MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

AUSTRALIA: PARADIGM FAILURE IN HEALTH TECHNOLOGY ASSESSMENT

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

LOGIT WORKING PAPER No 544 JUNE 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

This paper argues that Australian health technology assessment (HTA) has reached a point of paradigm failure. The claim is not that the contemporary reference-case framework requires refinement, improved modelling techniques, or revised utility instruments. Rather, the argument is that the framework itself was established without satisfying the standards required for quantitative science. The central constructs of the reference case—utility scores, quality-adjusted life years (QALYs), cost-effectiveness ratios, and simulation models—depend upon measurement properties that have never been demonstrated. Utilities are treated as though they possess ratio properties, QALYs are constructed through inadmissible multiplication, and simulation models extend these assumptions across hypothetical populations and time horizons. The result is a framework grounded in assumptions regarding measurement rather than measurement itself.

The paper presents evidence from a series of large language model interrogations of the Pharmaceutical Benefits Advisory Committee (PBAC) and seven leading Australian HTA research centers. These interrogations assessed endorsement of canonical propositions derived from representational measurement and scientific measurement theory. The findings reveal a consistent pattern of measurement inversion. Concepts fundamental to quantitative science—representational measurement, ratio measurement, dimensional homogeneity, unidimensionality, manifest and latent attributes, and Rasch measurement—receive little recognition, while propositions that conflict with these principles receive strong endorsement. The same investigations also reveal curriculum inversion, where students and researchers are introduced to utilities, QALYs, cost-effectiveness analysis, and simulation modelling without first being equipped to evaluate their measurement foundations.

The paper argues that these findings explain the persistence of the reference-case paradigm despite its failure to satisfy the requirements of measurement. The concepts necessary to identify the defect have themselves been largely absent from the Australian HTA knowledge base. The consequence is a self-reinforcing intellectual system in which arithmetic precedes measurement and numerical constructions are mistaken for evidence. The only viable response is reconstruction. A future HTA framework must begin with attributes, distinguish manifest from latent outcomes, employ linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes, and require all claims regarding therapy impact to be evaluable, replicable, and falsifiable. The future of HTA depends not upon reforming the reference case but upon replacing it with a framework grounded in measurement science.

INTRODUCTION

Paradigm failure has now been recognized in Australian health technology assessment (HTA). The evidence is clear. Interrogations of the Pharmaceutical Benefits Advisory Committee (PBAC) and seven leading Australian HTA research centers demonstrate a consistent failure to recognize the requirements of representational measurement, admissible arithmetic and scientific measurement theory. The result is a discipline that has become detached from the standards necessary to support quantitative claims regarding therapy impact.

Australia has long regarded itself as occupying a leading position in international HTA. Through PBAC, academic research centers, professional organizations and university training programs, it has played an influential role in promoting the reference-case framework as the preferred approach to therapy assessment. Utility scores, quality-adjusted life years (QALYs), cost-effectiveness ratios and simulation models have become established not only as accepted analytical tools but as the defining features of modern HTA. The widespread belief has been that the reference case represents the culmination of methodological progress in healthcare decision making.

Nothing could be further from the truth.

Judged against the standards of representational measurement, the reference case is an analytically impossible construct. Its central components depend upon arithmetic operations that require measurement properties which have never been demonstrated. Utility scores are treated as though they possess ratio properties. QALYs are constructed through multiplication without establishing that multiplication is admissible. Reference-case simulation models extend these assumptions across hypothetical populations and lifetime horizons. The resulting framework is not grounded in measurement but in assumptions regarding measurement.

The significance of this conclusion is often misunderstood. The problem is not that the reference case was once scientifically defensible and has gradually become less convincing. Nor is the problem that methodological refinement has failed to keep pace with advances in science. The defect was present from the outset. From the moment health-state valuations were transformed into utility scores and treated as though they represented measurable quantities, the framework departed from the requirements of measurement. The subsequent development of QALYs, cost-effectiveness ratios and simulation models merely extended and institutionalized this original error.

The reference case therefore did not begin with measurement and later lose its way. It began without measurement. The train never left the station.

This raises an obvious question. If the defect was present from the beginning, why did it remain largely invisible for more than four decades? Why was it accepted by policy makers, researchers, educators and professional organizations? Why was it embraced by institutions whose responsibility was to uphold scientific standards?

The answer emerges from the interrogation program undertaken in Australia and internationally. The importance of the measurement inversion and curriculum inversion interrogations is that they provide systematic evidence explaining why the problem persisted. The concepts required to identify the defect are themselves largely absent from the HTA knowledge base.

Across PBAC and the Australian research centers there is little recognition of representational measurement, admissible arithmetic, ratio measurement, unidimensionality, the distinction between manifest and latent attributes or the role of Rasch measurement in constructing measures of latent attribute possession. The result is a professional culture in which arithmetic is accepted while measurement is assumed. The analytical impossibility at the heart of the reference case remained invisible because the intellectual tools required to identify it were largely absent. This is unacceptable,

This is the significance of the present findings. For the first time, the interrogations provide consistent and reproducible evidence that the persistence of the reference case is not evidence of scientific validity. Rather, it reflects a curriculum and research environment that has institutionalized measurement inversion and reproduced it through policy, teaching and research for more than forty years.

The findings are particularly striking because they concern institutions that should have known better. PBAC occupies a central role in Australian healthcare decision making. Australian HTA research centers are responsible for training future researchers, analysts and policy advisers. Yet the interrogations reveal the same omissions across all institutions. The same assumptions are reproduced. The same absence of measurement science recurs. The same failure to confront the requirements of measurement is evident regardless of institution or setting.

The challenge facing Australian HTA is not impossible reform but reconstruction. The reference case offers no basis for reform because the source of failure lies in its design. Utilities, QALYs and reference-case simulation models are not imperfect implementations of sound measurement principles. They are failed constructs that never considered the requirements of measurement.

Reconstruction therefore becomes imperative. It cannot be avoided if HTA is to retain any credibility. HTA must be rebuilt upon a foundation that recognizes measurement as the prerequisite for quantitative claims, distinguishes manifest from latent attributes, employs valid ratio measures where required and utilizes Rasch measurement for latent attributes. Claims regarding therapy impact must be evaluable, replicable and capable of falsification. Evidence must be grounded in measurement rather than numerical construction.

The conclusion is stark but unavoidable. The reference case is a 40-year design failure. After more than four decades, the absence of measurement can no longer be treated as an oversight, a methodological limitation or a topic for future discussion. The evidence demonstrates that the paradigm itself is incapable of satisfying the standards required for quantitative science.

