RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

GERMANY: CURRICULUM INVERSION IN HEALTH TECHNOLOGY ASSESSMENT

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

LOGIT WORKING PAPER No 1470 JULY 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

This study examines the German health technology assessment (HTA) curriculum knowledge base as the educational complement to an earlier interrogation of the German HTA academic knowledge base that identified a pervasive pattern of measurement inversion. While the previous assessment demonstrated that the scientific principles required for lawful quantitative measurement receive little recognition within German HTA methodology, the present investigation asks how this methodological framework is transmitted to successive generations of researchers and practitioners. The central hypothesis is that measurement inversion is sustained through curriculum inversion, whereby students are introduced to accepted HTA methods before they are taught the scientific principles necessary to evaluate those methods.

The curriculum knowledge base was interrogated using ten canonical statements derived from the axioms of representational measurement. These statements encompass specification of the target attribute, the principal scales of measurement, 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 empirically falsifiable. Collectively, these concepts define the minimum scientific foundations required for quantitative therapy assessment. Across the interrogation, endorsement probabilities indicate strong recognition of directly observable, manifest attributes but consistently weak recognition of the principles governing measurement itself. Representational measurement, admissible arithmetic, measurement status, unidimensionality, Rasch measurement, and the distinction between linear ratio and Rasch logit ratio measurement occupy only a marginal position within the German curriculum knowledge base.

The findings identify a distinctive German pattern of curriculum inversion. Germany has established an international reputation for evidence-based medicine, comparative clinical effectiveness, benefit assessment, clinical epidemiology, and rigorous evaluation of patient-relevant outcomes. Unlike several other European jurisdictions, Germany has not adopted the cost-per-QALY reference case as the central basis for reimbursement decisions. Nevertheless, the interrogation demonstrates that movement away from the reference-case paradigm has not been accompanied by adoption of the scientific principles required for measurement-based HTA. Students become proficient in evidence synthesis, comparative benefit assessment, patient-reported outcomes, health economic evaluation, and decision modelling without first acquiring the conceptual framework necessary to determine whether the quantities entering those analyses constitute lawful measures.

The study concludes that German HTA requires reconstruction rather than incremental methodological refinement. A scientifically defensible curriculum must begin with explicit specification of the target attribute, distinguish manifest from latent attributes, employ lawful linear ratio and Rasch logit ratio measurement, and require that all therapy-impact claims be prospectively evaluable, independently replicable, and empirically falsifiable. Only by restoring representational measurement as the foundation of quantitative reasoning can German HTA reconcile its internationally respected tradition of methodological rigor with the scientific standards required for lawful quantitative measurement.

INTRODUCTION

The companion assessment of the German HTA academic knowledge base concluded that measurement inversion is a pervasive characteristic of contemporary German health technology assessment (HTA) [1] . The present study addresses a different but closely related question. If the scientific principles required for lawful quantitative claims receive little recognition within the academic knowledge base, to what extent are those same principles represented within the educational knowledge base responsible for training the next generation of HTA professionals?

The focus of the present assessment is therefore curriculum inversion. Whereas measurement inversion concerns the relationship between measurement and arithmetic within HTA methodology, curriculum inversion concerns the sequence in which those concepts are taught. A scientifically coherent curriculum would first establish the principles governing quantitative measurement and only then introduce the analytical methods that depend upon those principles. If students are instead introduced to utilities, economic evaluation, cost-effectiveness analysis, decision modelling, registry analysis, real-world evidence, and related HTA techniques without first acquiring the scientific foundations of representational measurement, then the curriculum itself has become inverted.

To examine this question, the German HTA curriculum knowledge base was interrogated using a fixed set of ten canonical curriculum statements. These statements represent the minimum scientific concepts that should be encountered within a curriculum intended to prepare graduates for quantitative therapy assessment. They encompass the specification of attributes, the identification of target attributes, the principal scales of measurement [2], the requirement that measurement precede quantitative analysis, the axioms of representational measurement [3] , unidimensionality, the distinction between manifest and latent attributes, the recognition that manifest and latent attributes require different forms of ratio measurement, Rasch logit ratio attribute possession [4] and the requirement that therapy-impact claims be prospectively evaluable and falsifiable.

