MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLOGY ASSESSMENT

THE END OF THE REFERENCE CASE: RECONSTRUCTING HTA

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

LOGIT WORKING PAPER No 345 JUNE 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

For more than four decades the reference case has occupied a central position within health technology assessment (HTA). Developed initially in the United Kingdom and subsequently adopted by reimbursement agencies, academic research centres and professional organizations worldwide, the reference case promised a standardized framework for evaluating competing therapies. Consistency, comparability, transparency and cost-effectiveness became its defining characteristics. Utilities, QALYs and simulation models emerged as the principal tools for informing resource allocation decisions. Yet despite its widespread acceptance, the reference case has rarely been examined from the perspective of measurement theory.

This paper makes clear that the central failure of the reference case is its institutionalization of measurement inversion: the reversal of the scientific sequence in which arithmetic precedes measurement. Drawing upon large language model interrogations of more than 230 HTA-related knowledge bases, including agencies, academic centers, journals, professional organizations and Colleges and Schools of Pharmacy, the analysis demonstrates a consistent pattern. Statements affirming representational measurement, ratio measurement and the principle that measurement must precede arithmetic receive weak endorsement, while propositions supporting utilities, QALYs and simulation-based decision making receive strong endorsement. Measurement inversion emerges as a defining fatal characteristic of the contemporary HTA paradigm.

The consequences are profound. The paper traces the progression from health-state descriptions and ordinal preference scores through utilities, QALYs and simulation models to cost-effectiveness claims. At each stage arithmetic operations are undertaken without first establishing the measurement properties of the entities involved. The result is arithmetic chaos: a sequence of inadmissible transformations in which each stage inherits the deficiencies of the previous one. The final outputs may be numerically sophisticated, but they remain numerical constructs rather than lawful measures of therapy impact.

The analysis concludes that the reference case has reached closure. Closure does not imply the disappearance of agencies or the abandonment of modelling. Rather, it signifies the exhaustion of the framework's scientific legitimacy and its inability to justify its claims according to the standards of quantitative science. Reconstruction therefore becomes essential to undo 40 years of analytical failure. The only framework for attribute claims is founded on two lawful forms of measurement: linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes. Together these provide the basis for explicit attribute claims, empirical evaluation, falsification and the continuing evolution of objective knowledge in HTA.

INTRODUCTION

The adoption of the reference case as the foundation of health technology assessment (HTA) represents one of the most influential developments in modern health policy. Originating in the United Kingdom and subsequently embraced by reimbursement agencies, academic research centers and professional organizations throughout the world, the reference case promised a

standardized framework for evaluating competing therapies. Consistency, comparability and transparency became its defining characteristics. Utilities, QALYs, simulation models and cost-effectiveness thresholds were presented as essential tools for rational decision making. Over time, the reference case acquired the status of orthodoxy, shaping both policy and professional education.

Unlike the United States, where multiple approaches coexist, in the UK the reference case occupies a central position within the National Institute for Health and Care Excellence (NICE) [i] . Its origins can be traced to the 1980s and early 1990s through the work of the York Centre for Health Economics and the Office of Health Economics, which sought to establish a common framework for economic evaluation in reimbursement decisions [ii iii]. Today, NICE guidelines continue to place cost-utility analysis, QALYs, simulation modelling and comparative cost-effectiveness claims at the center of evidence assessment, endorsing measurement inversion .

What is striking, however, is that while the reference case standardizes arithmetic, it pays little attention to the prior question of measurement. Recent large language model interrogations of HTA agencies, academic centers, journals and educational institutions have revealed a remarkably consistent pattern. Statements supporting representational measurement, ratio measurement and the principle that measurement must precede arithmetic receive weak endorsement, while propositions supporting utilities, QALYs and reference-case modelling receive strong endorsement. The result is measurement inversion: arithmetic routinely precedes measurement.

This finding is important because measurement inversion is not merely another methodological criticism. It is a diagnostic signal that the analytical framework has become detached from the constraints imposed by scales of measurement and the axioms of representational measurement. Once arithmetic is no longer governed by lawful measurement, arithmetic chaos follows. Numerical operations accumulate through utilities, QALYs and simulation models without establishing whether the underlying entities possess the properties necessary to support those operations.

