MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

AUSTRALIA: THERE ARE ONLY TWO LAWFUL MEASURES FOR THERAPY IMPACT ASSESSMENT

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

LOGIT WORKING PAPER No 1811 JUNE 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

Australia occupies a distinctive position in the history of health technology assessment (HTA). Through the Pharmaceutical Benefits Advisory Committee (PBAC), Australia became one of the earliest and most influential adopters of the reference-case approach to reimbursement decision making. Over four decades, Australian universities, research centres and policy agencies helped institutionalize a framework based upon utilities, QALYs and simulation modelling. This paper argues that Australia provides a revealing case study of how measurement inversion became embedded within a national HTA system despite the availability of a lawful alternative.

The central thesis is that the path taken by Australian HTA was unnecessary. Once therapy outcomes are recognized as either manifest or latent attributes, the requirements for measurement are straightforward. Manifest attributes require linear ratio measures. Latent attributes require Rasch logit ratio measures. Together these two forms of measurement provide a complete framework for the assessment of therapy impact. Rather than adopting this approach, Australian HTA embraced utility construction, QALY estimation and reference-case simulation modelling. The result was a framework in which arithmetic routinely displaced measurement.

The consequences have been substantial. Health-state descriptions were transformed into utility scores, utility scores became inputs to QALYs and QALYs became central components of simulation models used to support reimbursement decisions. Yet at no stage was there a demonstration that the entities entering these calculations possessed the measurement properties necessary to support the arithmetic operations imposed upon them. The result was measurement inversion, where arithmetic preceded measurement, and arithmetic chaos, where one inadmissible numerical operation built upon another.

Recent interrogations of the PBAC and leading Australian HTA research centres reveal a consistent pattern of endorsement for utilities, QALYs and simulation modelling alongside little recognition of representational measurement, dimensional homogeneity and Rasch measurement. The findings suggest that Australia has institutionalized a framework whose quantitative claims cannot be justified according to the standards of measurement science. The lesson is clear. The problem confronting Australian HTA is not a lack of analytical sophistication but a failure to distinguish between numbers and measures. The future of HTA in Australia lies not in refining the reference case but in reconstructing assessment around the only two lawful forms of measurement: linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes.

.

INTRODUCTION

For more than four decades health technology assessment (HTA) has operated under a fundamental misconception. The prevailing assumption has been that the principal challenge in evaluating therapies is to generate numbers. Once numerical values are available, whether utilities, QALYs, cost-effectiveness ratios or simulation outputs, they are assumed to provide a quantitative basis for decision making. The result has been an extraordinary expansion in analytical complexity. Health-state descriptions have been transformed into utilities, utilities into QALYs and QALYs into cost-

effectiveness claims supported by increasingly sophisticated simulation models. Yet throughout this process a more fundamental question has largely been ignored: what is being measured?

Australia provides a particularly revealing example of how this misconception became institutionalized. Through the Pharmaceutical Benefits Advisory Committee (PBAC), Australia was among the earliest countries to embrace the reference-case approach to reimbursement decision making. Over time, a network of university research centers, health economists and policy analysts emerged to support and refine this framework. Utilities, QALYs and simulation models became the accepted language of HTA. Yet remarkably little attention was given to the measurement foundations required to support these numerical constructions. The result was a system in which arithmetic came to dominate assessment while measurement itself remained largely unexamined [i].

The irony is that this path was entirely unnecessary. Once therapy outcomes are recognized as either manifest or latent attributes, the requirements for lawful measurement become straightforward. Manifest attributes require linear ratio measures. Latent attributes require Rasch logit ratio measures. Together these provide a complete framework for the assessment of therapy impact. There was no need to construct utilities from health-state descriptions, no need to create QALYs and no need to populate simulation models with entities whose measurement status was unknown. Australia nevertheless followed the reference-case pathway and, in doing so, became increasingly committed to a framework that substituted arithmetic for measurement[ii] .

The consequences are now becoming apparent. Recent interrogations of the PBAC and leading Australian HTA research centers reveal a consistent pattern of endorsement for utilities, QALYs and simulation modelling alongside little recognition of representational measurement, dimensional homogeneity and Rasch measurement [iii]. These findings point to measurement inversion: arithmetic preceding measurement. The result is arithmetic chaos, where one inadmissible numerical operation builds upon another until the final outputs emerge as cost-effectiveness claims, reimbursement recommendations and policy decisions. Australia therefore provides an important lesson for HTA. The problem has never been a lack of analytical sophistication. The problem has been a failure to recognize that there are only two lawful measures and that everything else is arithmetic.