AUSTRALIA: DEFINING PARADIGM FAILURE

It could reasonably be argued that, given the built-in design failure of the reference-case paradigm, talk of paradigm failure is some forty years overdue. The foundational defect was always present. Once health-state descriptions were converted into utility scores and treated as though they possessed the properties required for quantitative measurement, the conditions for measurement inversion were established. From this perspective, the subsequent development of QALYs, cost-effectiveness ratios, and simulation models merely extended and institutionalized an error that was already embedded within the framework.

For many practitioners, researchers, and policy-makers in Australia, however, the claim of paradigm failure remained largely unsubstantiated. The underlying design fault had never been examined systematically, and the concepts required to identify it were largely absent from the HTA knowledge base. As a result, the persistence of the reference case was often interpreted as evidence of scientific legitimacy rather than evidence of institutional acceptance.

The emergence of large language model interrogation has changed that position. For the first time, it has become possible to examine HTA knowledge bases systematically and consistently against the standards of representational measurement. The resulting interrogations provide substantive evidence not only of measurement inversion but also of curriculum inversion. They demonstrate that the concepts required to recognize measurement failure are themselves largely absent from the educational, research, and policy environments that support HTA. Paradigm failure is therefore no longer a theoretical assertion. It is supported by reproducible evidence showing that the contemporary HTA framework both embodies and reproduces assumptions that conflict with the requirements of quantitative science.

A paradigm fails when the assumptions that support it can no longer be reconciled with the standards that govern scientific inquiry. The failure may not be immediately visible. Institutions may continue to employ the paradigm, journals may continue to publish within it, and educational programs may continue to teach it. Yet once its foundational assumptions are shown to be inconsistent with accepted scientific standards, the paradigm loses its intellectual legitimacy. It was always a design failure,

Paradigm failure is therefore not a matter of disagreement over methods or competing analytical preferences. It occurs when the conceptual foundations of a framework are shown to be defective. The framework may continue to function administratively, but it can no longer claim scientific authority.

In the physical sciences, paradigm failure occurs when observations repeatedly conflict with a theoretical structure and the resulting anomalies cannot be accommodated within the existing framework. In measurement science, paradigm failure occurs when the quantities upon which a framework depends are shown not to possess the properties required for the operations being performed. The problem is not that the calculations are inaccurate. The problem is that the calculations are inadmissible.

Applied to health technology assessment, paradigm failure occurs when the central constructs of the field are shown to violate the requirements of measurement scales and representational measurement [1] [2] [3] . Utilities are treated as though they possess ratio properties without demonstration. QALYs are constructed through multiplication without establishing that multiplication is admissible. Reference-case simulation models extend these assumptions across hypothetical lifetimes and populations. If the measurement requirements necessary to support these operations are absent, then the framework loses its scientific foundation.

The relevance to Australia is immediate. The PBAC employs a reference-case framework in which utility-based evaluations and QALY-derived assessments play a central role in reimbursement decisions [4,5]. The practice interrogations of the major research centers indicate that the educational and research environments supporting HTA in Australia similarly give little recognition to representational measurement, scale theory, unidimensionality, latent attribute measurement and ratio measurement [6,7,8,9,10,11,12,13,14] . The result is a self-reinforcing system in which the reference case is both taught and applied while the measurement principles necessary to evaluate its scientific legitimacy remain largely absent. The issue is therefore not confined to a particular method or institution. It extends across policy, research and professional education.

A critical point is that paradigm failure does not imply that the framework ceases to exist. Paradigms often survive long after their foundations have been undermined. Educational programs continue to teach them. Research centers continue to apply them. Agencies continue to use them in decision making. What changes is their intellectual status. They become institutional conventions rather than scientific frameworks; constructs of historical interest.

Eventually, however, a choice must be made. Either the standards of measurement science are abandoned in order to preserve the paradigm, or the paradigm is abandoned in order to preserve the standards of measurement science. In the case of HTA, the relevant standards are those of representational measurement, dimensional homogeneity, unidimensionality, ratio measurement and falsifiable claims. If these standards are accepted, then utilities, QALYs and reference-case simulation models cannot be defended.

Paradigm failure therefore marks the point at which reform becomes impossible. The deficiencies are no longer matters of refinement or methodological adjustment. The foundations themselves have failed. Reconstruction becomes necessary because the existing structure cannot be repaired without abandoning the assumptions upon which it was built.

For Australia, this conclusion applies equally to the reference-case framework employed by the PBAC and to the educational and research structures that continue to reproduce it. The issue is not whether the current paradigm can be improved; it cannot. The issue is whether a framework constructed on arithmetic before measurement can survive once measurement is restored to its central place in scientific inquiry. If the answer is no, then what is being observed is not methodological weakness but paradigm failure.

The implications are unavoidable. The future of HTA in Australia cannot be secured through incremental reform, revised guidelines or increasingly sophisticated simulation techniques. The problem lies deeper, in the assumptions upon which the reference-case framework itself depends. Once those assumptions are shown to be incompatible with the requirements of measurement, the only remaining option is reconstruction. The task is not to repair the paradigm but to replace it with a framework in which measurement precedes arithmetic, quantitative claims are evaluable and falsifiable, and evidence is grounded in measures rather than numerical constructions.

The central argument advanced here is that any future HTA paradigm must conform to the requirements of ratio measurement. Measurement must precede arithmetic, and quantitative claims must be grounded in measures that satisfy the standards of representational measurement. Once this principle is accepted, only two legitimate forms of measurement remain available for the assessment of therapy impact. Manifest attributes, which are directly observable, require linear ratio measures. Latent attributes, which are not directly observable and must be inferred from empirical observations, require Rasch logit ratio measures. *There is no third category and no alternative pathway*. The future of HTA therefore rests upon the recognition that all evaluable claims regarding therapy impact must ultimately be supported by one of these two forms of ratio measurement [15].

AUSTRALIA: PARADIGM FAILURE - THE EVIDENCE FOR MEASUREMENT INVERSION

The evidence for measurement inversion in Australia is both extensive and consistent [16]. It is not confined to a single institution, a particular methodological preference or an isolated policy framework. Rather, it is evident across the entire HTA environment. Interrogations undertaken at the national level, together with separate assessments of the PBAC and 7 university HTA research centers. The results reveal a common pattern: quantitative claims are accepted and employed without first demonstrating that the quantities involved satisfy the requirements of measurement.

As detailed in the individual interrogation studies, the extent of measurement inversion was assessed through responses to a set of 24 canonical statements drawn from the principles of representational measurement and the requirements for valid quantitative claims. These statements comprised both true and false propositions. The true propositions reflected accepted principles of measurement science, while the false propositions represented assumptions embedded within the contemporary reference-case paradigm that are inconsistent with those principles.

The logic of the assessment was straightforward. In a knowledge base aligned with the standards of measurement science, the true propositions should attract high levels of endorsement and the false propositions should attract low levels of endorsement. Such a pattern would indicate recognition of the requirements for measurement, admissible arithmetic and valid quantitative claims.

