurement for latent constructs, and falsifiable therapy-impact claims are either weakly represented or effectively absent. At the same time, the curriculum knowledge base emphasizes comparative effectiveness assessment, utilities, QALYs, cost-effectiveness analysis, economic evaluation, decision modelling, thresholds, and reimbursement decision support. The result is an educational framework that teaches practitioners how to apply the reference-case paradigm before equipping them with the scientific principles required to determine whether that paradigm can support lawful quantitative claims. Curriculum inversion thus emerges not merely as a Dutch educational omission, but as one mechanism through which the European reference-case tradition has been transmitted and sustained.

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 for manifest and the application of Rasch models for latent attributes [5] [6]

[7]. 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 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 and decision makers. Curriculum inversion therefore serves as the educational mechanism through which measurement inversion becomes institutionalized within a discipline. In HTA this serves to support administrative decisions for therapy pricing and access.

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 Spanish 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 scales of measurement, the axioms of representational measurement, unidimensionality, latent attribute measurement or ratio 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 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.

THE DUTCH HTA CURRICULUM KNOWLEDGE BASE

The Dutch HTA curriculum knowledge base comprises the educational, methodological, and professional resources through which the principles and practice of health technology assessment

(HTA) are taught, interpreted, and disseminated throughout the Netherlands. It extends well beyond formal university curricula to encompass postgraduate education, professional training programs, methodological guidance, publications from national organizations, conference proceedings, continuing professional development, and the extensive academic literature on health economics and HTA. Collectively, these resources establish the conceptual framework through which successive generations of clinicians, pharmacists, economists, policy analysts, consultants, manufacturers, and researchers acquire their understanding of HTA and its role in healthcare decision making.

The Netherlands occupies a distinctive position within the international development of HTA. For more than three decades, Dutch researchers, universities, and policy organizations have played a leading role in advancing health economic evaluation, cost-utility analysis, decision modelling, and reference-case methodology. Dutch contributions have influenced not only national reimbursement policy but also the development of methodological guidance adopted throughout Europe. Consequently, the Dutch curriculum knowledge base is important not simply because it shapes education within the Netherlands but because it has contributed to defining the analytical framework that has become characteristic of European HTA more generally.

The purpose of identifying the Dutch curriculum knowledge base is to distinguish between the educational framework through which HTA is taught and the broader HTA knowledge base that encompasses assessment reports, reimbursement procedures, methodological guidance, and policy implementation. Nested within the wider Dutch HTA knowledge base is an educational framework whose primary function is to transmit accepted analytical methods to future practitioners. It determines the concepts regarded as fundamental, the analytical sequence in which they are presented, and the methodological assumptions that define professional competence. The present interrogation is directed toward this educational framework rather than toward the outcomes of individual reimbursement decisions or specific HTA reports.

The interrogation focuses on a fundamental scientific question. Does the Dutch curriculum knowledge base introduce students and practitioners first to the accepted principles of representational measurement before presenting the established methods of contemporary HTA? These principles include specification of the target attribute, the principal scales of measurement, the axioms of representational measurement, admissible arithmetic, dimensional homogeneity, unidimensionality, the distinction between manifest and latent attributes, the corresponding requirements for linear ratio and Rasch logit ratio measurement, and the requirement that therapy-impact claims be prospectively evaluable, independently replicable, and capable of empirical falsification. These principles represent the scientific foundations upon which all lawful quantitative claims ultimately depend.

The importance of this interrogation is heightened by the historical influence of the Netherlands on the development of the European reference-case tradition. If these scientific foundations are absent from one of the principal educational systems responsible for developing and disseminating contemporary HTA methodology, the implications extend well beyond the Netherlands itself. Curriculum inversion would then represent not merely a national educational deficiency but an important source from which measurement inversion has been propagated throughout Europe. The Dutch curriculum knowledge base therefore provides a particularly important case study because it occupies a pivotal position in the intellectual development of modern HTA. The interrogation seeks to determine whether the educational framework that has helped shape European HTA is itself grounded in the accepted principles of representational measurement or whether it instead reproduces methodological convention without first establishing the scientific conditions required for lawful quantitative measurement.

INTERROGATING THE DUTCH CURRICULUM KNOWLEDGE BASE

The objective of large language model (LLM) 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 Canada 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.