The purpose of the interrogation is not to evaluate individual lecturers, courses, or degree programs. Nor is it to judge the quality of German HTA research. Rather, it asks a more fundamental question: does the publicly accessible educational knowledge base expose students to the scientific principles that must precede quantitative claims regarding therapy impact? If these principles receive strong endorsement, the curriculum can be regarded as providing the conceptual foundations for measurement-based HTA. If they receive little endorsement while accepted HTA methods are strongly represented, then the curriculum demonstrates curriculum inversion.

The importance of this question extends beyond Germany. German universities educate researchers and practitioners who contribute to HTA agencies, research institutes, the pharmaceutical industry, consultancy, and the broader European HTA community. The concepts emphasized within the curriculum therefore influence not only professional education but also the future direction of HTA research, methodological development, journal publication, and reimbursement practice. The present assessment seeks to determine whether the German HTA educational environment equips future practitioners to evaluate the scientific legitimacy of accepted HTA methods or primarily prepares them to apply those methods without first establishing their measurement foundations.

As a member state of the European Union, Germany is subject to a further influence that reinforces both measurement inversion and curriculum inversion: the progressive harmonization of health technology assessment through the European HTA framework [5] . Over the past two decades, first through EUnetHTA and now through the implementation of the Joint European HTA Regulation, the European Union has developed a common methodological framework intended to standardize the assessment of therapeutic interventions across member states. While designed to improve consistency and reduce duplication, this framework has largely adopted the analytical assumptions of the NICE reference-case paradigm. In doing so, it has institutionalized a methodology in which utilities, cost-effectiveness analysis, and model-based evaluation are treated as established scientific practice without first demonstrating that they satisfy the requirements of representational measurement. . In practical terms, the EU exports measurement inversion and curriculum inversion to member states.

The consequence extends well beyond reimbursement methodology. Once incorporated into European guidance, the reference-case framework influences university curricula, professional training, journal publication standards, methodological guidance, peer review, and the education of future HTA practitioners. Measurement inversion therefore becomes curriculum inversion. Students are taught accepted analytical methods before they are taught the scientific principles required to determine whether those methods can legitimately support quantitative claims regarding therapy impact. In this way, methodological convention is reproduced through education, professional certification, and policy rather than being continually evaluated against the axioms of representational measurement.

Germany occupies a distinctive position within this European landscape. Unlike many member states, it has not adopted the NICE-style cost-per-QALY reference case as the central basis for reimbursement decision making. The AMNOG framework, together with the work of IQWiG and the Federal Joint Committee (G-BA), places primary emphasis on comparative clinical benefit rather than incremental cost per QALY. Nevertheless, this departure does not amount to a measurement-based reconstruction of HTA. The German framework continues to operate without explicit recognition that measurement must precede arithmetic, without distinguishing between linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes, and without grounding quantitative therapy-impact claims in the axioms of representational measurement. Germany has therefore moved beyond the reference case but not beyond the analytical framework from which the reference case emerged. It represents a case of change without change**:** the abandonment of one methodological expression of measurement inversion without the adoption of the scientific principles necessary to replace it.

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 GERMAN CURRICULUM KNOWLEDGE BASE

The first step in any LLM curriculum interrogation is to define the knowledge base to be interrogated. The validity of the assessment depends upon ensuring that the knowledge base reflects the information environment that shapes HTA education, research training, professional formation, and policy-facing practice. A curriculum knowledge base is more than a collection of course descriptions. It represents the accumulated concepts, assumptions, methods, definitions, and analytical frameworks through which students, researchers, agency staff, consultants, and professional practitioners acquire their understanding of HTA.

In the present assessment, the German HTA curriculum knowledge base is defined as the publicly accessible body of materials that contributes to HTA education and professional formation in Germany. This includes university program descriptions, course outlines, doctoral training resources, research-center educational materials, methodological guidance, IQWiG and G-BA documents, EUnetHTA-related materials, seminars, workshops, faculty publications, professional reports, conference presentations, policy briefs, journal articles, and other resources through which HTA concepts are communicated and reinforced. Taken together, these materials provide a representation of the concepts and analytical frameworks that students, researchers, and professional staff are most likely to encounter within the German HTA educational environment.