The interrogations revealed the opposite pattern, Statements that should have been strongly endorsed received consistently low endorsement probabilities, while statements that should have been rejected received consistently high endorsement probabilities. The result was not merely a lack of knowledge regarding measurement theory. It was evidence of reversal. The knowledge bases systematically favored propositions that conflicted with the principles of representational measurement while failing to endorse propositions that reflected those principles.

This reversal is the defining characteristic of measurement inversion. The issue is not simply that measurement concepts are absent. The issue is that the analytical framework actively supports assumptions that contradict the requirements of measurement science. The interrogations therefore provide evidence not only of omission but of endorsement of error. Knowledge bases associated with HTA agencies, research centers and educational programs repeatedly accepted propositions that treated arithmetic as independent of measurement, assumed ratio properties where none had been demonstrated and supported quantitative claims without first establishing the measurement status of the underlying quantities.

The significance of this finding is considerable. If the interrogations had merely shown low awareness of measurement principles, the conclusion might have been limited to deficiencies in curriculum coverage or professional training. Instead, the results indicate the presence of an intellectual framework that systematically privileges false measurement propositions over true ones. The problem is therefore not ignorance alone. It is the institutionalization of assumptions that are incompatible with the standards of quantitative science that has lasted for some 40 years.

The evidence for measurement inversion rests on this reversal. True propositions are rejected. False propositions are endorsed. The resulting pattern demonstrates that the reference-case paradigm is not simply disconnected from measurement science; it operates according to assumptions that are fundamentally opposed to it.

This finding is important because measurement inversion lies at the heart of the contemporary failed HTA paradigm. Measurement inversion occurs when arithmetic precedes measurement. Instead of establishing the measurement properties of a quantity and then determining which mathematical operations are admissible, HTA begins with arithmetic and assumes that measurement has already been achieved. Numbers are treated as measures because they are numerical. The scientific burden of demonstrating measurement is quietly abandoned.

The separate interrogation of the PBAC reached a similar conclusion. The PBAC's decision-making framework relies heavily upon utility-based evaluations and QALY-derived assessments. These constructs occupy a central role in determining comparative value and informing reimbursement decisions. Yet their use presupposes that utility scores possess ratio properties capable of supporting multiplication by time. No demonstration of these properties is provided. The legitimacy of the arithmetic is assumed rather than established.

This assumption is critical because the QALY depends entirely upon multiplication. Survival time is unquestionably a ratio measure. The utility score is therefore required to function as a proportional adjustment factor. Unless the utility score itself possesses ratio properties, the multiplication is inadmissible. The resulting QALY becomes a numerical construction rather than a measure. The PBAC interrogation revealed no evidence that this foundational issue occupies a meaningful place within the analytical framework.

The interrogations of the research centers point to the same conclusion. All institutions support research and training activities associated with HTA. All contribute to the intellectual environment from which future researchers, analysts and policy advisers emerge. Yet no interrogation revealed meaningful recognition of the measurement requirements necessary to support utility-based claims. Utilities, QALYs and simulation modelling are accepted components of analysis, while the measurement foundations upon which they depend remain largely invisible.

The significance of this consistency should not be underestimated. Independent interrogations of policy agencies, research centers and academic institutions all point in the same direction. The issue is not confined to one organization or one group of researchers. It reflects a broader intellectual framework in which arithmetic has displaced measurement as the starting point for quantitative claims. A national consensus to promote paradigm failure

The consequences become particularly apparent when examining the role of utilities. Utility scores are routinely interpreted as though they represent proportions of health. A utility value of 0.8 is interpreted as 80 percent of full health; a value of 0.5 as 50 percent of full health. Yet proportional interpretation requires ratio measurement. A ratio measure requires a fixed non-arbitrary zero and meaningful proportional comparisons. Neither condition has been demonstrated for utility scores. The proportional interpretation is simply assumed and then embedded within the arithmetic of the QALY.

The same problem extends to reference-case simulation models. These models do not generate new measures. They manipulate quantities already assumed to be valid. If the utility scores entering the model lack ratio properties, every QALY generated by the model inherits the same defect. Increasing analytical sophistication cannot compensate for defective inputs. A simulation model cannot create measurement where measurement does not exist.

What makes the Australian evidence particularly compelling is its uniformity. Whether one examines HTA nationally, the PBAC specifically or the seven university-based research environments, the same pattern emerges. There is no identifiable point at which the requirements of representational measurement are systematically asserted. There is no identifiable point at which utility scores are required to demonstrate ratio properties. There is no identifiable point at which the admissibility of ratio arithmetic is established.

This is the defining characteristic of measurement inversion. Numerical operations are accepted first, while measurement is assumed to follow automatically. Arithmetic acquires authority simply because it is numerical. The distinction between a number and a measure disappears. All fail to understand the role of ratio measurement

The consequence is that the central constructs of Australian HTA rest upon assumptions rather than demonstrated measurement. Utilities are assumed to be ratio measures. QALYs are assumed to be valid quantitative outcomes. Simulation models are assumed to generate evidence. Yet the measurement requirements necessary to support these assumptions remain unmet.

Taken together, the national interrogation and the separate assessments provide compelling evidence that measurement inversion is not an isolated methodological weakness but a defining characteristic of the Australian HTA paradigm. The problem is therefore not one of implementation or technical refinement. It is foundational. The paradigm rests upon arithmetic before measurement. Once measurement is restored to its central role in scientific inquiry, the deficiencies become impossible to ignore. Measurement inversion is not merely present within Australian HTA; it is embedded within the framework itself. This is why the issue is no longer one of reform but of paradigm failure.

AUSTRALIA: PARADIGM FAILURE - THE EVIDENCE FOR CURRICULUM INVERSION

If measurement inversion explains why contemporary HTA fails the standards of measurement science, curriculum inversion explains why that failure has persisted. The two phenomena are closely connected. Measurement inversion concerns the analytical framework itself. Curriculum inversion concerns the educational environment that reproduces and sustains that framework. Together they provide compelling evidence of sustained paradigm failure in Australian HTA [17].

The assessment of curriculum inversion was restricted to the PBAC and seven leading Australian HTA research centers . The purpose was not to evaluate individual courses, lecture content, or training materials in isolation, but to determine whether the concepts required to support valid quantitative claims occupy a meaningful place within the educational and research environment that prepares future HTA practitioners.

The logic of the assessment was straightforward. A curriculum consistent with the standards of measurement science would be expected to show strong endorsement of concepts central to measurement and quantitative reasoning. These include attributes, scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement, dimensional homogeneity, Rasch measurement, and falsifiable claims. High endorsement probabilities would indicate that students, researchers, and policy analysts are likely to encounter these concepts during their professional development and therefore possess the intellectual tools necessary to evaluate the validity of quantitative claims. Conversely, low endorsement probabilities would indicate that these concepts are largely absent from the curriculum environment.