The implications are profound. For more than four decades the reference case has generated claims concerning cost-effectiveness, value for money and resource allocation. Yet if the measurement foundations of the framework are absent, these claims cannot be rescued by larger datasets, more elaborate sensitivity analyses or increasingly sophisticated simulation techniques. The problem is not technical but structural. The reference case was designed to standardize arithmetic rather than establish measurement. As a consequence, measurement inversion became institutionalized, arithmetic chaos became inevitable and closure followed. Closure does not mean that agencies disappear or models cease to be constructed. It means that the framework no longer possesses the capacity to justify its claims according to the standards of quantitative science. The future of HTA therefore lies not in reforming the reference case but in replacing it with a framework grounded in lawful measurement, explicit attribute claims and empirical evaluation.

JUSTIFYING THE REFERENCE CASE

The success of the reference case rests on a set of assumptions that are rarely challenged. It is presented as a standardized framework for evaluating therapies, promoting consistency across

assessments, transparency in decision making and comparability between competing interventions. These objectives have become so widely accepted that they are often treated as self-evident virtues and, for many, define modern HTA itself.

The attraction is obvious. Health systems face continual resource allocation decisions, and a framework that applies common assumptions, outcome measures and decision rules appears to offer a rational basis for comparing therapies. If every intervention is assessed according to the same procedures, the resulting recommendations are assumed to possess scientific legitimacy. The reference case therefore presents itself not merely as a methodology but as a guarantor of objectivity.

The difficulty is that consistency is not the same as validity. A framework may apply the same calculations repeatedly and transparently while failing to establish that the entities entering those calculations possess lawful measurement properties. Standardization can ensure that the same arithmetic is applied in every assessment; it cannot establish that the arithmetic itself is meaningful.

This distinction is crucial because the justifications offered for the reference case focus almost entirely on the administration of arithmetic. Standardization, comparability, transparency and reproducibility are administrative virtues. They say nothing about whether utilities, QALYs and model parameters satisfy the requirements imposed by scales of measurement and the axioms of representational measurement.

The omission is remarkable. The reference case was developed to support quantitative claims regarding therapy impact, yet measurement occupies no central role in its justification. There is no requirement that outcomes be ratio measures, no distinction between manifest and latent attributes and little consideration of whether the arithmetic operations employed are admissible. Instead, the framework assumes that utilities, QALYs and model outputs are suitable for arithmetic manipulation and focuses on ensuring that those manipulations are applied consistently.

The consequence is a confusion between consistency of arithmetic and validity of measurement. A standardized framework built upon inadmissible arithmetic does not become scientifically legitimate because it is applied consistently. It merely reproduces the same errors more efficiently. The critical question is therefore not whether the reference case achieves standardization, but whether it can support meaningful claims regarding therapy impact. That question leads directly to the first principle of quantitative science: measurement must precede arithmetic. It is this requirement that the reference case fails to address and which ultimately explains the emergence of measurement inversion, arithmetic chaos and closure.

MEASUREMENT BEFORE ARITHMETIC

The principle that measurement must precede arithmetic is neither controversial nor new. It is one of the foundational requirements of quantitative science. Before numbers can be added, multiplied, averaged or incorporated into mathematical models, it must first be demonstrated that the entities being manipulated possess the measurement properties necessary to support those operations. Arithmetic is therefore conditional upon measurement. Without measurement, arithmetic becomes little more than numerical manipulation without established scientific meaning.

Although this principle has long been embedded in the physical sciences, its formal articulation within the social sciences emerged during the twentieth century. Stevens' classification of measurement scales in 1946 distinguished nominal, ordinal, interval and ratio scales and specified the arithmetic operations appropriate to each [iv]. This was followed by the formalization of representational measurement theory through the work of Krantz, Luce, Suppes and Tversky in 1971 [v]. Together, these developments established a simple but profound principle: not all numbers are measures and not all measures support the same arithmetic operations.

At the same time, Rasch measurement is entirely absent from the reference-case framework. First proposed by Georg Rasch in 1960 and subsequently linked by Benjamin Wright in 1977 to the requirements of fundamental and representational measurement, Rasch provides a rigorous methodology for the development of instruments intended to measure latent attributes [vi] [vii]. Indeed, Rasch measurement remains the only recognized approach that offers the necessary and sufficient conditions for transforming subjective observations into lawful ratio measures of latent trait possession.

This omission is particularly striking because latent attributes are implicit throughout the reference-case framework. Time trade-off exercises, health-state descriptions, utility generation and patient-reported outcomes all assume the existence of latent characteristics such as symptom burden, treatment satisfaction, functional status and quality of life. Yet rather than measuring these attributes directly, the reference case relies upon ordinal preference scores and composite utility constructions that are subsequently treated as though they were quantitative measures.

