MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

LEGACY HTA: IS THIS THE WORST WE COULD HAVE DONE?

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

LOGIT WORKING PAPER No 7987 AUGUST 2026

www.maimonresearch.com

Tucson AZ

BACKGROUND

In discussions with colleagues in health economics and technology assessment, I have attempted on a number of occasions to debate the requirements of lawful measurement for the quantitative evaluation of therapy claims. The principal obstacle has been an almost universal unfamiliarity with the scales of measurement and the axioms of representational measurement; not to mention the standards of normal science. My own training included these requirements, including the critical place of ratio measurement in determining the admissibility of arithmetic operations. I therefore found this apparent vacuum in HTA training difficult to understand, particularly in a discipline that presents itself as presumptively quantitative.

The immediate catalyst for this brief paper was a recent discussion in which the defense of legacy HTA was reduced to the familiar proposition: ***"This is the best we can do."*** The comment was offered not as a defense of the measurement properties of utilities, QALYs or cost-effectiveness ratios, but apparently as a reason for accepting the established framework despite its deficiencies. My immediate response was: ***"Are you sure it is not the worst that could have been done?"***

That exchange raises a more serious question than the rhetoric might suggest. If a presumptively quantitative discipline constructs its analytical framework without first establishing the measurement properties of the quantities it manipulates, then longevity, widespread adoption and methodological sophistication cannot establish scientific legitimacy. The issue is not whether legacy HTA is imperfect; all empirical science is imperfect. It is whether the arithmetic at the center of legacy HTA was ever admissible in the first place. This paper argues that the defense that ***"This is the best we can do"*** fails because it ignores what must first be demonstrated: that the numbers entering the analysis are measures and possess the properties required for the arithmetic operations performed upon them. Legacy HTA is therefore better characterized as a presumptive quantitative discipline: it assumes measurement in order to perform arithmetic rather than establishing measurement before arithmetic. To achieve this, it ignores all required standards for lawful measurement

INTRODUCTION

Legacy health technology assessment (HTA) is commonly defended on pragmatic grounds. Utilities, QALYs, ICERs and assumption-driven reference-case models may be imperfect, but healthcare decisions must be made and, it is argued, this is the best we can do. If so, it is a poor effort. The defense is false. Legacy HTA is not the best we could have done. It is the worst we could possibly have done**.**

The failure is not that legacy HTA operates under uncertainty. All empirical science does. The failure is that an entire presumptive quantitative discipline was constructed without first establishing lawful measurement: the axioms of representational measurement and scales of measurement [1 2 3] . Multidimensional health-state descriptions were converted into numbers, those numbers were assigned the mathematical properties required to create QALYs, QALYs became denominators for ICERs, and assumption-driven models projected the resulting constructs into imaginary futures.

Measurement did not determine the arithmetic. The arithmetic demanded the numbers and the properties those numbers were required to possess. The fact that this numeric manipulation denied the axioms of representational measurement and the arithmetic implications of the scales of measurement was never a recognized concern.

Worse followed. Universities taught the framework, journals published it, professional organizations promoted it and HTA agencies required it. Juggling numbers became institutional orthodoxy, transmitted to a global audience who had no defense.

The most serious consequence of legacy HTA is educational. For decades, universities have trained successive generations of students to manipulate numbers without first establishing whether the attributes in question have been lawfully measured or whether the arithmetic performed on them is admissible. Numerical practice has been taught as quantitative science without satisfying the prior requirements of measurement. More disturbing still, there appears to be no recognized duty of care requiring faculty and institutions to distinguish substantive quantitative science from numerical nonsense before transmitting it to students. The result is not simply methodological failure in contemporary HTA practice, but its systematic reproduction under the authority and prestige of graduate education. Students are entitled to expect that methods presented as quantitative science satisfy the elementary requirements of measurement. Where that standard is neither recognized nor enforced, institutional authority risks becoming a mechanism for perpetuating error rather than correcting it.

1. THE NUMERICAL CHARADE

Legacy HTA begins not with measurement of a defined attribute but with the valuation of health states. Through procedures such as the time trade-off or standard gamble, descriptions of health states are assigned numerical values intended to represent preferences for those states. This starting point effectively severs the subsequent analysis from the axioms of representational measurement and the scales of measurement for allowable arithmetic. The prior scientific questions have not been answered: what is the attribute to be measured; is it unidimensional; is it manifest or latent; what empirical properties does it possess; and what measurement structure is required to represent it quantitatively? Without those foundations, there is no demonstrated measure to which arithmetic operations can be applied.,

Legacy HTA instead proceeds in the opposite direction. Numbers are juggled because there is a required construct to be justified. A numerical value is required to transform a health-state description into a utility; the utility is required to multiply time to create the QALY; QALYs are differenced to produce an incremental outcome; and incremental QALYs are required as the denominator of the cost-per-QALY ICER. At each stage, the numerical construct is assumed to possess whatever mathematical properties are required by the next arithmetic operation.