THE NETHERLANDS AND CURRICULUM INVERSION

The Netherlands has been one of the principal architects of modern European HTA. Dutch health economists and policy researchers have had a major influence on the development of cost-utility analysis, reference-case methodology, economic evaluation guidelines, and the widespread use of utilities and QALYs. Through Dutch universities, advisory groups, and collaborations with the UK and Scandinavian countries, these ideas became embedded not only within the Netherlands but throughout Europe. Dutch methodological guidance has consistently emphasized transparency, consistency, and standardized decision making. The interrogation suggests, however, that these achievements occurred within an administrative rather than a measurement framework. The issue is therefore not that Dutch HTA lacks methodological sophistication. It is that the sophistication has been directed toward improving the operation of a framework whose measurement foundations remain largely unexamined.

TABLE 1: CURRICULUM CONTENT ENDORSEMENT: NETHERLANDS KNOWLEDGE BASE

CANONICAL STATEMENT	CATEGORICAL PROBABILITY	NORMALIZED LOGIT
An attribute is the specific outcome of interest in a therapy assessment	0.20	-1.50
Every therapy assessment begins with specification of the target attribute	0.10	-2.00
The principal scales of measurement (nominal, ordinal, interval and ratio) have different properties and support different forms of analysis	0.15	-1.75
The measurement status of a target attribute must be established before quantitative claims can be advanced	0.10	-2.00
The axioms of representational measurement underpin quantitative claims	0.05	-2.50
Attributes must be demonstrated to be unidimensional before measurement is possible	0.10	-2.00
A manifest attribute is directly observable and capable of supporting empirical observation	0.25	-1.25
A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute	0.10	-2.00
Manifest and latent attributes require different forms of ratio measurement	0.05	-2.50
Therapy impact claims must be falsifiable	0.15	-1.75

This gives the Dutch case particular significance. If countries with relatively limited HTA traditions exhibit measurement inversion, this might reasonably be attributed to adoption of international practice. The Netherlands cannot be interpreted in this way. It has itself been a leading exporter of that practice. Consequently, the interrogation indicates that curriculum inversion has become institutionalized within one of the countries that helped define the contemporary European reference model. Rather than transmitting representational measurement as the foundation of quantitative science, the curriculum transmits economic evaluation, utilities, QALYs, and decision modelling as established analytical methods. The educational consequence is that successive generations of HTA practitioners acquire increasing technical expertise in applying the reference-case framework without first examining whether its constituent quantities satisfy the conditions required for lawful measurement.

The broader implication extends well beyond the Netherlands. Because Dutch methodological guidance has influenced HTA development across Europe, the Netherlands occupies a pivotal position in the genealogy of the European reference-case paradigm. If curriculum inversion is present within the Dutch knowledge base, then one of the principal sources from which other jurisdictions have drawn their methodological assumptions is itself characterized by the same inversion. This strengthens the conclusion that measurement inversion is not the product of isolated national practices but a structural feature of the European HTA tradition.

Statement-by-Statement Review

1. An attribute is the specific outcome of interest in a therapy assessment — Probability 0.20; Logit -1.50

The Netherlands curriculum knowledge base gives only weak endorsement to the idea that an attribute is the specific outcome of interest in therapy assessment. Dutch HTA education and guidance give substantial attention to economic evaluation, cost-effectiveness, QALYs, modelling, disease burden, and reimbursement decision making. Zorginstituut Nederland describes its economic evaluation guideline as the starting point for conducting economic evaluations in healthcare, while Dutch HTA training resources emphasize cost-effectiveness methods, modelling, QALYs, and thresholds. The attribute itself is not typically presented as the primary object of measurement.

2. Every therapy assessment begins with specification of the target attribute — Probability 0.10; Logit -2.00

This proposition receives very weak endorsement. The educational sequence appears to begin with evidence appraisal, economic evaluation, health outcomes, and modelling rather than with specification of the target attribute. Students and practitioners are taught how to assess interventions using accepted HTA methods before being required to identify precisely what is being measured. The result is curriculum inversion: analytical technique precedes attribute specification.

3. The principal scales of measurement have different properties and support different forms of analysis — Probability 0.15; Logit -1.75

The Netherlands curriculum knowledge base shows limited recognition of the principal scales of measurement as constraints on analysis. Dutch health economics education clearly teaches quantitative methods, applied econometrics, cost-effectiveness, and HTA, but there is little indication that nominal, ordinal, interval, and ratio scales are positioned as foundational requirements before arithmetic proceeds. Numerical analysis is therefore taught more strongly than the measurement theory governing which numerical operations are admissible.