The consequence is that the central Rasch insight is ignored: measurement begins with the attribute itself. If an attribute is latent, then a Rasch instrument must be developed to establish whether observations can support a lawful measurement structure. Instead, the reference case bypasses this requirement entirely, assuming that ordinal responses can be transformed into utilities and then manipulated arithmetically to create QALYs and cost-effectiveness claims.

Equally important, the reference case fails to recognize the fundamental distinction between manifest and latent attributes. Manifest attributes, such as hospital admissions, hospital days and prescription counts, possess observable ratio properties. Latent attributes require Rasch measurement. Health-state descriptions ignore this distinction by bundling together multiple manifest and latent characteristics within a single composite structure.

The implications are immediate. Ordinal scales support ranking but nothing more. They indicate relative position but provide no basis for multiplication, division or meaningful estimates of magnitude. Ratio scales, by contrast, possess a meaningful non-arbitrary zero and constant units, supporting the full range of arithmetic operations. Arithmetic is therefore constrained by the properties of the scale. It cannot be undertaken independently of how the numbers were generated.

For therapy assessment, the requirements are equally clear. Claims regarding therapy impact must begin with measurable attributes. Manifest attributes, such as hospital admissions, physician visits, emergency department utilization and hospital days, require linear ratio measures. Latent attributes, such as symptom burden, quality of life, treatment satisfaction and need fulfilment,

require Rasch logit ratio measures. These are the only lawful routes to quantitative claims regarding therapy impact.

This omission is not a minor technical oversight. It changes the entire character of the framework. Once measurement ceases to govern arithmetic, there is no principled basis for determining which numerical operations are admissible. Numbers become detached from their measurement properties and can be manipulated without constraint. This reversal of the scientific sequence—arithmetic before measurement—is the defining characteristic of measurement inversion and the starting point for the arithmetic chaos that follows.

THE MEASUREMENT INVERSION SIGNAL

The emergence of large language models has created, for the first time, the opportunity to undertake systematic interrogation of institutional knowledge bases. Rather than focusing on the views of individual researchers or isolated publications, large language models allow assessment of the concepts, assumptions and methodological commitments embedded within entire disciplines. By presenting a common set of canonical statements derived from the scales of measurement and the axioms of representational measurement, it becomes possible to evaluate whether a knowledge system recognizes the conditions necessary for lawful quantitative claims.

Applied to health technology assessment, the results are striking. Interrogations of more than 230 HTA-related knowledge bases, including reimbursement agencies, academic research centres, journals, professional organizations and Colleges and Schools of Pharmacy, reveal a remarkably consistent pattern. Statements affirming representational measurement, ratio measurement, unidimensionality and the principle that measurement must precede arithmetic receive consistently weak endorsement. In contrast, propositions supporting utilities, QALYs, reference-case simulation models and cost-effectiveness claims receive consistently strong endorsement. The result is a reproducible profile of measurement inversion across countries, institutions and professional communities.

These findings are important because they provide a diagnostic evaluation of the HTA paradigm itself. The issue is no longer whether individual models are flawed or whether particular assumptions should be revised. The interrogations point to a deeper problem: a systematic failure to recognize the role of measurement as the foundation of quantitative inquiry. Measurement inversion emerges as the defining yet fatal characteristic of the contemporary HTA knowledge base.

Measurement inversion is therefore not merely another methodological criticism. It is a diagnostic signal. It indicates that the analytical framework has become detached from the constraints imposed by scales of measurement and the axioms of representational measurement. Once this occurs, the distinction between admissible and inadmissible arithmetic operations begins to disappear. If measurement inversion is the signal; arithmetic chaos is the consequence. Once measurement no longer governs arithmetic, there is no scientific basis for deciding which numerical operations are permissible. In the reference case, numbers are treated as measures simply because they are expressed numerically. The result is arithmetic chaos and eventual closure.

MEASUREMENT INVERSION, ARITHMETIC CHAOS AND THE REFERENCE CASE

If measurement inversion is the signal, arithmetic chaos is the consequence. The reference-case framework provides perhaps the clearest example of how this process unfolds. At each stage of model construction, arithmetic operations are undertaken without first establishing whether the entities involved satisfy the requirements of representational measurement. The result is a sequence of numerical manipulations in which each stage inherits the deficiencies of the stage before it. What begins as a failure to establish measurement ultimately culminates in elaborate numerical constructions that are presented as evidence of therapy value.