The required operation determines the presumed properties of the number rather than demonstrated measurement properties determining the admissible operation. Once that reversal has occurred, increasingly sophisticated modelling, simulation and uncertainty analysis cannot restore the missing measurement foundation. Legacy HTA becomes detached from the constraints governing lawful quantitative analysis: it does not move from measurement to arithmetic; it moves from required arithmetic to presumed measurement properties. The ratio-scale properties required to support its multiplicative and divisive arithmetic are neither demonstrated beforehand nor treated as constraints upon what may subsequently be calculated.

This is the defining feature of legacy HTA as a presumptive quantitative discipline. Every step on the path toward a reference-case claim for cost-effectiveness introduces another unforced error of numerical manipulation. A utility is required and therefore constructed; multiplication by time is required and therefore performed; QALY differences are required and therefore calculated; an incremental ratio is required and therefore created. At no stage does the required arithmetic derive from measurement properties established beforehand. The remarkable result is the global acceptance of an analytical framework whose arithmetic has persisted largely because the measurement standards required to challenge it were absent from the knowledge base within which analysts were educated and subsequently practiced.

The program-content evidence provides perhaps the clearest demonstration of this condition. Across the knowledge bases interrogated, the recurring empirical pattern is not merely weak representation of measurement theory but the absence of representational measurement and the scales of measurement as the organizing scientific foundation for quantitative analysis, alongside extensive or complete representation of the methods of legacy HTA. Faculty teach and students encounter utilities, QALYs, cost-effectiveness analysis, ICERs, decision modelling, simulation and uncertainty analysis without the prior measurement framework required to determine whether the underlying arithmetic is admissible. The result is a self-reproducing epistemic environment in which the assumptions of legacy HTA appear legitimate because the scientific framework capable of exposing their measurement failure is effectively absent.

This is what makes the situation more than an historical methodological mistake. Legacy HTA has become an epistemically closed quantitative system: its assumptions generate the numbers, its required arithmetic determines the properties those numbers are presumed to possess, and its models elaborate the resulting claims, while the external measurement standards capable of challenging the entire sequence remain outside the framework. What appears from within as an increasingly sophisticated quantitative discipline appears from the standpoint of representational measurement as something very different: a presumptive quantitative discipline sustained by a sequence of unforced arithmetic errors. the system contains everything necessary to reproduce itself but not the knowledge necessary to challenge its foundations.

Legacy HTA has a unique place as a discipline. It is constructed from a process of numerical manipulation that would not be recognized, let alone permitted, in any other quantitative discipline. It is unique, but in the worst possible way: it has succeeded in institutionalizing practices that the elementary requirements of measurement should have prevented from occurring in the first place. The remarkable feature is not simply that errors were made, but that they became the accepted methodology of the discipline and were subsequently embedded in guidelines, curricula, journals and decision processes. Legacy HTA may display the architecture of normal science, but it denies a fundamental reality of quantitative science: arithmetic is constrained by measurement;

measurement is not created by arithmetic**.** In that sense, its apparent quantitative sophistication conceals rather than corrects its scientific failure

2. WHERE IS THE EVIDENCE FOR MEASUREMENT?

The defense that legacy HTA represents “the best we can do” would at least be intelligible if its methods had emerged from sustained engagement with measurement theory and survived systematic assessment against the requirements of representational measurement. The empirical evidence demonstrates the opposite. Across interrogations of HTA knowledge bases, curricula and university program content, the recurring finding is not that representational measurement was considered and rejected. It is that representational measurement is effectively absent while the methods of legacy HTA are comprehensively represented.

It is only now, with the emergence of large-scale knowledge bases and their systematic interrogation, that the scale of the legacy HTA disaster in measurement terms has become visible**.** More than 200 interrogations by Maimon Research have exposed not merely individual methodological errors but a remarkably consistent pattern across agencies, universities, research centers, journals, curricula and national HTA environments [4]. The significance of this evidence lies in its breadth. What might once have been dismissed as an omission within a particular guideline, institution or curriculum can now be recognized as a characteristic of the legacy HTA knowledge base itself.