The importance of defining a national curriculum knowledge base extends beyond formal course content. National HTA systems do more than apply methods. They define what counts as legitimate evidence, what questions are regarded as relevant, what assumptions are accepted, what analytical techniques are treated as standard, and what competencies are expected of future practitioners. Over time, these priorities shape the intellectual culture of HTA education. Concepts that are consistently emphasized become embedded in professional practice. Concepts that are neglected disappear from education, research, agency guidance, journal expectations, and policy assessment.

A German HTA curriculum designed to support scientific evaluation of therapeutic interventions would be expected to expose students and practitioners to the foundational principles required for quantitative reasoning. These include specification of the target attribute, the principal scales of measurement, representational measurement, admissible arithmetic, unidimensionality, dimensional homogeneity, the distinction between manifest and latent attributes, linear ratio measurement for manifest attributes, Rasch logit ratio measurement for latent attributes, and the requirement that therapy-impact claims be capable of empirical evaluation, replication, and potential falsification.

The objective of the present interrogation is therefore not to evaluate a single German university, research center, agency, course, faculty member, or teaching style. Rather, it is to assess the broader national educational environment represented by the German HTA curriculum knowledge base. The question is straightforward: to what extent are the concepts required for measurement-based therapy assessment present, absent, or weakly represented within the German educational and methodological environment? The answer indicates whether future German HTA researchers, analysts, agency staff, 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 established analytical techniques.

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.

GERMANY AND CURRICULUM INVERSION

The interrogation of the German health technology assessment (HTA) curriculum knowledge base demonstrates a striking contrast between strong recognition of directly observable clinical outcomes and very weak recognition of the scientific principles governing quantitative measurement. Across the ten canonical curriculum statements, endorsement probabilities range from only 0.05 to 0.85. While the curriculum gives overwhelming support to the concept of manifest, directly observable attributes, nearly all of the concepts required to establish lawful quantitative measurement receive consistently weak endorsement. The resulting profile indicates that German HTA education emphasizes evidence acquisition and clinical observation far more strongly than the scientific foundations of measurement itself.

TABLE 1: CURRICULUM CONTENT ENDORSEMENT - GERMANY

CANONICAL STATEMENT	CATEGORICAL PROBABILITY	NORMALIZED LOGIT
An attribute is the specific outcome of interest in a therapy assessment	0.20	-1.40
Every therapy assessment begins with specification of the target attribute	0.15	-1.75
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	0.05	-2.50
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	-2.20
A manifest attribute is directly observable and capable of supporting empirical observation	0.85	+1.75
A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute	0.15	-1.75
Manifest and latent attributes require different forms of ratio measurement	0.05	-2,50
Therapy impact claims must be falsifiable	0.15	-1.75

The strongest endorsement is given to the statement that a manifest attribute is directly observable and capable of supporting empirical observation (probability 0.85). This finding is consistent with

Germany's internationally recognized emphasis on patient-relevant clinical outcomes, comparative clinical effectiveness, evidence-based medicine, and rigorous evaluation of observable therapy endpoints. The curriculum therefore demonstrates considerable strength in recognizing the importance of empirically observable outcomes. More modest recognition is given to the concept of an attribute as the outcome of interest in therapy assessment (probability 0.20), while specification of the target attribute, the principal scales of measurement, recognition of latent attributes, and the requirement that therapy-impact claims be falsifiable each receive only weak endorsement, with probabilities ranging between 0.15 and 0.20.

Recognition becomes markedly weaker for the principles that distinguish scientific measurement from numerical analysis. The requirement that attributes be demonstrated to be unidimensional before measurement is possible receives a probability of only 0.10, while the requirement that the measurement status of a target attribute be established before quantitative claims are advanced receives only 0.05. These findings indicate that German HTA education gives little attention to determining whether an empirical quantity actually possesses the mathematical properties required to support quantitative inference. Existing outcome measures, patient-reported scores, and preference-based instruments appear largely to be accepted without first demonstrating that they constitute lawful measures.