The purpose of this paper is therefore straightforward. It is to demonstrate that, for the assessment of therapy impact, there are only two lawful measures. Recognizing this fact not only explains the emergence of measurement inversion and arithmetic chaos within Australian HTA, but also provides the foundation for its reconstruction.

The neglect of this question has had profound consequences. Across the physical sciences, engineering and the natural sciences, arithmetic follows measurement. Quantitative operations are undertaken only after the measurement properties of the entities involved have been established. In HTA, however, the reverse has become commonplace. Numerical operations are routinely applied to constructs whose measurement status is either unknown or assumed. This phenomenon, described throughout this series of papers as measurement inversion, has created the conditions for arithmetic chaos. Numbers are multiplied, divided, aggregated and incorporated into simulation

models without first demonstrating that the underlying attributes possess the properties necessary to support those operations.

The source of much of this confusion lies in a failure to distinguish between two fundamentally different classes of attributes. Some attributes are manifest. They are directly observable and can be measured through straightforward counting or observation. Hospital admissions, hospital days, emergency department visits, physician encounters, treatment persistence and resource utilization are examples. These attributes support linear ratio measurement with meaningful zeros and interpretable magnitudes. Other attributes are latent. Pain, fatigue, depression, physical functioning, symptom burden, treatment satisfaction and need fulfilment cannot be observed directly. Their existence must be inferred from observations. These attributes require a completely different measurement framework.

Once this distinction is recognized, the path forward becomes remarkably clear. There are only two lawful forms of measurement available for the assessment of therapy impact. Manifest attributes require linear ratio measures. Latent attributes require Rasch logit ratio measures. Together these provide the only basis for constructing empirically evaluable, replicable and falsifiable claims regarding treatment outcomes. No other measurement framework is required. More importantly, no other measurement framework is capable of supporting the arithmetic operations necessary for quantitative science.

The significance of this conclusion extends far beyond a technical debate over measurement theory. It challenges the foundations of the contemporary reference-case paradigm. Utilities, QALYs and simulation models are all dependent upon assumptions concerning measurement that are rarely examined and seldom justified. By contrast, a framework built upon linear ratio measures and Rasch logit ratio measures restores the principle that measurement must precede arithmetic. It provides a direct route from observable evidence to evaluable claims without the need for utility construction, imaginary future worlds or non-falsifiable simulations.

The purpose of this paper is therefore straightforward. It is to demonstrate that, for the assessment of therapy impact, there are only two lawful measures. Recognizing this fact not only resolves the confusion that has characterized modern HTA but also provides the foundation for its reconstruction. Measurement inversion and arithmetic chaos are not inevitable. They are the consequence of abandoning the discipline of measurement. The future of HTA depends upon restoring it.

LINEAR RATIO MEASURES AND MANIFEST ATTRIBUTES

The first lawful form of measurement relevant to therapy assessment is the linear ratio measure. Although the concept is straightforward, it has been largely overshadowed within contemporary HTA by an emphasis on utilities, QALYs and simulation models. Yet it is linear ratio measurement that provides the foundation for quantitative science. Without ratio measurement there can be no meaningful multiplication, division, comparison of magnitudes or quantitative inference. The assessment of therapy impact therefore begins not with utilities or health-state descriptions but with the identification of attributes that can be measured directly.

Manifest attributes are directly observable characteristics. Their existence does not have to be inferred from responses to questionnaires, preference exercises or health-state classifications. They can be counted, observed or recorded. Examples include hospital admissions, hospital days, emergency department visits, physician encounters, ICU hours, treatment persistence, medication possession, treatment switching and resource utilization. These are not hypothetical constructs. They are observable events that occur within the real world of patient care.

The defining feature of a manifest attribute is that it supports a meaningful zero. A patient with zero hospital admissions has experienced no admissions. A patient with zero emergency department visits has not attended an emergency department. Similarly, a patient with ten hospital days has experienced twice as many hospital days as a patient with five hospital days. The magnitudes are interpretable because the attribute possesses ratio properties. Arithmetic operations involving addition, subtraction, multiplication and division are therefore admissible.

This may appear self-evident, but its implications for HTA are profound. Once an attribute is measured on a linear ratio scale, claims regarding therapy impact become immediately evaluable. A manufacturer can claim that a therapy reduces hospital admissions by a specified amount over a defined period. The claim can be tested empirically. Data can be collected, analyzed and replicated. Independent investigators can attempt to confirm or refute the result. The claim is therefore falsifiable. This is precisely how quantitative science is intended to operate.