More disturbing is what these interrogations reveal about the transmission of the failure. Measurement inversion, curriculum inversion and program-content inversion are not independent phenomena. They are mutually reinforcing expressions of the same epistemic failure**.** Measurement inversion establishes the intellectual condition: the requirements of lawful measurement are absent or weakly recognized while the numerical conventions of legacy HTA are accepted. Curriculum inversion reproduces that condition educationally: students encounter utilities, QALYs, ICERs, modelling and simulation without first encountering the measurement principles required to determine whether their arithmetic is admissible. Program-content inversion then demonstrates the institutional result: sophisticated university and research environments can comprehensively represent the techniques of legacy HTA while providing no explicit foundation in representational measurement.

The importance of the more than 200 interrogations is therefore cumulative**.** They transform the argument from a criticism of particular methods into evidence of a discipline-wide and global knowledge failure**.** The recurring pattern across different institutions and jurisdictions is difficult to explain as coincidence or local oversight. It indicates that generations of analysts have been trained within, and subsequently reproduced, a presumptive quantitative framework without access to the measurement standards necessary to recognize its foundational errors. What has now become visible is not simply that legacy HTA made mistakes. It constructed an international knowledge system capable of reproducing those mistakes while remaining largely unaware that they were mistakes. It survives because **:** the knowledge required to recognize the failure was itself missing from the knowledge base.

This acceptance of measurement inversion is not restricted to universities and their continuing reproduction of false measurement within the legacy HTA epistemic framework. The same pattern extends across the other institutions that create, disseminate and enforce the methodological conventions of the discipline. At the apex stands the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), whose global reach through methodological guidance, task forces, conferences, journals, educational activities and regional networks has made it one of the principal vehicles through which the legacy HTA framework has been transmitted internationally. Rather than providing a challenge to measurement inversion, the ISPOR knowledge base has helped normalize and reproduce the same presumptive quantitative framework. Yet that knowledge base shows a virtual absence of any understanding of the standards for lawful measurement

The significance of this influence becomes apparent when country and regional knowledge bases are interrogated. Despite substantial differences in healthcare systems, reimbursement arrangements and institutional structures, the same methodological architecture repeatedly emerges: utilities, QALYs, cost-effectiveness analysis, ICERs, preference-based outcomes, decision modelling, simulation and uncertainty analysis are extensively represented, while representational measurement and the constraints imposed by scales of measurement are absent or marginal. The international uniformity is striking. ISPOR communities across Europe, Asia-Pacific and the Americas effectively sing from the same methodological hymnbook. What is replicated globally is not a science of measurement but a common system of numerical manipulation in which the properties required for arithmetic are presumed rather than demonstrated.

This helps explain why measurement inversion has proved so resistant to challenge. The epistemic system is mutually reinforcing. Universities educate students in the accepted methods; professional organizations codify and disseminate them; journals publish research employing them; HTA agencies incorporate them into guidelines; consultants and manufacturers construct submissions around them; and subsequent generations of students encounter their widespread use as evidence that the framework must be scientifically legitimate. Each component validates the others. The result is a global legacy HTA epistemic memeplex capable of reproducing itself across institutions and jurisdictions while excluding from its core knowledge base the measurement principles that would expose its foundational failure.

Whether challenges grounded in the requirements of lawful measurement will have a significant impact on legacy HTA is a more difficult question. Too many people and institutions have too much invested in the existing framework to make rapid change likely. Acceptance of the measurement critique would not require a minor methodological adjustment or the addition of another technique to the HTA toolkit. It would require acknowledgement that much of the established utility–QALY–ICER architecture rests upon numerical operations whose measurement foundations were never demonstrated. The implications extend across academic careers, graduate curricula, textbooks, methodological guidelines, journal publications, professional organizations, consulting practices, manufacturer submissions and decades of agency decision-making.

The barriers are therefore institutional as well as intellectual. For universities, acceptance would require reconsideration of what has been taught to successive generations of students. For professional organizations, it would require reassessment of methodological standards promoted internationally. For journals, it would call into question an extensive published literature. For HTA agencies, it would challenge reference cases and decision frameworks that have acquired authority through decades of use. The cost of admitting measurement failure consequently increases with the very institutional success of the legacy framework.

This creates a powerful asymmetry. Rejecting the representational measurement challenge preserves careers, methods, institutions and established decision processes; accepting it requires substantial reconstruction. Yet none of these considerations constitutes a scientific defense of legacy HTA. Institutional inconvenience cannot make inadmissible arithmetic admissible. The process of deconstruction makes this clear. The relevant question is therefore not whether the transition to lawful measurement would be disruptive but how long a presumptive quantitative discipline can continue to defend established practice once the measurement failure underlying that practice has been made explicit. The question is not just would students enroll in a program that continues to endorse Legacy HTA but would the community of potential HTA services continue to pay attention.