The weakest endorsement is reserved for the axioms of representational measurement and the principle that manifest and latent attributes require different forms of ratio measurement, both receiving probabilities of only 0.05. These are among the most fundamental concepts in measurement science because they determine whether arithmetic operations are scientifically admissible. Their near absence from the curriculum indicates that students are unlikely to encounter the principles governing admissible transformations, dimensional homogeneity, or the distinction between linear ratio measurement for manifest attributes and Rasch logit ratio measurement for latent attributes. Consequently, numerical outputs derived from fundamentally different types of data may be treated as though they possess equivalent measurement status.

Taken together, these findings demonstrate that the German HTA curriculum exhibits the defining characteristics of curriculum inversion. Students become highly proficient in evaluating clinical evidence, patient-relevant outcomes, comparative effectiveness, economic evaluation, utilities, quality-adjusted life years (QALYs), and decision-analytic modelling, yet they receive little formal instruction in the scientific principles that determine whether the quantities entering those analyses are lawful measures. Germany therefore presents a distinctive form of curriculum inversion. Unlike many countries where even observable attributes receive only modest recognition, the German curriculum displays considerable strength in empirical observation while simultaneously neglecting the representational measurement framework required to transform observation into lawful quantitative science. The result is an educational system in which measurement inversion is perpetuated not through lack of methodological sophistication, but through failure to establish the scientific foundations upon which all quantitative claims ultimately depend.

STATEMENT-BY-STATEMENT REVIEW

Statement 1: An attribute is the specific outcome of interest in a therapy assessment
Probability 0.20; Normalized Logit –1.40

Recognition of the concept of an attribute is limited within the German HTA curriculum knowledge base. Although German HTA gives considerable attention to patient-relevant outcomes, clinical effectiveness, and comparative benefit, these outcomes are not consistently introduced as formally specified attributes that must be identified before measurement begins. Consequently, assessment frequently proceeds directly to available endpoints, clinical evidence, or economic evaluation without first establishing precisely what characteristic of therapy impact is to be measured. The low probability indicates that explicit specification of the attribute is not yet embedded as the starting point for quantitative evaluation.

Statement 2: Every therapy assessment begins with specification of the target attribute
Probability 0.15; Normalized Logit –1.75

This principle receives only weak endorsement. The German curriculum appears to begin therapy assessment with accepted HTA procedures, evidence review, comparative effectiveness, or benefit assessment rather than with formal specification of the target attribute. This reverses the scientific sequence required by representational measurement. The target attribute should first be identified, followed by determination of whether it is manifest or latent and selection of the appropriate measurement strategy. Without this sequence, available methodologies rather than scientific measurement determine the structure of the assessment.

Statement 3: The principal scales of measurement (nominal, ordinal, interval and ratio) have different properties and support different forms of analysis
Probability 0.15; Normalized Logit –1.75

Recognition of the principal scales of measurement is weak. While German students are likely to encounter these concepts within statistics and research methods, there is little indication that the mathematical implications of scale type are integrated into HTA methodology. In particular, the curriculum provides little evidence that students are taught that arithmetic operations must be constrained by the properties of the underlying measurement scale. Ordinal observations therefore risk being manipulated as though they possessed interval or ratio properties without demonstrating that these operations are scientifically admissible.

Statement 4: The measurement status of a target attribute must be established before quantitative claims can be advanced
Probability 0.05; Normalized Logit –2.50

This statement receives one of the weakest endorsements in the interrogation. It represents the defining characteristic of measurement inversion. German HTA education appears to assume that clinical outcomes, utility measures, patient-reported scores, and composite indices possess suitable measurement properties without first establishing whether they satisfy the conditions required for lawful quantitative inference. Consequently, arithmetic and modelling proceed before measurement has been demonstrated. The scientific question of whether the underlying quantity is actually measurable receives very little attention within the curriculum.

Statement 5: The axioms of representational measurement underpin quantitative claims
Probability 0.05; Normalized Logit –2.50

Recognition of representational measurement is almost entirely absent. This finding is particularly important because the axioms of representational measurement determine whether empirical observations have been mapped into numerical relations capable of supporting quantitative analysis. Germany has a strong international reputation for methodological rigor in evidence-based medicine and benefit assessment, yet this rigor does not appear to extend to the scientific foundations of measurement itself. Students are therefore unlikely to encounter representational measurement as the framework governing the legitimacy of quantitative claims.