4. The measurement status of a target attribute must be established before quantitative claims can be advanced — Probability 0.10; Logit -2.00

This receives very weak endorsement. The Dutch knowledge base emphasizes the conduct of economic evaluations in an "unambiguous manner" so that interventions can be compared, but comparison is not the same as measurement. The interrogation indicates that the measurement status of the target attribute is generally assumed rather than demonstrated. Claims proceed through economic evaluation rather than first through a measurement justification.

5. The axioms of representational measurement underpin quantitative claims — Probability 0.05; Logit -2.50

This receives the lowest endorsement. There is little evidence that the axioms of representational measurement are presented as the scientific foundation for Dutch HTA education. Concepts such as admissible transformations, dimensional homogeneity, cancellation, unidimensionality, and measurement before arithmetic are not prominent in the curriculum knowledge base. This is the central indication of curriculum inversion.

6. Attributes must be demonstrated to be unidimensional before measurement is possible — Probability 0.10; Logit -2.00

The requirement for unidimensionality receives very weak endorsement. Dutch HTA practice, like other European HTA systems, relies heavily on quality-of-life measures, utility instruments, preference scores, and QALYs. Yet the curriculum knowledge base does not appear to require demonstration that these constructs represent a single measurable attribute before they are used in quantitative claims. Multidimensional scoring is therefore treated as though it can support measurement without first satisfying the conditions required for measurement.

7. A manifest attribute is directly observable and capable of supporting empirical observation — Probability 0.25; Logit -1.25

This statement receives the strongest endorsement, but still only modestly. Dutch HTA education recognizes observable clinical outcomes, such as survival, events, hospitalizations, adverse events, and resource use. However, these are usually treated as clinical endpoints or cost inputs rather than as manifest attributes requiring linear ratio measurement. The curriculum recognizes observation more readily than it recognizes measurement.

8. A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute — Probability 0.10; Logit -2.00

Recognition of latent attribute measurement is weak. Dutch HTA education includes health-related quality of life, QALYs, and patient outcomes, but these are generally introduced through utility instruments and economic evaluation rather than through a prior theory of latent measurement. The requirement that latent attributes require a measurement model capable of estimating possession of the attribute is therefore weakly represented.

9. Manifest and latent attributes require different forms of ratio measurement — Probability 0.05; Logit -2.50

This proposition receives virtually no endorsement. The Dutch curriculum knowledge base does not appear to distinguish systematically between linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes. This is a decisive failure because it removes the gateway distinction required for measurement-based HTA. Without this distinction, clinical events, resource utilization, quality-of-life scores, utilities, and QALYs can be brought into the same analytical framework despite lacking equivalent measurement status.

10. Therapy impact claims must be falsifiable — Probability 0.15; Logit -1.75

The requirement for falsifiable therapy-impact claims receives weak endorsement. Dutch HTA education and guidance emphasize economic evaluation, modelling, cost-effectiveness, uncertainty, and policy relevance. The revised Dutch guideline is described as guidance for economic evaluations to inform health policy decisions from a broad societal perspective. This is decision support, not a falsification framework. A falsifiable claim would require prospective specification, lawful measurement, a defined comparator, a timeframe, success criteria, independent replication, and conditions for rejection.

MANIFEST AND LATENT ATTRIBUTES: THE MISSING FOUNDATION OF DUTCH HTA

A central finding of the Dutch curriculum interrogation is the absence of any explicit framework distinguishing manifest from latent attributes and the corresponding measurement requirements that follow from this distinction. This omission is particularly significant because the Netherlands has been one of the principal architects of the contemporary European reference-case tradition. The manifest-latent distinction is not a methodological refinement but one of the foundational concepts of representational measurement. Without it there is no coherent basis for determining how therapy outcomes should be measured, what constitutes a lawful quantitative claim, or whether the arithmetic applied within HTA is scientifically admissible.

Manifest attributes are directly observable. Their existence and magnitude can be established through empirical observation without the need for an intervening measurement model. Examples include survival time, hospital admissions, emergency department visits, medication adherence, treatment discontinuation, adverse events, laboratory values, and healthcare resource utilization. These attributes are observable phenomena that can be counted, timed, or otherwise recorded directly. When properly specified, they support linear ratio measures with a meaningful zero and lawful arithmetic operations. The scientific task is straightforward: define the attribute, establish the unit of observation, specify the observation period, and subject the resulting claim to empirical evaluation and independent replication.