The contrast with the reference-case framework is striking. Rather than beginning with directly observable attributes, the reference case focuses on health-state descriptions, utility construction and simulation models. The result is a framework in which arithmetic operations are applied to entities whose measurement status remains uncertain. Linear ratio measures avoid this problem entirely. The measurement properties of the attribute are established before arithmetic is undertaken. Measurement precedes arithmetic rather than the reverse.

Manifest attributes also provide a direct link between clinical outcomes and health system objectives. A reduction in hospital admissions, emergency department visits or treatment discontinuations has immediate practical significance. The outcome is understandable to clinicians, patients, payers and policy makers. Unlike utility scores, the measure does not require interpretation through preference algorithms or hypothetical valuations. The attribute itself is the outcome of interest.

Importantly, the use of linear ratio measures does not imply that all therapy impact can be reduced to observable events. Many important treatment outcomes involve latent attributes such as pain, fatigue, functioning and quality of life. These cannot be measured directly and require a different measurement framework. The existence of latent attributes, however, does not diminish the importance of manifest attributes. Rather, it reinforces the need to distinguish clearly between the two.

The reconstruction of HTA therefore begins with a recognition that manifest attributes require linear ratio measurement. This is not a matter of methodological preference but a consequence of the requirements of representational measurement. When the attribute is directly observable, the objective is to construct a measure that preserves meaningful magnitudes, supports admissible

arithmetic operations and allows empirical evaluation. Linear ratio scales satisfy these requirements. They provide a lawful basis for quantitative claims regarding therapy impact and establish the first of the two measurement frameworks necessary for a scientifically credible HTA.

The importance of this conclusion cannot be overstated. Once manifest attributes are recognized as requiring linear ratio measurement, much of the complexity that characterizes contemporary HTA becomes unnecessary. Claims can be framed directly in terms of observable outcomes, evaluated through explicit protocols and subjected to replication and falsification. The result is a return to the principles that govern quantitative inquiry in every other scientific discipline. For manifest attributes, the path is clear: linear ratio measurement provides the only lawful foundation for assessing therapy impact.

RASCH LOGIT RATIO MEASURES AND LATENT ATTRIBUTES

While manifest attributes can be measured directly through observation and counting, many of the outcomes of greatest interest in therapy assessment are latent attributes. These include pain, fatigue, depression, anxiety, physical functioning, symptom burden, treatment satisfaction and need fulfilment. Such attributes cannot be observed directly. A patient cannot be examined and assigned a direct measure of pain in the same way that a hospital admission can be counted or a length of stay recorded. The existence of the attribute must instead be inferred from observable indicators. This distinction is crucial because latent attributes require a fundamentally different measurement framework.

The challenge of latent measurement has long been misunderstood within HTA. The conventional approach has been to administer questionnaires, assign numerical scores to responses and aggregate those scores to create a summary measure. The resulting totals are often treated as though they were quantitative measures capable of supporting arithmetic operations and comparisons of therapy impact. Yet the assignment of numbers to questionnaire responses does not create measurement. The resulting scores remain ordinal structures. They provide rankings but do not establish equal intervals, meaningful units or ratio properties.

This is where Rasch measurement assumes central importance. Developed by Georg Rasch and subsequently formalized within the framework of representational measurement, the Rasch model provides the only established method for transforming observations of latent attributes into lawful measures [iv]. The objective is not to generate scores but to construct a measurement system in which persons and items are located simultaneously on a common latent continuum. The resulting measure is expressed in logits, the natural logarithm of the odds ratio.

The importance of the Rasch framework lies in its strict measurement requirements. The model demands unidimensionality, local independence and invariance. Items must collectively represent a single underlying attribute. Responses must be independent once the latent trait is taken into account. Most importantly, the relationship between persons and items must be invariant across populations and contexts. These requirements are not optional. They are the conditions necessary for measurement.

The outcome of interest in Rasch measurement is not a questionnaire score. It is the possession of the latent trait itself. A patient's position on the latent continuum represents the extent to which the attribute is possessed relative to other persons and items within the measurement system. Therapy impact is therefore assessed through changes in possession of the latent trait rather than changes in arbitrary ordinal scores. This distinction separates Rasch measurement from conventional psychometric approaches, item response theory models and preference-based utility instruments.