The process begins with health-state descriptions. These are intended to capture aspects of health, functioning and well-being that may be affected by disease or treatment. Yet health-state descriptions are not measures. They are collections of characteristics and observations, often combining multiple dimensions of health within a single descriptive framework. They may provide useful information regarding patient experience, but they do not possess established measurement properties. Nevertheless, they form the basis for preference elicitation exercises such as time trade-off and standard gamble techniques.

These exercises generate preference rankings and valuations. The resulting scores are ordinal. They indicate relative preference among alternatives but do not establish equal intervals, constant units or a meaningful zero. Ordinal structures support ranking and nothing more. They do not support multiplication, division or the construction of quantitative relationships. At this point the analytical sequence should stop because the conditions necessary for further arithmetic have not been satisfied. Instead, the reference-case framework proceeds to the next stage.

Preference scores are transformed into utility values intended to represent the relative desirability of different health states. Yet this transformation changes nothing fundamental. Utilities remain composite constructs derived from ordinal preferences. They are neither unidimensional measures nor entities with demonstrated ratio properties. Indeed, many utility systems permit negative values for health states judged to be worse than death. This feature alone is inconsistent with the requirements of ratio measurement, which requires a meaningful non-arbitrary zero and does not admit negative values. Utilities therefore remain numerical constructs rather than lawful measures.

The next step introduces the QALY. Utilities are multiplied by time in an attempt to create a common metric capable of comparing therapies across disease areas. This is the point at which arithmetic chaos becomes explicit. Time is a lawful ratio measure, but multiplication requires that both entities entering the operation possess ratio properties. Utilities do not satisfy this requirement. Multiplying a composite ordinal construct by time cannot create a lawful measure. The resulting QALY therefore inherits all of the measurement deficiencies present in the utility structure from which it is derived.

The process then moves to simulation modelling. QALYs, costs, transition probabilities, compliance assumptions, persistence estimates and numerous other inputs are incorporated into analytical models that may project outcomes over decades or even entire lifetimes. Many of these inputs have no established measurement status. Costs frequently combine heterogeneous resource

units into composite monetary aggregates. Assumptions are treated as though they were measurable quantities. Transition probabilities are combined with utility values and projected treatment effects. Increasing model complexity creates an appearance of analytical sophistication while simultaneously obscuring the absence of lawful measurement foundations.

The final output is typically expressed as a cost-effectiveness ratio, most commonly cost per QALY gained. By this stage the framework has travelled a considerable distance from measurable attributes. The ratio combines entities whose measurement properties remain unresolved and presents the result as evidence for pricing and reimbursement decisions. The apparent precision of the estimate conceals the fact that it rests upon a succession of inadmissible arithmetic operations. Every stage of the analytical sequence depends upon entities that fail to satisfy the requirements necessary for quantitative interpretation.

This is arithmetic chaos. Once arithmetic becomes detached from measurement, there is no longer a principled basis for distinguishing admissible from inadmissible operations. Any numerical entity can be manipulated by any other numerical entity simply because both are expressed as numbers. Utilities can be multiplied by time. Composite costs can be divided by composite outcomes. Assumptions can be treated as quantities. The discipline that measurement imposes upon quantitative inquiry disappears.

The importance of this conclusion should not be underestimated. Arithmetic chaos is not a criticism of any single component of the reference case. It is a criticism of the entire sequence. Health-state descriptions, ordinal preferences, utilities, QALYs and simulation outputs are linked together through a chain of calculations that ignores the constraints imposed by scales of measurement and the axioms of representational measurement. Measurement inversion is the signal; arithmetic chaos is the consequence. The remarkable feature is not that arithmetic chaos emerged, but that this process remained largely unchallenged for more than four decades while becoming the dominant framework for therapy assessment globally.

THE REFERENCE CASE: NOT FIT FOR PURPOSE

The purpose of the reference case is straightforward. It is intended to support claims regarding the comparative value of therapies and assist decision makers in allocating health care resources. To achieve this objective, the framework must be capable of generating meaningful quantitative claims regarding therapy impact. The question is therefore not whether the reference case is widely used, internally consistent or administratively convenient. The question is whether it is fit for purpose.

The answer follows directly from the preceding analysis. A framework designed to evaluate therapy impact must begin with measurement. Therapy claims are claims about attributes. A new therapy is expected to reduce symptoms, improve functioning, enhance quality of life, reduce hospital admissions or influence some other manifest or latent attribute. If these claims are to support quantitative inference, the attributes concerned must first be measured according to recognized scientific standards.