Latent attributes present a fundamentally different measurement problem. Pain, fatigue, depression, anxiety, physical functioning, treatment satisfaction, health confidence, and need fulfilment cannot be observed directly. Their existence must be inferred from observable indicators, typically responses to structured questionnaire items. Consequently, latent attributes require a measurement model capable of estimating possession of the attribute while satisfying the axioms of representational measurement. Within HTA this requirement is met through Rasch measurement, which provides a logit ratio measure possessing the properties required for lawful quantitative inference.

The importance of this distinction cannot be overstated. Manifest and latent attributes require fundamentally different forms of ratio measurement. Manifest attributes require linear ratio measures; latent attributes require Rasch logit ratio measures. These are the only two measurement frameworks capable of supporting lawful quantitative claims regarding therapy impact. The distinction therefore determines the entire measurement strategy adopted within HTA.

The interrogation suggests that the Dutch curriculum knowledge base does not recognize this distinction as an organizing principle for HTA education. Instead, outcomes are grouped within broad categories such as patient-reported outcomes, quality of life, utility measurement, clinical effectiveness, and value assessment. These categories are descriptively useful but scientifically inadequate because they fail to distinguish between directly observable attributes and attributes requiring a measurement model. The measurement requirements governing each type of outcome consequently remain obscured.

Once this distinction disappears, virtually any numerical output can be treated as though it possesses equivalent measurement status. Utility scores, preference weights, composite indices, symptom scales, observational counts, and resource-use measures are brought together within the same analytical framework despite possessing fundamentally different measurement properties. The consequence is a loss of measurement discipline in which numerical construction is mistaken for quantitative measurement.

For the Netherlands this finding carries particular significance. Dutch researchers and institutions have played a leading role in developing and disseminating the European reference-case tradition. Yet the interrogation indicates that the educational framework underpinning that tradition does not begin with the most fundamental distinction required by representational measurement. The result is that successive generations of practitioners become highly proficient in economic evaluation, cost-utility analysis, and decision modelling without first asking the elementary scientific question: what type of attribute is being measured, and what form of measurement is required?

THE ABSENCE OF RASCH

The absence of the manifest-latent distinction also explains the near absence of Rasch measurement within the Dutch curriculum knowledge base. If latent attributes are not explicitly recognized as requiring a measurement model, there is no perceived need to introduce the one measurement framework capable of constructing lawful quantitative measures of latent attribute possession. Instead, ordinal questionnaire responses are transformed into utility scores, indices, and QALYs and subsequently treated as though measurement has already been achieved. The measurement problem is therefore bypassed rather than solved.

From the perspective of curriculum design, this represents a classic example of curriculum inversion. Students are introduced to utilities, QALYs, patient-reported outcomes, and reference-case modelling before they are taught to distinguish between manifest and latent attributes. Consequently, they are never encouraged to ask the critical scientific question that should precede every HTA submission: What is the target attribute, is it manifest or latent, and which lawful form of ratio measurement is required?

For a country that has exercised such influence over the development of European HTA, this omission is particularly consequential. The interrogation suggests that the Dutch curriculum knowledge base has helped disseminate methodological sophistication without first establishing the scientific foundations of measurement. Until the distinction between manifest and latent attributes becomes the organizing principle of HTA education, curriculum inversion will continue to reproduce measurement inversion, not only within the Netherlands but throughout the broader European HTA tradition that Dutch methodology has helped shape.

One of the most striking findings from the interrogation of the Dutch HTA curriculum knowledge base is not simply the absence of representational measurement but the near-complete absence of Rasch measurement and its role in the assessment of latent attributes. This omission is particularly significant because the Netherlands has been one of the principal contributors to the development of modern European HTA. Dutch researchers and institutions have played a leading role in advancing cost-utility analysis, reference-case methodology, and economic evaluation. Yet the educational framework underpinning this tradition gives little recognition to the one measurement framework capable of constructing lawful quantitative measures of latent attributes. The issue is not whether the term *Rasch* occasionally appears in specialist publications or psychometric research. The issue is whether Rasch measurement is recognized as the scientific foundation for measuring latent therapy outcomes. The interrogation indicates that it is not.