Statement 6: Attributes must be demonstrated to be unidimensional before measurement is possible
Probability 0.10; Normalized Logit –2.20

Recognition of unidimensionality is very limited. The curriculum provides little indication that an attribute must first be demonstrated to represent a single measurable construct before numerical measurement can proceed. Composite indices and multidimensional patient-reported instruments therefore risk being accepted as quantitative measures despite combining multiple distinct attributes. Without unidimensionality, there is no scientific basis for treating a composite score as a measure of a single therapy outcome.

Statement 7: A manifest attribute is directly observable and capable of supporting empirical observation
Probability 0.85; Normalized Logit +1.75

This statement receives by far the strongest endorsement within the German interrogation. The result reflects Germany's well-established emphasis on patient-relevant clinical outcomes, observable endpoints, comparative effectiveness, and rigorous empirical evaluation. Manifest attributes such as mortality, morbidity, adverse events, hospitalization, and treatment discontinuation occupy a central position within German HTA. However, recognition of directly observable outcomes should not be confused with recognition of measurement itself. Even manifest attributes require explicit definition, lawful units, appropriate ratio measurement, and prospectively specified protocols before quantitative claims can be regarded as scientifically valid.

Statement 8: A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute
Probability 0.15; Normalized Logit –1.75

Recognition of latent attributes is weak. Although patient-reported outcomes are widely employed within German HTA, there is little evidence that the curriculum systematically teaches that latent constructs require formal measurement models before quantitative claims can be advanced. Instead, questionnaire scores, utility values, and preference-based measures appear to be accepted largely through convention rather than through demonstration that the underlying latent attribute has been lawfully measured. Rasch measurement, the only established framework for constructing ratio measures of latent attributes, occupies little place within the curriculum.

Statement 9: Manifest and latent attributes require different forms of ratio measurement
Probability 0.05; Normalized Logit –2.50

This statement receives one of the weakest endorsements in the interrogation. The German curriculum knowledge base demonstrates little recognition that manifest attributes require linear ratio measurement while latent attributes require Rasch logit ratio measurement. This distinction lies at the heart of representational measurement because it determines the appropriate measurement framework for every therapy outcome. Without it, directly observable outcomes, patient-reported scores, utilities, preference weights, and composite indices are easily incorporated into a common analytical framework despite possessing fundamentally different measurement properties.

Statement 10: Therapy impact claims must be falsifiable
Probability 0.15; Normalized Logit –1.75

Recognition that therapy-impact claims must be prospectively evaluable, independently replicable, and capable of empirical falsification is weak. German HTA places considerable emphasis on evidence quality, comparative effectiveness, and patient benefit, yet these activities are not equivalent to designing claims that can subsequently be confirmed or refuted through observation. The curriculum therefore appears to emphasize assessment of existing evidence more strongly than the formulation of scientifically testable claims. As a consequence, many quantitative conclusions remain dependent upon accepted methodological conventions rather than upon prospective empirical evaluation.

MANIFEST AND LATENT ATTRIBUTES: THE MISSING FOUNDATION OF GERMAN HTA

A central finding of the German 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 distinction between manifest and latent attributes is not a methodological refinement but one of the foundational principles 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 employed within HTA is scientifically admissible. The distinction establishes the entire logic of measurement and determines the appropriate quantitative framework for evaluating therapy impact.

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 represent observable phenomena that can be counted, timed, or otherwise recorded directly. When properly specified they support linear ratio measures possessing a meaningful zero and therefore permit lawful arithmetic operations. The scientific task is straightforward: define the attribute, specify the unit of observation, determine the observation period, and subject the resulting claim to empirical evaluation, independent replication, and potential falsification.

Latent attributes present an entirely different measurement problem. Pain, fatigue, depression, anxiety, physical functioning, treatment satisfaction, health confidence, and need fulfilment cannot be observed directly. Their existence must instead be inferred from observable indicators, most commonly 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 mathematical properties necessary for lawful quantitative inference. Without such a model, ordinal observations remain ordinal observations irrespective of the sophistication of subsequent statistical analysis.