The reference case does not satisfy this requirement. Instead, it begins with valuation. Health-state descriptions are converted into preference scores, preference scores into utilities, utilities into QALYs and QALYs into simulation outputs. At no stage is there a demonstration that the entities entering these calculations possess the measurement properties necessary to support the arithmetic operations imposed upon them. The framework therefore substitutes numerical manipulation for measurement.

This is not a minor methodological weakness. It is a structural failure. The reference case was designed to standardize arithmetic rather than establish lawful measurement. Consistency, comparability and transparency are repeatedly presented as virtues of the framework, yet these characteristics are irrelevant if the underlying quantities are not measures. A calculation does not become scientifically meaningful because it is applied consistently. Standardization merely ensures that the same errors are reproduced across all assessments.

The consequence is that the reference case cannot achieve its stated objective. A framework built upon measurement inversion and arithmetic chaos cannot generate credible claims regarding therapy impact. It may generate numbers. It may generate ratios. It may generate recommendations. What it cannot generate is evidence grounded in lawful measurement.

The problem is often misunderstood as one of model refinement. More sophisticated utilities, larger simulation models, additional sensitivity analyses and increasingly complex assumptions are presented as solutions. They are not. Such modifications leave the underlying problem untouched because the difficulty lies not with particular assumptions but with the architecture itself. Arithmetic chaos is embedded in the structure of the framework.

The implication is unavoidable. The reference case is not a flawed framework awaiting improvement. It is a dead end. The framework cannot be rehabilitated because its foundations are incompatible with the requirements of representational measurement. For more than four decades HTA has attempted to construct quantitative claims from entities that do not possess the properties necessary to support those claims. The result is a closed analytical structure that appears scientific while lacking the conditions required for scientific legitimacy.

CLOSURE

The concept of closure has a precise meaning. It does not imply that agencies cease to function, journals cease publication or simulation models cease to be constructed. Nor does it imply that practitioners abandon the reference case overnight. Closure refers to something more fundamental: the exhaustion of a framework's scientific legitimacy.

A framework reaches closure when it no longer possesses the capacity to justify its claims according to accepted scientific standards. In the physical and biological sciences, progress depends upon continual exposure of claims to empirical challenge. Measurement structures are examined, assumptions are tested and theories remain vulnerable to falsification. Closure occurs when this process breaks down and the framework becomes incapable of recognizing or correcting its own foundational errors.

The reference case has reached precisely this point. For more than four decades the framework has focused on refining utilities, expanding simulation models, increasing computational sophistication and developing increasingly elaborate sensitivity analyses. Yet none of these developments address the prior question of measurement. The framework evaluates uncertainty within the model while ignoring the validity of the quantities entering the model. It refines arithmetic while neglecting measurement.

This is why closure follows inevitably from measurement inversion. Once arithmetic is allowed to proceed independently of lawful measurement, the framework loses the ability to determine whether its outputs possess quantitative meaning. Numerical complexity becomes a substitute for scientific legitimacy. Larger models create an appearance of progress while leaving the foundational problem untouched.

The significance of closure is that it changes the nature of the debate. The issue is no longer whether the reference case can be improved. The issue is whether the framework possesses the scientific foundations necessary to support improvement. The evidence presented here suggests that it does not. Measurement inversion is not a technical defect that can be corrected through incremental reform. It is a structural characteristic of the framework itself.

Closure therefore marks the end of the reference case as a framework for therapy assessment. This conclusion does not arise from disagreement over assumptions, discount rates, utility algorithms or model specifications. It arises because the framework abandoned the principle that measurement must precede arithmetic. Once that principle was abandoned, arithmetic chaos became inevitable. Recognizing closure was merely a matter of time.

The recognition of closure is not a pessimistic conclusion. On the contrary, it creates the opportunity for transition. The end of the reference case is not the end of HTA. It is the end of a framework that substituted valuation for measurement and simulation for empirical evaluation. What follows is the task that should have been undertaken from the beginning: the development of therapy assessment grounded in lawful measurement, explicit claims and falsification.

RECONSTRUCTING HTA: THE TWO MEASURES

The closure of the reference case does not create a vacuum in health technology assessment. On the contrary, it creates an opportunity to reconstruct HTA on foundations consistent with the requirements of quantitative science. The abandonment of utilities, QALYs, simulation models and cost-per-QALY claims does not leave decision makers without evidence. Rather, it removes a framework built upon measurement inversion and arithmetic chaos and replaces it with one grounded in lawful measurement, empirical evaluation and falsification.