The significance of this distinction is important. Curriculum inversion does not occur merely because particular topics are omitted. It occurs when students are introduced to analytical techniques before they are introduced to the principles required to evaluate those techniques. In HTA this means that utilities, QALYs, cost-effectiveness analysis, and simulation modelling are taught and applied without first establishing the measurement foundations necessary to determine whether the resulting claims are scientifically defensible.

The interrogation results revealed a remarkably consistent pattern across both the PBAC and the Australian research centers. Concepts central to measurement science received low levels of endorsement, while concepts associated with the application of the reference-case paradigm received substantially stronger support. The implication is that future practitioners are being trained to operate within the framework rather than to evaluate its scientific legitimacy. This is the defining characteristic of curriculum inversion.

While there was recognition of attributes and, to a lesser extent, falsifiable claims, there was little recognition of representational measurement, scale theory, unidimensionality, latent attribute measurement, ratio measurement, Rasch measurement and the distinction between manifest and latent forms of measurement. The findings point to curriculum inversion, where the concepts required to evaluate the scientific legitimacy of the reference-case paradigm are largely absent from the educational structures responsible for reproducing that paradigm.

A scientific curriculum should equip students with the concepts necessary to evaluate the validity of quantitative claims. In HTA, this would require instruction in attributes, scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement, dimensional homogeneity and falsifiable claims. These concepts are not optional additions to professional training. They are the foundations upon which all quantitative reasoning depends. Without them, students may learn analytical techniques, but they cannot determine whether the outputs generated by those techniques are scientifically meaningful..

This omission is evident in the treatment of measurement scales. Across all interrogations, there is little indication that students are systematically exposed to the differences between nominal, ordinal, interval and ratio scales or to the implications these distinctions have for quantitative analysis. Yet these distinctions determine whether arithmetic operations are admissible. Without them, there is no basis for deciding whether multiplication, division or averaging can be performed meaningfully. The result is that students are likely to encounter numerical outputs without understanding the scale properties upon which those outputs depend.

The same pattern is evident in the treatment of representational measurement. The proposition that quantitative claims must be grounded in the axioms of measurement receives almost no endorsement. This finding is particularly important because representational measurement provides the scientific framework that distinguishes measures from numerical assignments. It establishes the conditions under which arithmetic is lawful and determines whether quantitative claims are defensible. Its absence leaves students without the conceptual tools necessary to evaluate the legitimacy of the methods they are taught to apply.

The interrogation results also reveal little recognition of unidimensionality. Measurement requires that an attribute represent a single dimension. Without unidimensionality, numerical aggregation may be possible, but measurement is not. Yet students are exposed to multidimensional instruments, composite indices and utility-based frameworks without systematic instruction in this requirement. The consequence is that aggregation is easily mistaken for measurement.

Equally significant is the treatment of manifest and latent attributes. The curriculum gives only limited recognition to the distinction between directly observable outcomes and latent constructs that require a measurement model. More importantly, there is almost no recognition that manifest and latent attributes require fundamentally different measurement frameworks. . Students are therefore unlikely to appreciate that survival, hospital admissions and treatment persistence pose different measurement problems from pain, fatigue, quality of life or patient functioning. The distinction between direct observation and latent attribute measurement is largely lost.

These omissions are not random. They occur precisely in those areas most likely to challenge the assumptions that underpin the reference-case paradigm. Students are taught utilities, QALYs, cost-effectiveness analysis and simulation modelling. They are not taught the measurement principles necessary to determine whether utilities support proportional interpretation, whether QALYs satisfy dimensional homogeneity or whether simulation outputs are grounded in valid measures. The result is an educational framework that reproduces the assumptions of the reference case while neglecting the concepts required to evaluate it critically.

This is the essence of curriculum inversion. Instead of beginning with measurement and proceeding to analysis, the curriculum begins with analytical techniques and assumes that measurement has already been established. The educational sequence mirrors the analytical sequence observed in measurement inversion. Arithmetic precedes measurement in practice because arithmetic precedes measurement in education.

The consequences extend beyond the classroom. Graduates move into research centers, policy agencies and professional organizations carrying the same conceptual limitations. They become researchers, reviewers, educators and decision makers. The assumptions embedded in the curriculum are then reproduced through research publications, HTA submissions, reimbursement decisions and future educational programs. Curriculum inversion therefore becomes the mechanism through which measurement inversion is sustained from one generation to the next.

Curriculum inversion thus provides the missing explanation for the persistence of measurement inversion. The former reproduces the latter. Together they reveal a paradigm that has become detached from the standards of measurement science while continuing to present itself as a quantitative discipline. This is why the challenge facing HTA in Australia is not merely methodological reform. It is the reconstruction of both the analytical framework and the curriculum that sustains it.

AUSTRALIA: UNDERSTANDING RATIO MEASUREMENT

The fatal weakness of contemporary health technology assessment (HTA) in Australia is not a failure of modelling, statistical analysis or computational sophistication. It is a failure to understand ratio measurement [18]. This may appear an unlikely criticism given the technical complexity of modern HTA, yet it lies at the heart of the paradigm failure documented throughout this report. Once the requirements imposed by ratio measurement are understood, the intellectual foundations of utilities, QALYs and reference-case simulation models become impossible to defend.

Ratio measurement occupies a unique position within science. A ratio scale possesses a true zero and supports all arithmetic operations, including multiplication and division. The importance of this property cannot be overstated. Multiplication is not a universal mathematical operation that can be applied to any numbers that happen to be available. Multiplication is admissible only when the quantities involved possess the measurement properties required to support proportional interpretation. If these properties are absent, the arithmetic result may be a number, but it is not a measure.

This principle is recognized throughout the physical sciences. No scientist would multiply quantities drawn from arbitrary scales and claim that the result represented a meaningful measure. The scale properties of the quantities must first be established. Measurement precedes arithmetic.

The tragedy of the reference-case paradigm is that this principle was never applied to the quantities upon which the framework depends. The architects of utility-based assessment assumed that preference scores could function as ratio measures. Health-state valuations obtained from time trade-off, standard gamble and related exercises were treated as though they represented proportions of health. A utility score of 0.8 came to be interpreted as 80 percent of full health, while a score of 0.5 was interpreted as 50 percent of full health. Yet no demonstration was provided that these values possessed the properties required for proportional interpretation.

This failure proved decisive because the entire QALY framework depends upon multiplication. Survival time is unquestionably a ratio measure. To multiply survival time by a utility score, the utility score must itself function as a ratio measure. If it does not, the operation is inadmissible. The resulting QALY is not a measure of quality-adjusted survival but a numerical construction generated through unlawful arithmetic.