The importance of this distinction cannot be overstated. Manifest attributes require linear ratio measures, whereas 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. Failure to recognize this distinction inevitably leads to inappropriate treatment of fundamentally different forms of evidence as though they possessed equivalent measurement status.

The interrogation indicates that the German curriculum knowledge base does not recognize this distinction as an organizing principle for HTA education. Instead, outcomes are grouped within broad methodological categories such as patient-relevant outcomes, patient-reported outcomes, quality of life, utility measurement, comparative clinical effectiveness, evidence synthesis, benefit assessment, and economic evaluation. Although useful for administrative and descriptive purposes, these categories are scientifically inadequate because they fail to distinguish directly observable attributes from latent attributes requiring a formal measurement model. The measurement requirements governing each type of outcome therefore remain obscured.

Once this distinction disappears, almost any numerical output can be treated as though it possesses equivalent quantitative status. Utility scores, preference weights, composite indices, symptom scales, observational counts, registry-derived outcomes, and resource-use measures become incorporated within a common analytical framework despite possessing fundamentally different measurement properties. The result is a progressive erosion of measurement discipline in which numerical construction is mistaken for quantitative measurement and arithmetic replaces measurement as the foundation of scientific inference.

For Germany this finding is particularly significant. German universities, research institutes, and national HTA organizations have established an international reputation for evidence-based medicine, comparative clinical effectiveness, benefit assessment, clinical epidemiology, health economics, and rigorous evaluation of patient-relevant outcomes. Indeed, the present interrogation demonstrates exceptionally strong recognition of directly observable, manifest attributes. The curriculum interrogation indicates, however, that the educational framework supporting these achievements does not proceed from the fundamental distinction required by representational measurement. Successive generations of practitioners therefore become highly proficient in evidence synthesis, benefit assessment, health economic evaluation, cost-utility analysis, and decision modelling without first addressing the elementary scientific question that should precede every quantitative assessment: what type of attribute is being measured, and what form of measurement is scientifically required? Until the distinction between manifest and latent attributes becomes the organizing principle of German HTA education, curriculum inversion will continue to reproduce measurement inversion, limiting the scientific foundations upon which quantitative claims regarding therapy impact are constructed.

THE ABSENCE OF RASCH

The absence of the manifest-latent distinction also explains the near absence of Rasch measurement within the German 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, composite indices, and quality-adjusted life years (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 patient-relevant outcomes, utilities, QALYs, patient-reported outcomes, benefit assessment, health economic evaluation, and reference-case modelling before they are taught to distinguish between manifest and latent attributes. Consequently, they are never encouraged to ask the fundamental scientific question that should precede every HTA assessment: What is the target attribute, is it manifest or latent, and which lawful form of ratio measurement is required?

For Germany this omission is particularly significant because health technology assessment occupies a central position in decisions concerning the comparative benefit, reimbursement, pricing, and adoption of pharmaceuticals, medical devices, diagnostics, and other healthcare interventions. German HTA is internationally recognized for its emphasis on patient-relevant outcomes and comparative clinical evidence. Indeed, the present interrogation demonstrates exceptionally strong recognition of directly observable, manifest attributes. Yet the curriculum suggests that these methodological strengths have been developed without first establishing the scientific foundations of measurement itself. Until the distinction between manifest and latent attributes becomes the organizing principle of German HTA education, curriculum inversion will continue to reproduce measurement inversion throughout German teaching, research, and policy evaluation.

One of the most striking findings from the interrogation of the German HTA curriculum knowledge base is therefore 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 Germany has established an international reputation for evidence-based medicine, clinical epidemiology, comparative effectiveness research, patient-relevant outcome assessment, health economics, and health technology assessment. German researchers have made important methodological contributions to evidence synthesis and benefit assessment, yet the educational framework underpinning these activities 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 within specialist psychometric publications or methodological 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 lawful measurement have been satisfied.