The resulting Rasch logit scale possesses ratio properties through the odds structure underlying the model. Changes in logits can be interpreted quantitatively and transformed into odds ratios. The measurement framework therefore supports meaningful comparisons and provides a lawful basis for assessing therapy impact in latent attributes. Unlike summed ordinal scores, the Rasch measure is not an artifact of arbitrary scoring conventions. It is the outcome of a measurement model specifically designed to satisfy the requirements of representational measurement.

The implications for HTA are substantial. Many outcomes currently represented through utilities, quality-of-life instruments and patient-reported outcome measures are latent attributes. Yet the conventional approach treats questionnaire responses as though arithmetic operations alone can create measurement. Rasch measurement demonstrates that this is not the case. Measurement requires a formal measurement model, not a scoring algorithm. The objective is not to produce numbers but to construct lawful measures.

The reconstruction of HTA therefore requires recognition of a simple principle. Manifest attributes demand linear ratio measures. Latent attributes demand Rasch logit ratio measures. There is no third category. Once this distinction is accepted, much of the confusion surrounding patient-reported outcomes, quality-of-life instruments and utility construction disappears. The question is no longer how to manipulate ordinal scores through increasingly sophisticated algorithms. The question is whether the latent attribute has been measured according to the requirements of Rasch measurement.

For latent attributes, the conclusion is clear. Rasch logit ratio measurement provides the only lawful basis for assessing therapy impact. It therefore represents the second and final measurement framework required for a scientifically credible system of health technology assessment. Once combined with linear ratio measures for manifest attributes, the foundation exists for a reconstructed HTA grounded in lawful measurement, empirical evaluation and falsification rather than measurement inversion and arithmetic chaos.

WHY THERE ARE NO OTHER MEASURES

The argument advanced in this paper is intentionally simple. For the assessment of therapy impact there are only two lawful forms of measurement: linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes. This conclusion raises an obvious question. What about the numerous measurement systems, utility instruments, patient-reported outcome measures and composite indices that currently dominate HTA? If these are not lawful measures, what are they?

The answer lies in a failure to distinguish between numbers and measures. The assignment of numbers to observations does not, in itself, create measurement. Numbers may be used as labels, rankings, classifications or scores without satisfying the requirements of representational measurement. The critical issue is not whether a numerical value exists but whether the numerical value preserves the properties of the attribute being measured and supports admissible arithmetic operations.

This distinction explains why so many contemporary HTA constructs fail as measures. Utility instruments such as the EQ-5D and the Health Utilities Index begin with health-state descriptions and preference valuations. The resulting utility scores are generated through scoring algorithms that combine attribute levels and preference weights. Yet the existence of an algorithm does not establish measurement. Health-state classifications are not measures, preference structures are ordinal and arithmetic operations applied to ordinal structures cannot create ratio measures. The final utility score is therefore a numerical construct rather than a lawful measure.

The same criticism applies to summed questionnaire scores. Instruments such as quality-of-life scales, symptom inventories and treatment satisfaction questionnaires frequently assign numbers to response categories and aggregate the results into a total score. This practice remains widespread despite a simple limitation: ordinal scores support ranking, not arithmetic. Summing ordinal responses does not transform them into interval or ratio measures. The resulting total score remains an ordinal construct irrespective of the complexity of the scoring procedure.

The QALY provides perhaps the most influential example. The QALY is often presented as a universal measure of health benefit capable of supporting comparisons across diseases and interventions. Yet the QALY depends entirely upon the utility values from which it is constructed. If the utility lacks ratio properties, multiplication by time cannot create a lawful measure. The resulting QALY is therefore not a ratio measure but the numerical consequence of inadmissible arithmetic. The same conclusion applies to the cost-effectiveness ratios and simulation outputs that subsequently depend upon QALYs as model inputs.

This is why the search for alternative measurement frameworks within HTA has been so unproductive. The problem is not that existing instruments require refinement. The problem is that they attempt to achieve measurement through arithmetic rather than through measurement itself. Utility algorithms, summed scores, preference weights and simulation models all share a common assumption: that numerical manipulation can compensate for the absence of lawful measurement. It cannot.

Once the distinction between manifest and latent attributes is recognized, the situation becomes much clearer. Manifest attributes require direct observation and linear ratio measurement. Latent attributes require Rasch measurement and logit ratio scales. Every outcome relevant to therapy assessment falls into one of these two categories. There is no intermediate class of attribute requiring a separate measurement framework. There is no need for utility construction, preference algorithms or composite indices. The apparent diversity of measurement approaches in HTA is largely an illusion created by the proliferation of scoring systems that generate numbers without generating measures.