The key shift is from simulated claims to evaluable claims. For more than four decades HTA has focused on constructing elaborate models intended to predict future outcomes. Increasing attention has been given to refining assumptions, improving model structures and conducting sensitivity analyses. Yet these activities have done little to address the central question of whether the entities entering the framework possess the measurement properties necessary to support quantitative

claims. Reconstruction begins by rejecting this approach and returning to the first principle of quantitative science: measurement must precede arithmetic.

Once the scales of measurement and the axioms of representational measurement are recognized, the path forward becomes straightforward. Therapy impact assessment must begin with attributes. These attributes fall into two distinct classes: manifest attributes and latent attributes. Each requires its own measurement framework and each must be evaluated separately. The long-standing ambition of creating a single universal outcome metric capable of combining all aspects of therapy impact must therefore be abandoned.

Manifest attributes are directly observable. Examples include hospital admissions, emergency department visits, physician consultations, hospital days, treatment persistence and therapy switching. These attributes require linear ratio measures. Because they possess meaningful units and a non-arbitrary zero, they support lawful arithmetic operations and provide the basis for credible, evaluable and replicable claims regarding therapy impact.

Latent attributes are different. They are not directly observable but are inferred from structured observations. Examples include symptom burden, fatigue, treatment satisfaction, need fulfilment and quality of life. These attributes require Rasch measurement. Through the construction of a Rasch logit ratio scale, subjective observations are transformed into lawful measures of latent attribute possession. Therapy impact is then assessed through changes in possession of the latent trait rather than through manipulation of ordinal scores or composite indices.

The result is a new HTA built around only two forms of measurement: the linear ratio measure for manifest attributes and the Rasch logit ratio measure for latent attributes. Together they provide the foundation for explicit attribute claims, protocol-driven evaluation, empirical assessment and falsification. Instead of generating hypothetical claims through simulation, HTA becomes a research program directed toward the accumulation of objective knowledge. The future of HTA therefore lies not in numerical storytelling but in measurable attributes, evaluable claims and the continuing empirical assessment of therapy impact.

CONCLUSION

The reconstruction of health technology assessment does not require the invention of new principles. It requires a return to principles that quantitative science has long recognized. The central lesson of this review is that the reference case failed because it standardized arithmetic while neglecting measurement. The consequence was measurement inversion, the emergence of arithmetic chaos and ultimately the closure of a framework that could no longer justify its claims according to the standards of representational measurement.

Closure, however, should not be viewed as a crisis. It should be viewed as an opportunity. The abandonment of utilities, QALYs, simulation-based cost-effectiveness claims and reference-case modelling does not leave HTA without direction. Rather, it clears the ground for a framework capable of supporting meaningful quantitative claims regarding therapy impact. The challenge is not to rescue the reference case but to replace it.

The path forward is straightforward. Therapy impact assessment must recognize the distinction between manifest and latent attributes and the different measurement structures required for each. Manifest attributes require linear ratio measures. Latent attributes require Rasch logit ratio measures. These are the only lawful routes to quantitative claims once the scales of measurement and the axioms of representational measurement are respected.

The result is a new HTA centered on measurable attributes, explicit claims, empirical evaluation and falsification. Attention shifts from imaginary futures generated by simulation models to observable evidence generated through protocol-driven assessment. In place of numerical storytelling emerges a cumulative research program dedicated to the evolution of objective knowledge. The future of HTA therefore lies not in refining the reference case, but in constructing a discipline founded on lawful measurement and the scientific evaluation of therapy impact.

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

[i] Langley P. United Kingdom: The National Institute for Health and Care Excellence: The absence of representational measurement. Logit Working Paper No 701 https://maimonresearch.com/logit-working-paper-no-701-march-2026-rev-8/

[ii] Langley P. United Kingdom: The York Centre for Health Economics – A canonical knowledge base interrogation and the absence of measurement. Logit Working Paper No 780 https://maimonresearch.com/logit-working-paper-no-780-april-2026/

[iii] Langley P, United Kingdom: The Office of Health Economics – The inversion of representational measurement. Logit Working Paper No 25, https://maimonresearch.com/logit-working-paper-no-25/

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

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

[vi] Rasch G, Probabilistic Models for some Intelligence and Attainment Tests. Chicago: University of Chicago Press, 1980 [An edited version of the original 1960 publication]

[vii] Wright B. Solving measurement problems with the Rasch Model. *J Educational Measurement.* 1977;14(2):97-116