This places Rasch in a unique position. Rasch is not simply another psychometric technique competing with item response theory, PROMIS, utility instruments, or preference-based scoring systems. It addresses an entirely different scientific problem. Its purpose is to determine whether ordinal observations can be transformed into a quantitative measure of possession of a latent attribute. In doing so, it provides the only established framework capable of demonstrating whether the conditions required for measurement have been satisfied.

The distinction is fundamental. Patient-reported outcomes begin with ordinal responses to questionnaire items describing pain, fatigue, anxiety, physical functioning, treatment satisfaction, need fulfilment, or other latent attributes. These responses provide information about order but not quantity. Arithmetic performed directly on ordinal observations cannot create measurement. Summed scores, averages, weighted algorithms, utility values, and preference functions remain numerical constructions unless it has first been demonstrated that the underlying latent attribute has been measured. Numerical manipulation is not measurement.

Rasch measurement was developed precisely to solve this problem. Through the conjoint calibration of persons and items, the Rasch model estimates the location of individuals on a latent continuum while simultaneously testing whether the observations satisfy the conditions required for lawful measurement. Unidimensionality, specific objectivity, invariance, item fit, responsecategory functioning, local independence, and differential item functioning are not optional methodological refinements. They are the empirical conditions that must be satisfied before a claim regarding possession of a latent attribute can be advanced. Rasch therefore provides both the measurement model and the empirical tests necessary to determine whether measurement has actually been achieved.

The interrogation indicates that this perspective is largely absent from the Dutch curriculum knowledge base. Students and practitioners are introduced to patient-reported outcomes, utilities, quality-of-life instruments, preference weights, QALYs, and economic evaluation without first confronting the measurement problem these constructs are intended to solve. The educational sequence moves directly from ordinal responses to scoring systems and economic models. The essential intermediate step—demonstrating that a latent attribute has been measured—is effectively bypassed.

This omission has profound implications. Without Rasch measurement, latent attributes remain unmeasured. Utility scores, composite indices, and preference algorithms may generate numerical outputs, but they do not establish that the underlying construct has been measured. The existence of a number should never be confused with the existence of a measure. Nevertheless, much of contemporary Dutch HTA proceeds as though this distinction does not exist.

Students consequently learn how utilities are generated, how QALYs are constructed, and how reference-case models are populated, yet they are not taught how latent attributes themselves can be measured. The curriculum therefore introduces numerical representations of patient benefit before introducing the scientific framework required to justify those representations. This is curriculum inversion in its clearest form.

For the Netherlands this conclusion carries particular weight. One of the countries that helped shape the European reference-case tradition has not incorporated into its educational framework the only established measurement theory capable of supporting quantitative claims for latent therapy outcomes. Until Rasch measurement assumes its proper place within Dutch HTA education, latent attributes will continue to be represented through scores, utilities, and indices rather than lawful measures, and the distinction between scoring and measurement will remain obscured. The consequence is that curriculum inversion continues to reinforce the broader pattern of measurement inversion that characterizes contemporary European HTA.

CONCLUSION: CURRICULUM INVERSION AND THE FUTURE OF HTA IN THE NETHERLANDS

The companion interrogation of the Dutch HTA knowledge base demonstrated that contemporary health technology assessment in the Netherlands is characterized by measurement inversion. The accepted principles of representational measurement receive little recognition, while utilities, QALYs, cost-effectiveness analysis, and reference-case modelling continue to be treated as though they provide scientifically valid quantitative measures of therapy impact. The present study has addressed the complementary question of how such a framework has been created, sustained, and transmitted across successive generations of practitioners. The answer is curriculum inversion.

The interrogation demonstrates that the Dutch HTA curriculum knowledge base introduces students and practitioners to the methods of contemporary HTA before introducing them to the scientific principles required to evaluate those methods. Educational emphasis is placed upon comparative effectiveness, economic evaluation, utilities, QALYs, decision modelling, evidence synthesis, and reimbursement methodology rather than upon specification of the target attribute, representational measurement, admissible arithmetic, unidimensionality, manifest and latent attributes, Rasch measurement, dimensional homogeneity, and the requirement that therapy-impact claims be prospectively evaluable, independently replicable, and capable of falsification. Graduates therefore acquire technical competence in constructing the reference-case framework without first acquiring the scientific framework necessary to determine whether that framework supports lawful quantitative claims.