This conclusion explains the emergence of measurement inversion and arithmetic chaos throughout contemporary HTA. The discipline has attempted to create quantitative claims without first establishing lawful measurement. The result has been a proliferation of numerical constructs that are routinely treated as measures despite lacking the properties required to support that interpretation. The solution is not further methodological complexity. The solution is to return to the two measurement frameworks that satisfy the requirements of representational measurement. Linear ratio measures and Rasch logit ratio measures are not simply preferable alternatives. They are the only lawful measures available for the assessment of therapy impact.

THE END OF ARITHMETIC CHAOS

Arithmetic chaos is not an inevitable feature of health technology assessment. It is the consequence of abandoning a simple scientific principle: measurement must precede arithmetic. Once this principle is ignored, any numerical construct can be manipulated as though it were a measure. Health-state descriptions become utilities, utilities become QALYs, QALYs become simulation outputs and simulation outputs become cost-effectiveness claims. At every stage arithmetic operations are undertaken without establishing whether the entities involved possess the measurement properties necessary to support those operations. The result is a framework in which numerical sophistication substitutes for measurement.

The consequences of this inversion have been profound. For more than four decades HTA has devoted enormous intellectual effort to the refinement of utilities, preference algorithms, simulation models and cost-effectiveness frameworks. Yet each of these developments rests upon the assumption that lawful measurement has already been achieved. The evidence suggests otherwise. The widespread acceptance of utilities and QALYs has obscured a more fundamental problem: the absence of measures capable of supporting the arithmetic operations imposed upon them.

Recognition that there are only two lawful measures immediately changes the discussion. Once therapy outcomes are classified as either manifest or latent attributes, the measurement requirements become clear. Manifest attributes require linear ratio measures. Latent attributes require Rasch logit ratio measures. The distinction eliminates the need for utility construction, preference weighting systems and composite indices. Therapy impact can be assessed directly through attributes that possess demonstrable measurement properties.

The implications for HTA are transformative. Formulary submissions no longer require hypothetical future worlds populated by utility scores and simulated QALYs. Instead, manufacturers present claims based upon measurable attributes supported by explicit protocols. These claims can be evaluated empirically, replicated by independent investigators and subjected to falsification. The focus shifts from the generation of imaginary estimates to the assessment of real-world evidence.

The same transformation applies to disease-area reviews and therapeutic-class evaluations. Rather than comparing therapies through cost-per-QALY estimates derived from simulation models, comparisons can be based upon measurable outcomes relevant to patients, clinicians and health systems. Hospital admissions, treatment persistence, resource utilization, symptom burden,

physical functioning and need fulfilment can all be assessed through lawful measurement frameworks. The resulting evidence is transparent, evaluable and directly relevant to decision making.

Most importantly, the recognition of only two lawful measures restores the distinction between numbers and measures. Numerical outputs are no longer accepted simply because they are produced by sophisticated algorithms. Every claim must demonstrate the measurement properties of the attribute involved. Arithmetic becomes subordinate to measurement rather than a substitute for it.

This is why reconstruction is both possible and necessary. The problem confronting HTA is not a lack of analytical sophistication. The problem is that analytical sophistication has been allowed to replace measurement. Once the discipline returns to linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes, the foundations of arithmetic chaos disappear. Measurement inversion is reversed, lawful measurement is restored and HTA can once again function as a quantitative science. The path forward is therefore neither complex nor mysterious. It begins with a simple recognition: there are only two lawful measures, and everything else is arithmetic.

IMPLICATIONS FOR THE RECONSTRUCTION OF HTA

Recognition that there are only two lawful measures has profound implications for the future of health technology assessment. The issue is not merely one of measurement theory. It affects every stage of therapy evaluation, from clinical development and formulary submissions to disease-area reviews, therapeutic-class assessments and professional education. Once the distinction between manifest and latent attributes is restored, reconstruction becomes a practical process rather than an abstract methodological debate.

The first implication concerns formulary submissions. The reference-case paradigm assumes that the central objective of a submission is the construction of a cost-effectiveness model supported by utilities, QALYs and simulation outputs. Reconstruction reverses this emphasis. The purpose of a submission becomes the presentation of empirically evaluable claims regarding therapy impact. These claims must be expressed either as linear ratio measures for manifest attributes or Rasch logit ratio measures for latent attributes. Each claim must be supported by an explicit protocol describing the target population, comparator, attribute definition, measurement framework, evaluation period, success criteria and replication strategy. The focus shifts from hypothetical future projections to measurable outcomes.