The consequences extend even further. Reference-case simulation models depend upon QALYs as their central outcome measure. If the QALY lacks measurement status, every estimate generated by the model inherits the same defect. The complexity of the simulation is irrelevant. Sophisticated mathematics cannot rescue quantities that fail the requirements of measurement. Arithmetic cannot create measurement where measurement does not exist.

What makes this failure particularly significant in Australia is that the curriculum interrogations indicate little recognition of ratio measurement or the scale theory upon which it depends. Students are exposed to utilities, QALYs and simulation models but are not systematically introduced to the distinctions between nominal, ordinal, interval and ratio scales. They are not taught that different scales support different forms of arithmetic. Most importantly, they are not apparently taught that multiplication requires ratio measurement.

The consequence is predictable. Graduates learn how to calculate QALYs without learning whether QALYs are scientifically possible. They learn how to populate simulation models without learning whether the quantities entering those models qualify as measures. The scientific challenge is never encountered because the curriculum does not provide the concepts necessary to formulate it.

This omission helps explain why the reference-case paradigm remains largely unchallenged. If students are never introduced to the requirements of ratio measurement, they have no basis for questioning utility scores, QALYs or simulation outputs. The arithmetic appears legitimate simply because it is numerical. Measurement inversion becomes normalized because the measurement foundations are absent from professional education.

Yet once ratio measurement is restored to its central place, the situation changes dramatically. The entire structure of utility-based assessment depends upon a proposition that has never been demonstrated: that utility scores possess ratio properties. If this proposition fails, the QALY fails. If the QALY fails, reference-case simulation models fail. The collapse is not partial. It is complete because each component depends upon the one that precedes it.

This is why ratio measurement occupies such a critical place in the argument for paradigm failure. The issue is not whether utilities can be refined, whether valuation methods can be improved or whether simulation models can be made more sophisticated. The issue is whether the quantities entering those analyses satisfy the requirements of ratio measurement. If they do not, the resulting arithmetic is inadmissible regardless of the complexity of the analytical framework.

The evidence from Australia suggests that this fundamental issue remains largely invisible within both policy and education. The PBAC continues to employ a reference-case framework dependent upon utility-based assessment. Universities continue to teach the methods associated with that framework. Yet the scientific principles necessary to evaluate its legitimacy receive little attention.

The result is a self-reinforcing system in which arithmetic is accepted while measurement is assumed.

Understanding ratio measurement therefore changes everything. It reveals that the central weakness of the HTA paradigm is not methodological but conceptual. The failure to recognize the constraints imposed by ratio measurement undermines utilities, destroys the QALY and removes the foundation upon which reference-case simulation modelling depends. Once this is understood, the conclusion is unavoidable. The issue confronting HTA in Australia is not reform but reconstruction. Ratio measurement does not merely challenge the reference case; it exposes why the reference case cannot survive. A design fault that guaranteed rejection.

AUSTRALIA: MANIFEST AND LATENT ATTRIBUTES

The importance of ratio measurement has been further obscured by the failure of contemporary HTA to distinguish between manifest and latent attributes. Although these represent fundamentally different classes of outcomes, both require ratio measurement if quantitative claims regarding therapy impact are to be scientifically defensible. The crucial point is that the form of ratio measurement differs according to the nature of the attribute being assessed.

For manifest attributes, the measurement problem is relatively straightforward. Manifest attributes are directly observable. Survival time, hospital admissions, emergency department visits, medication possession, treatment discontinuation, adverse events, and resource utilization can all be observed, counted, or timed directly. When appropriately specified, these attributes support linear ratio measures characterized by a meaningful zero and admissible arithmetic operations. The measurement task is therefore one of observation and quantification.

Latent attributes present an entirely different challenge. Pain, fatigue, depression, anxiety, physical functioning, treatment satisfaction, need fulfilment, and quality of life are not directly observable. Their existence must be inferred from observable indicators such as questionnaire responses or behavioral observations. Consequently, latent attributes cannot be measured through direct observation alone. They require a measurement model capable of transforming ordinal observations into a quantitative estimate of attribute possession.

This is where Rasch assumes a unique importance. Rasch measurement provides the only established framework capable of constructing a lawful measure of latent attribute possession while simultaneously testing whether the requirements for measurement have been satisfied. Through the conjoint calibration of persons and items, Rasch establishes a latent continuum, evaluates unidimensionality, tests invariance, and determines whether observations support quantitative interpretation. The resulting logit scale is not merely another scoring system. It is a measurement framework specifically designed for latent attributes.

The significance of this distinction for HTA is profound. Had the discipline been forced from the outset to classify every target attribute as either manifest or latent, attention would inevitably have focused on the corresponding measurement requirements. Investigators would have been required to identify whether a proposed outcome demanded direct ratio measurement or a latent measurement model. The measurement problem would have become unavoidable. The present reliance on utilities, preference scores, and composite indices would have been difficult to sustain because the question would immediately arise: what is the attribute, and how is it being measured?

Instead, contemporary HTA largely bypassed this distinction. Outcomes with fundamentally different measurement properties were incorporated into a common evaluative framework. Directly observable events, patient-reported outcomes, preference scores, and utility algorithms were treated as though they occupied the same measurement status. Once this occurred, the need to distinguish between linear ratio measures and Rasch logit ratio measures disappeared from view. The measurement problem was effectively replaced by a scoring problem.

The consequence was that Rasch measurement never assumed the central role it should have occupied within HTA. If latent attributes are recognized as requiring their own form of ratio measurement, Rasch becomes indispensable. If the distinction between manifest and latent attributes is ignored, Rasch appears optional, peripheral, or unnecessary. The history of HTA reflects the latter path. The discipline embraced utilities, QALYs, and simulation models while largely overlooking the one framework capable of providing lawful measurement for latent outcomes.

Had HTA adopted the manifest-latent distinction at the outset, the history of the discipline would have been very different. Utilities would have been immediately challenged as measures of latent attributes. QALYs would have been challenged as inadmissible arithmetic. Rasch measurement would have occupied a central position in the assessment of patient-reported outcomes. The contemporary reference case might never have emerged in its present form.

Recognition of the manifest-latent distinction therefore changes the entire structure of HTA. It redirects attention from numerical constructions to measurement, from scores to attributes, and from preference weights to possession of measurable outcomes. Most importantly, it establishes that there are only two legitimate pathways to quantitative assessment: linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes. Once this distinction is accepted, the absence of Rasch from contemporary HTA is revealed not as a minor methodological omission but as one of the defining features of the paradigm's measurement failure.