The distinction is fundamental. Patient-reported outcomes begin with ordinal responses to questionnaire items describing pain, fatigue, depression, anxiety, physical functioning, treatment satisfaction, need fulfilment, and 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, response-category 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 German curriculum knowledge base. Students and practitioners are introduced to patient-reported outcomes, utilities, quality-of-life instruments, preference weights, QALYs, comparative benefit assessment, 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 German HTA proceeds as though this distinction does not exist. Students consequently learn how utilities are generated, how QALYs are constructed, how patient benefit is synthesized, and how economic 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 Germany the implications are direct. A healthcare system internationally recognized for its commitment to evidence-based medicine, patient-relevant outcomes, comparative benefit assessment, and rigorous clinical evaluation should equally require that the quantities informing reimbursement and policy decisions satisfy the accepted standards of representational measurement. Until Rasch measurement assumes its proper place within German HTA education, latent attributes will continue to be represented through scores, utilities, and composite indices rather than lawful measures, and the distinction between scoring and measurement will remain obscured. Curriculum inversion will therefore continue to reinforce the broader pattern of measurement inversion that characterizes contemporary HTA, limiting the scientific legitimacy of quantitative claims regarding therapy impact.

CONCLUSION

The German HTA research and training knowledge base faces a straightforward scientific test. It should ask whether every graduate entering HTA practice can answer the ten canonical curriculum questions that define the scientific foundations of quantitative measurement. Can they define the target attribute? Can they explain why every therapy assessment must begin with specification of that attribute? Can they distinguish nominal, ordinal, interval, and ratio scales and explain the mathematical operations each supports? Can they explain why the measurement status of an attribute must be established before any quantitative claim can be advanced? Can they describe the role of representational measurement? Can they explain why unidimensionality is a prerequisite for measurement? Can they distinguish manifest from latent attributes? Can they explain why latent attributes require a formal measurement model? Can they distinguish linear ratio measurement from Rasch logit ratio measurement? Can they explain why every therapy-impact claim must be prospectively evaluable, independently replicable, and capable of empirical falsification?

If the answer to these questions is no, the problem extends far beyond pedagogy. It becomes a professional and scientific issue. German HTA education would be certifying competence in evidence synthesis, comparative benefit assessment, patient-relevant outcomes, economic evaluation, and established HTA procedures without certifying competence in the scientific principles required to determine whether those procedures support lawful quantitative claims. This is the defining characteristic of curriculum inversion. Students become proficient in applying accepted methods without first acquiring the measurement framework needed to evaluate their scientific legitimacy.

The conclusion is therefore direct. The German HTA research and training knowledge base demonstrates considerable strength in clinical evidence appraisal, comparative benefit assessment, evidence-based medicine, health economic evaluation, and institutional decision-making. The present interrogation also indicates exceptionally strong recognition of directly observable, manifest attributes. At the same time, it reveals a profound weakness in the scientific foundations of measurement. Representational measurement, the distinction between manifest and latent attributes, Rasch measurement, admissible arithmetic, and the conditions required for lawful quantitative inference occupy only a marginal position within the curriculum. This finding does not diminish the methodological sophistication of German HTA. Rather, it identifies the scientific foundation that must be restored if HTA is to claim standing as a quantitative scientific discipline.

Scientific HTA must begin with measurement rather than arithmetic; with explicit specification of the target attribute rather than acceptance of existing instruments; with representational measurement rather than numerical convention; with lawful linear ratio and Rasch logit ratio measurement rather than composite scores and utilities; and with prospectively evaluable, independently replicable, and empirically falsifiable therapy-impact claims rather than continued reliance upon methodological inheritance. Only then can German HTA fully reconcile its internationally respected tradition of rigorous evidence assessment with the equally rigorous standards required for quantitative measurement.

To facilitate this transition, Maimon Research LLC has developed a comprehensive nine-unit HTA Reconstruction Program [6]. 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] Germany: Academic Research Centers and the endorsement of measurement failure in health technology assessment. Logit Working Paper No 106. https://maimonresearch.com/logit-working-paper-no-106/

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

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

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

[5] European Union: The EU institutionalizes false measurement as the basis for member state health technology assessment through EUnetHTA and EU HTA. Logit Working Paper No 239. https://maimonresearch.com/logit-working-paper-no-239-february-2026-rev-8/

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