The second implication concerns clinical development. Manufacturers can no longer regard formulary submissions as a separate activity undertaken after regulatory approval has been secured. The development of attribute claims must begin during clinical development. Manifest and latent outcomes should be identified during Phase II and Phase III studies, with measurement strategies embedded in trial design. There can be no discontinuity between the evidence generated for regulatory approval and the evidence required for therapy assessment. Both should be part of a single measurement strategy directed toward evaluable claims.

The third implication concerns disease-area and therapeutic-class reviews. Under the reference-case framework, these reviews are often dominated by simulation outputs, utility estimates and cost-per-QALY comparisons. Reconstruction replaces these constructs with measurable attributes. Therapies are compared according to their demonstrated impact on manifest and latent outcomes. The result is a framework that is transparent, reproducible and directly linked to patient experience and health-system objectives. Claims can be tested, challenged and revised as new evidence emerges.

The fourth implication concerns professional education. For more than four decades, HTA education has focused on utilities, QALYs, simulation modelling and economic evaluation. Reconstruction requires a different set of skills. Researchers, policy analysts, clinicians and pharmacy faculty must become familiar with representational measurement, dimensional homogeneity, Rasch measurement, protocol design, empirical evaluation and falsification. The emphasis moves from model construction to measurement construction. Understanding how to create lawful measures becomes more important than understanding how to manipulate numerical outputs.

Finally, reconstruction changes the purpose of HTA itself. The objective is no longer to generate estimates of value through increasingly elaborate simulations. The objective is to support credible claims regarding therapy impact that can be evaluated in real-world populations. Measurement regains its primary role. Arithmetic becomes subordinate to measurement rather than a substitute for it. The consequence is a framework grounded in observation, replication and falsification rather than hypothetical projections and non-evaluable claims.

Once it is recognized that there are only two lawful measures, the path forward becomes surprisingly clear. Linear ratio measures provide the framework for manifest attributes. Rasch logit ratio measures provide the framework for latent attributes. Together they offer the foundation for a reconstructed HTA capable of supporting the standards of quantitative science. The challenge is not conceptual but institutional. It requires a willingness to move beyond the reference case and to embrace a measurement-based paradigm for the assessment of therapy impact.

CONCLUSION

The central argument of this paper is straightforward. For the assessment of therapy impact there are only two lawful forms of measurement: linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes. This conclusion follows directly from the requirements of representational measurement and provides a clear foundation for quantitative claims regarding health outcomes. Once these two measurement frameworks are recognized, much of the complexity that has come to characterize contemporary HTA is revealed as unnecessary.

The history of the reference-case paradigm can be interpreted as a gradual movement away from measurement and toward arithmetic. Utilities, QALYs, cost-effectiveness ratios and simulation models were developed in the belief that increasingly sophisticated numerical methods would provide a stronger basis for decision making. Yet sophistication cannot compensate for the absence of lawful measurement. Arithmetic operations applied to entities whose measurement properties have not been established cannot create quantitative meaning. The result is measurement inversion,

where arithmetic precedes measurement, and arithmetic chaos, where numerical constructs are treated as though they were measures despite the absence of supporting evidence.

The consequences have been substantial. For more than four decades HTA has relied upon utilities derived from health-state descriptions, preference valuations and scoring algorithms. These utilities have been transformed into QALYs, incorporated into simulation models and used to support reimbursement decisions affecting millions of patients. Yet the legitimacy of the entire framework depends upon assumptions concerning measurement that remain unresolved. The problem is not the quality of the arithmetic. The problem is that arithmetic has been allowed to replace measurement.

Recognition that there are only two lawful measures provides a way forward. Manifest attributes can be assessed through linear ratio measures that support direct observation, replication and falsification. Latent attributes can be assessed through Rasch logit ratio measures that satisfy the requirements of measurement for unobservable traits. Together these frameworks provide the basis for empirically evaluable claims regarding therapy impact without recourse to utilities, QALYs or imaginary future worlds generated by simulation models.

The challenge facing HTA is therefore not one of refinement but reconstruction. The objective is not to improve the reference case but to replace it with a framework grounded in lawful measurement. Once measurement is restored to its proper place, arithmetic becomes a tool rather than a substitute for scientific inquiry. The future of HTA depends upon a return to first principles: measurement precedes arithmetic, claims must be evaluable and falsifiable, and quantitative science begins not with numbers but with measures. There are only two lawful measures. Everything else is arithmetic.

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

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

[iii] Maimon Research https://maimonresearch.com/ai-llm-rest-of-world/#australia

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