AUSTRALIA: THE ABSENCE OF RASCH

One of the most striking findings from the Australian measurement and curriculum interrogation is the near-complete absence of Rasch ratio measurement and its role in the assessment of latent attributes [19]. This omission is critical in defining the paradigm failure of contemporary HTA. important because it reveals a fundamental weakness in the educational and methodological framework that underpins contemporary health technology assessment. The issue is not whether the term "Rasch" appears occasionally in conference abstracts, research presentations, or specialist publications. The issue is whether Rasch measurement is recognized as the essential framework for constructing quantitative measures of latent attributes. The interrogations say that it is not.

This omission is particularly significant because HTA places considerable emphasis on patient-centered outcomes, quality of life, symptom burden, functional status, treatment satisfaction, patient experience, and similar constructs. These are all latent attributes. They cannot be directly observed in the same way that hospital admissions, survival time, medication possession, or adverse events can be observed. Latent attributes exist, but they are not directly measurable through counting, timing, or simple observation. Their measurement requires a formal ratio measurement model.

This is where Rasch occupies a unique position. Rasch is not simply another psychometric technique competing with item response theory, PROMIS, utility instruments, or preference-based scoring systems. Rasch addresses a fundamentally different question. It asks whether ordinal observations can be transformed into a quantitative measure of possession of a latent attribute. In doing so, it provides the only established framework capable of demonstrating whether the conditions required for measurement have been satisfied for latent attributes.

The distinction is critical. Patient-reported outcomes typically begin with ordinal responses to questionnaire items. Patients may indicate levels of pain, fatigue, anxiety, mobility limitations, or functional difficulties. These responses are rankings. They provide information about order but not quantity. Arithmetic performed directly on ordinal observations cannot create measurement. Summing scores, averaging responses, applying weights, or generating utility algorithms does not transform ordinal observations into quantitative measures. Numerical manipulation is not measurement.

The Rasch model was developed in the 1950s precisely to address this problem. Through the conjoint calibration of persons and items, Rasch analysis estimates the location of respondents on a latent continuum while simultaneously testing whether the data satisfy the requirements for ratio measurement. Unidimensionality, invariance, item fit, category functioning, local independence, and differential item functioning are not optional refinements. They are the conditions that must be satisfied before claims regarding possession of a latent attribute can be advanced. Rasch therefore provides both a measurement model and a set of empirical tests for determining whether measurement is possible.

The interrogations make clear that this perspective is absent from the Australian HTA educational framework. Students and practitioners are introduced to patient-reported outcomes, utility instruments, preference weights, quality-of-life measures, and value assessment methodologies without first confronting the measurement problem those constructs are intended to address. The curriculum appears to move directly from patient responses to scoring systems and economic evaluation. The intermediate step, demonstrating that a latent attribute has been measured, is effectively bypassed.

This omission has important consequences. Without Rasch ratio measurement, latent attributes remain latent. Utility scores, composite indices, and preference-weighted algorithms may generate numerical outputs, but they do not establish that the underlying construct has been measured. The existence of a number should not be confused with the existence of a measure. Yet much of contemporary HTA proceeds as though this distinction does not matter.

The result is that students are trained to accept numerical representations of quality of life, patient benefit, symptom burden, and treatment impact without being introduced to the Rasch framework required to determine whether those representations possess measurement properties. They learn how utilities are generated, how QALYs are constructed, and how economic models are populated, but they are not taught how latent attributes can be measured. The educational sequence is therefore inverted. Numerical outputs are presented before the conditions required to justify those outputs.

The absence of Rasch is consequently more than a methodological omission. It is a defining characteristic of curriculum inversion. The curriculum recognizes the importance of latent attributes but fails to recognize the only framework capable of transforming observations of those attributes into quantitative ratio measures. This leaves faculty, students and practitioners with a vocabulary of scores, utilities, and indices but without an understanding of measurement itself. Until Rasch measurement assumes its proper place within HTA education, latent attributes will continue to be represented through numerical constructions rather than lawful measures, and the distinction between scoring and measurement will remain obscured.

AUSTRALIA: THE HISTORICAL SPECTACLE

One of the most remarkable aspects of the HTA paradigm is not the nature of its failure but its longevity. Errors occur in every scientific discipline. Hypotheses are proposed, tested, challenged and, when necessary, abandoned. This process is fundamental to scientific progress. What makes the history of HTA unusual, if not unique, is that a framework that failed the requirements of measurement became established as the dominant model for evaluating therapy impact and remained largely unchallenged for more than four decades at a global level.

The scale of this enterprise is difficult to overstate. Since the emergence of utility-based assessment in the late twentieth century, thousands, and perhaps tens of thousands, of papers have been published employing utility scores, QALYs, cost-effectiveness ratios and reference-case simulation models. Entire academic careers have been devoted to the development and refinement of these methods (e.g., the latest PubMed hit for QALYS yields 28,870 responses). Research centers have been established to advance them. Professional societies have promoted them. Governments have incorporated them into reimbursement decision making. Universities have embedded them within curricula and professional training programs. An extensive international infrastructure emerged to support and sustain the reference-case paradigm.

Viewed from within the discipline, this growth was often interpreted as evidence of success. The expanding literature, increasing methodological sophistication and widespread institutional adoption appeared to demonstrate intellectual maturity. Yet institutional acceptance and scientific validity are not the same thing. The crucial question was never how many studies employed the framework but whether the framework itself satisfied the standards required for quantitative claims.

The historical spectacle lies in the fact that this question remained largely unasked. Utility scores were accepted as though they possessed ratio properties. QALYs were accepted as though multiplication was admissible. Simulation models were accepted as though they generated quantitative evidence. At every stage, arithmetic was treated as self-justifying. The prior requirement that measurement be demonstrated before arithmetic could proceed was largely ignored.

This is not a criticism of individual investigators. Most researchers worked within the intellectual environment they inherited; they were intellectually locked in. The concepts, methods and assumptions of the reference case were presented as established knowledge. Young investigators learned them from textbooks, mentors and professional training programs. Manuscripts employing these methods passed through peer review. Research grants supported further methodological development. Professional advancement often depended upon participation in the framework rather than its criticism.

The consequence was the creation of a self-reinforcing intellectual system or, in the terminology of Richard Dawkins, a memeplex. Authors wrote papers based upon accepted assumptions. Reviewers evaluated submissions using the same assumptions. Editors selected manuscripts from a pool of work operating within the same conceptual boundaries. Research centers trained students using the same methods. Policy agencies relied upon the resulting analyses. The framework reproduced itself because the standards used to evaluate it were derived from the framework itself.

The significance of the recent measurement and curriculum interrogations is that they provide, for the first time, a systematic explanation for how this occurred. The persistence of the reference case cannot be explained simply by professional inertia or institutional conservatism. The interrogations reveal something more fundamental. The concepts required to identify the problem are themselves largely absent from the HTA knowledge base. Representational measurement, admissible arithmetic, dimensional homogeneity, unidimensionality, ratio measurement, the distinction between manifest and latent attributes and the role of Rasch measurement occupy little place in the curriculum, research and policy structures of the discipline.