For the Netherlands this conclusion carries particular significance. Dutch universities, researchers, and policy organizations have been among the principal architects of the European reference-case tradition. They have helped shape the methodological framework subsequently adopted across much of Europe. The interrogation therefore indicates that curriculum inversion is not simply a feature of Dutch HTA; it is one of the mechanisms through which measurement inversion has been disseminated internationally. The European reference-case paradigm was not built upon the foundations of representational measurement. It was built upon administrative necessity. Every major component of that framework—utilities, QALYs, cost-effectiveness ratios, and reference-case simulation modelling—depends upon quantities that fail the accepted requirements for lawful measurement. The result is methodological sophistication applied to measurement impossibilities.

This conclusion leaves no room for incremental reform. A framework whose constituent quantities fail the axioms of representational measurement cannot be rescued through improved modelling, better utility instruments, more sophisticated statistics, expanded real-world evidence, or increasingly elaborate simulations. None of these activities can create measurement where measurement does not exist. Arithmetic cannot rescue quantities that are not measures. Forty years of methodological refinement have not resolved this problem because the problem lies beneath the methodology itself.

The implications for Dutch universities are therefore direct. Continuing to educate students in the construction and application of reference-case models without first teaching the scientific principles governing lawful measurement is incompatible with the central purpose of higher education. Universities exist to advance objective knowledge, not to perpetuate analytical traditions whose measurement foundations have collapsed. Students should certainly understand the historical importance of utilities, QALYs, and reference-case modelling, but they should understand them as examples of a paradigm that failed to satisfy the accepted standards governing quantitative science. They should not be taught that these constructs provide valid measures of therapy impact because they do not.

The only scientifically defensible path forward is reconstruction. Every HTA submission should begin with explicit specification of the target attribute. Manifest attributes should be evaluated through lawful linear ratio measures. Latent attributes should be evaluated through Rasch logit ratio measures that satisfy the axioms of representational measurement. Every quantitative claim should be accompanied by a prospective protocol specifying the attribute, measurement method, comparator, evaluation period, replication strategy, and conditions for empirical falsification. Measurement must precede arithmetic in every assessment without exception.

Taken together, the companion interrogations of measurement inversion and curriculum inversion demonstrate that the Dutch reference-case tradition has reached its intellectual limit. It has succeeded in standardizing administrative decision making while failing to establish the scientific legitimacy of the quantities upon which those decisions depend. That paradigm has no scientific future because it rests upon measurement impossibilities. Its place is no longer at the center of HTA education but in the history of the discipline as an instructive example of how methodological sophistication can evolve independently of measurement science.

To facilitate this transition, Maimon Research LLC has developed a comprehensive nine-unit HTA Reconstruction Program [8]. The program provides a systematic introduction to representational measurement, the theory of attributes, the principal scales of measurement, admissible arithmetic, dimensional homogeneity, manifest and latent attributes, Rasch logit ratio measurement, protocol development, and the construction of evaluable, replicable, and falsifiable claims regarding therapy impact. Its purpose is not to modify the existing reference-case paradigm but to replace it with a scientific framework in which measurement once again precedes arithmetic.

The program has been designed for universities, HTA agencies, reimbursement organizations, research centers, professional societies, pharmaceutical companies, and health economists seeking a transition from assumption-driven modelling to scientifically defensible measurement. It provides a structured pathway for professional development while establishing the competencies required for the next generation of HTA practitioners. In this way, it offers not simply a critique of the existing paradigm but a practical route toward the reconstruction of HTA as a measurement-based scientific discipline.

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] The Netherlands: The Dutch contribution to nonsense on stilts in health technology assessment. Logit Working Paper No 122. https://maimonresearch.com/logit-working-paper-no-122/

[2] The Netherlands: Measurement failure in national health technology assessment guidelines. Logit Working Paper No 120. https://maimonresearch.com/logit-working-paper-no-120/

[3] The Netherlands: Measurement failure vt Dutch academic research for health technology assessment. Log it Working Paper No 121. https://maimonresearch.com/logit-working-paper-no-121/

[4] The end of the reference case: Reconstructing HTA. Logit Working Paper no 345. https://maimonresearch.com/logit-working-paper-no-345-june-2026/

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

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

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

[8] Maimon Research. HTA Reconstruction. https://maimonresearch.com/hta-reconstruction-program-and-fees/