This finding is important because it shifts the discussion from methodological criticism to paradigm failure. The problem is not merely that HTA adopted a flawed framework. The problem is that the intellectual tools required to recognize the flaw were largely excluded from the memeplex. Researchers could not easily challenge assumptions regarding measurement because the principles necessary to formulate that challenge were absent from their professional education and research environment.

A further implication of paradigm failure is that what has been presented as evidence within HTA was not discovered through measurement but constructed through convention. The reference-case framework did not begin by identifying measurable attributes and then building evidence upon those measures. Instead, it began with assumptions regarding utilities, QALYs, and simulation modelling and treated the resulting numerical outputs as evidence. The status of those outputs as evidence was therefore not established by measurement but conferred by acceptance within the HTA community itself. What emerged was a professional consensus regarding what counted as evidence rather than a framework grounded in demonstrated measurement properties. The persistence of the reference case reflects the stability of that consensus rather than the successful resolution of the measurement questions upon which quantitative claims ultimately depend.

The result is one of the most striking episodes in the history of applied research. An emerging discipline adopted an analytical framework whose central constructs failed the requirements of measurement. Rather than examining those foundations, successive generations of researchers devoted their efforts to elaborating the framework. More sophisticated utility instruments were developed. More complex simulation models were constructed. Larger datasets were assembled. Increasingly advanced analytical techniques were employed. Yet the fundamental question remained untouched: are the quantities entering these analyses measures?

This is not unprecedented. The history of science contains examples of theories and practices that persisted for long periods despite foundational weaknesses. What distinguishes HTA is that the persistence of the paradigm was reinforced by the absence of the very concepts required to challenge it. The framework became self-reinforcing because measurement itself disappeared from view. As a consequence, criticism focused on methodological refinement rather than on the more fundamental question of whether the framework could satisfy the requirements of quantitative science.

From the perspective of measurement science, the answer is clear. The issue is not that the reference case occasionally produced misleading results. The issue is that its central constructs never satisfied the conditions necessary to support the arithmetic upon which they depended. The historical spectacle is therefore not merely the persistence of error. It is the persistence of error on an extraordinary scale.

Future historians of HTA may regard this period as a cautionary example of how scientific disciplines can become detached from their own foundations. The spectacle is not that the methods lacked sophistication. The spectacle is that increasing sophistication concealed an elementary failure. For more than forty years, an immense intellectual effort was devoted to refining arithmetic while neglecting measurement. The result was not the accumulation of quantitative knowledge but the institutionalization of measurement inversion.

That is why the present moment represents more than a methodological dispute. It marks the point at which an entire analytical tradition must confront the standards it neglected. The historical spectacle is coming to an end. What remains is the task of reconstruction.

AUSTRALIA: MEETING PROFESSIONAL STANDARDS

The evidence presented in this paper points to a conclusion that can no longer be avoided. The professional standards required to evaluate health technologies scientifically are not present in the Australian HTA curriculum knowledge base. This is not a matter of incomplete coverage or a need for modest curriculum adjustment. The omission concerns the foundations of quantitative science itself. Students and researchers are introduced to the reference case, utilities, QALYs, economic evaluation and simulation modelling, but they are not systematically introduced to the measurement principles necessary to determine whether these constructs can support valid quantitative claims.

This distinction is decisive. A professional curriculum should not simply train students to operate within an existing framework; to pull levers. It should equip them to evaluate whether that framework satisfies the standards of scientific inquiry. In HTA, that means understanding attributes, scales of measurement, representational measurement, unidimensionality, dimensional homogeneity, manifest and latent attributes, Rasch measurement, ratio measurement and falsifiable claims. These are not optional topics. They are the intellectual tools required to distinguish a measure from a number, lawful arithmetic from inadmissible arithmetic, and evidence from numerical construction. Without them the claim to provide a professional training falls short.

The Australian curriculum interrogations indicate that these tools are largely absent. The concepts most essential for evaluating the reference case receive the weakest endorsement. Scales of measurement are not central. Representational measurement is effectively invisible. Unidimensionality receives minimal recognition. Rasch and latent attribute measurement is absent. The distinction between manifest and latent forms of ratio measurement is almost absent. As a result, the curriculum does not prepare graduates to evaluate the measurement foundations of HTA. It prepares them to apply the reference case.

The consequence is professionally disastrous. A lifetime commitment to the reference case is no longer defensible once the requirements of measurement are recognized. The reference case is not a neutral analytical tool awaiting refinement. It is a sequence of dependent errors. Time trade-off responses are treated as proportional values without demonstrating ratio properties. Health-state valuations are converted into utility weights without establishing measurement status. Utilities are multiplied by survival time without proving that multiplication is admissible. QALYs are then inserted into simulation models that project lifetime claims using quantities whose measurement foundations remain unproven. One error follows another. The reference case has nothing to offer.

These errors remain largely invisible so long as scales of measurement and representational measurement are absent from professional training. If students are never taught that different scales support different arithmetic operations, they cannot see why utility scores may not support multiplication. If they are never taught that ratio measurement requires a true zero, they cannot see why bounded utility values do not automatically function as proportions. If they are never taught dimensional homogeneity, they cannot see why multiplying ordinal or interval preference scores by ratio time is inadmissible. If they are never taught the distinction between manifest and latent attributes and the contribution of Rasch measurement, they cannot see why quality of life requires a different measurement framework from survival time or hospitalization.

This is why curriculum reconstruction is unavoidable. The status quo curriculum is unsustainable because it reproduces the very paradigm that has failed. It trains students to accept utilities, QALYs and reference-case simulation models while withholding the measurement framework that would expose their weaknesses. This is not professional preparation. It is professional enculturation into a defective paradigm. Welcome to the memeplex.

The required alternative is already clear. A reconstructed curriculum must begin with measurement. Students must first learn that every therapy assessment concerns an attribute. They must then distinguish manifest from latent attributes. Manifest attributes, such as survival time, hospital admissions, emergency department visits and treatment persistence, may support direct linear ratio measurement. Latent attributes, such as pain, fatigue, functional status, need fulfilment and quality of life, require a measurement model capable of estimating possession of the attribute. Rasch measurement provides the necessary framework for constructing measures of latent attribute possession. Only after measurement has been demonstrated can arithmetic and claims assessment proceed.

A professional HTA curriculum must therefore move from simulation method training to scientific training. It must teach students how to construct evaluable claims, specify measurement models, determine scale properties, assess admissible arithmetic and frame claims capable of empirical falsification. The objective should not be proficiency in the reference case. The objective should be competence in determining whether any proposed analytical framework satisfies the requirements of measurement and normal science.

A further consequence of paradigm failure is that the educational resources required for reconstruction are largely absent. Existing HTA textbooks and teaching materials were developed within the reference-case tradition and therefore reproduce many of the assumptions that have contributed to measurement inversion. They provide extensive coverage of utilities, QALYs, cost-effectiveness analysis and simulation modelling, but virtually no coverage of representational measurement, ratio measurement, dimensional homogeneity, Rasch measurement, manifest and latent attributes or falsifiable claims. As a consequence, there is currently no widely adopted textbook framework available to guide the transition from arithmetic-based assessment to measurement-based assessment. *To address this gap, Maimon Research LLC has developed a nine-unit educational program designed specifically to support curriculum reconstruction* [20]*. The program introduces the principles of measurement science, distinguishes manifest from latent attributes, outlines the requirements for valid ratio measurement, presents Rasch measurement for latent attributes and provides a framework for developing evaluable and falsifiable claims regarding therapy impact. The objective is not to modify the reference case but to provide the educational foundation for its replacement.*

For Australia, the implications are unambiguous. The PBAC and the academic research centers occupy important positions in policy, research and professional development. If these institutions continue to reproduce a curriculum that omits measurement science while endorsing reference-case methods, the result will be continued measurement inversion and irrelevance. Graduates will enter professional roles able to conduct analyses but unable to judge whether the analyses rest on valid measures.

AUSTRALIA: CONCLUSIONS

The findings reported in this assessment leave little room for ambiguity. The problem confronting HTA is not one of methodological refinement, improved modelling techniques, revised utility instruments or updated reference-case guidance. The problem is foundational. The standards required for quantitative claims have been absent from the framework from the beginning. What has been documented across agencies, research centers, academic programs and professional curricula is not an isolated methodological weakness but a pervasive pattern of measurement inversion and curriculum inversion. The same errors appear repeatedly because the concepts required to recognize those errors are themselves largely absent from the knowledge base.

This conclusion has important implications. There can be no return to the status quo. The reference-case paradigm is not a framework that can be repaired through incremental modification because the defects are embedded in its foundations. Utilities remain ordinal numerical constructions. QALYs remain mathematically inadmissible because multiplication requires ratio properties that have never been demonstrated. Reference-case simulation models remain exercises in numerical speculation because the quantities entering those models do not satisfy the requirements of measurement. No amount of statistical sophistication, computational complexity or methodological elaboration can transform non-measures into measures. Arithmetic cannot create measurement where measurement does not exist.

It is therefore misleading to speak of reform. Reform implies that a structure remains fundamentally sound and requires adjustment. The evidence presented here points to a different conclusion. The contemporary HTA framework never satisfied the standards required for quantitative science. The problem is not that the framework has deteriorated over time. The problem is that it was established without first securing the measurement foundations upon which quantitative claims depend. The reference-case paradigm did not gradually lose its way. It began with assumptions regarding utility measurement, QALY construction and simulation modelling that failed to satisfy the requirements of representational measurement. Four decades of methodological development have merely extended and elaborated assumptions that should have been challenged at the outset.

The consequence is that only one path remains open. HTA must be reconstructed upon the foundations of measurement science. These standards are not optional methodological preferences. They are the minimum requirements for scientific inquiry. A discipline that wishes to make quantitative claims cannot exempt itself from the standards governing quantitative science. The future of HTA therefore depends not upon defending the reference case but upon abandoning it. The task ahead is reconstruction: replacing a framework built on arithmetic before measurement with one grounded in measurement before arithmetic. Only then can HTA move beyond numerical constructions and establish a credible scientific foundation for the evaluation of therapeutic interventions.

ACKNOWLEDGEMENT

I acknowledge that I have used OpenAI technologies, including the large language model, to assist in the development of this work. All final decisions, interpretations, and responsibilities for the content rest solely with me.

REFERENCES

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

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

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

[4] Australia: Pharmaceutical Benefits Advisory Committee (PBAC) – Decisions without measurement. Logit Working Paper No 34. https://maimonresearch.com/logit-working-paper-no-34/

[5] Australia: The end of the PBAC reference case – Reconstructing HTA around measurement. Logit Working Paper No 172. https://maimonresearch.com/logit-working-paper-no-173-june-2026/

[6] Australia: A national consensus on the absence of measurement in health technology assessment. Logit Working Paper No 32. https://maimonresearch.com/logit-working-paper-no-32/

[7] Australia: Closed epistemic reproduction with five Australian HTA academic knowledge bases sustaining measurement inversion. Logit Working Paper No 810. https://maimonresearch.com/logit-working-paper-no-810-april-2026/

[8] Australia: The absence of representational measurement and the Melbourne School of Population and Global Health (MSPGH) University of Melbourne. Logit Working Paper No 724. https://maimonresearch.com/logit-working-paper-no-724-february-2026/

[9] Australia: The absence of representational measurement and the Centre for Health Economics (CHE) Monash University. Logit Working Paper No 722. https://maimonresearch.com/logit-working-paper-no-722-february-2026/

[10] Australia: The absence of representational measurement and the Health Economics Group (HEG) University of Adelaide. Logit Working Paper No 723. https://maimonresearch.com/logit-working-paper-no-723-february-2026/

[11] Australia: The absence of representational measurement and the Centre for Applied Health Economics (CAHE) Griffith University. Logit Working Paper No 727. https://maimonresearch.com/logit-working-paper-no-727-february-2026-2/

[12] Australia: The absence of representational measurement and the Leeder Centre for Health Policy, University of Sydney. Logit Working Paper No 791. https://maimonresearch.com/logit-working-paper-no-791-april-2026/

[13] Australia: The absence of representational measurement and the Centre for Health Economics Research and Evaluation (CHERE) University of Technology Sydney. Logit Working Paper No 721. https://maimonresearch.com/logit-working-paper-no-721-february-2026/

[14] Australia: The absence of representational measurement and the Deakin University Health Economics and HTA knowledge base. Logit Working Paper No 281 https://maimonresearch.com/logit-working-paper-no-281-may-2026/

[15] Australia: There are only two lawful measures for therapy impact assessment. Logit Working Paper No 1811. https://maimonresearch.com/logit-working-paper-no-1811-june-2026/

[16] Maimon Research LLC. Website Interrogations: Australia https://maimonresearch.com/ai-llm-rest-of-world/#australia

[17] Australia: The missing science of measurement in health technology assessment. Logit Working Paper No 1204. https://maimonresearch.com/logit-working-paper-no-1204-june-2026/

[18] Australia: There are only two lawful measures for therapy impact assessment. Logit Working Paper No 1811. https://maimonresearch.com/logit-working-paper-no-1811-june-2026/

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

[20] Mamon Research LLC. https://maimonresearch.com/hta-reconstruction-program-and-fees/