Reviewing these results, one is reminded of a variation on Oscar Wilde's description of fox hunting. Legacy HTA has become the immeasurable in pursuit of the incalculable.

This empirical evidence destroys the claim that legacy HTA is the best we can do. A discipline cannot establish that its methods represent the best available quantitative approach when the knowledge base within which that judgment is made excludes the scientific principles required to determine what can legitimately be measured and calculated. Legacy HTA did not examine lawful measurement, discover that its requirements were impractical, and reluctantly select utilities and QALYs as second-best approximations. Lawful measurement was never established as the governing standard.

The application of lawful measurement begins with attributes. Manifest attributes such as survival time, hospital days or walking distance require linear ratio measurement where the necessary empirical properties can be established. Latent attributes such as pain, fatigue, physical function or need fulfilment require a fundamentally different measurement pathway, with Rasch measurement providing an invariant quantitative continuum. Therapy impact can then be expressed through attribute-specific, empirically evaluable claims. There is no scientific requirement to collapse heterogeneous therapy consequences into one number. Decision-makers may consider multiple legitimate claims together with costs, safety, unmet need, equity and other relevant considerations. Science measures. Decision-makers decide.

3. THE EPISTEMIC PRISON

Legacy HTA is the worst we could have done not simply because its central arithmetic is inadmissible. Scientific disciplines make mistakes. Errors are not the enemy of science; they are indispensable to its progress because they can be exposed, rejected and replaced. A visibly defective methodology invites criticism. Legacy HTA achieved something far more damaging. It succeeded.

Universities taught it. Textbooks codified it. Journals published it. Professional organizations endorsed it. HTA agencies required it. Consultants sold it. Manufacturers complied with it. Students were trained in it and built careers around it. What began as ignorance of lawful measurement became accepted methodological practice, and accepted methodological practice became institutional authority. Institutional acceptance transformed measurement ignorance into methodological orthodoxy.

The educational consequences are profound. Students can undertake graduate education in health economics, pharmacoeconomics, outcomes research and HTA and become highly proficient in health-state valuation, QALYs, ICERs, Markov models, simulation and probabilistic sensitivity analysis without being introduced to the axioms of representational measurement, admissible transformations, dimensional homogeneity, the constraints imposed by scales of measurement or the requirement for ratio properties before multiplication and division. They are taught how to manipulate numbers without being given the scientific knowledge required to determine whether those numbers can legitimately be manipulated.

The program-content findings therefore document something much more serious than a curricular omission. They describe an epistemic prison. Faculty teach what they themselves were taught. Students become researchers and future faculty. Journals expect analyses constructed within the accepted framework. Professional organizations reproduce its methodological standards. Agencies incorporate those standards into submission guidelines. Manufacturers and consultants must conform if their analyses are to be accepted. Successful conformity then reinforces the authority of the framework The system validates itself.

Forty years of use is not evidence of measurement. Agency endorsement is not evidence of measurement. Publication in prestigious journals is not evidence of measurement. Professional consensus is not evidence of measurement. The quantitative claim carries its own burden of proof. The epistemic prison blocks precisely this challenge. If students are never taught the measurement principles required to question the accepted framework, the framework appears unproblematic. Researchers refine models rather than interrogate their foundations. More sophisticated simulations, mapping algorithms and uncertainty analyses become signs of methodological progress while the prior measurement failure remains untouched. This prevents the evolution of objective knowledge. A scientific discipline advances when claims are exposed to empirical challenge, failure and replacement. An assumption-driven reference-case framework instead rewards refinement within the accepted paradigm. The fundamental propositions are protected by convention rather than repeatedly exposed to attempts at falsification.

If a demonstration were required of what happens when an analytical framework puts lawful measurement entirely to one side, it would be difficult to construct a more satisfactory example than the legacy HTA reference case. Indeed, it can be regarded as an almost ideal negative exemplar of quantitative science**.** Step by step, it demonstrates precisely why the axioms of representational measurement and the arithmetic constraints imposed by scales of measurement are indispensable. Begin without a defined unidimensional attribute; substitute the valuation of multidimensional health states for measurement; assign numerical utilities without first establishing the measurement properties required of them; presume the properties necessary to multiply those utilities by time; construct QALYs; subtract them to create incremental effects;

divide costs by the resulting numerical differences; and finally project the entire construction through an assumption-driven reference-case model. At every stage, the arithmetic is determined by what the framework requires rather than constrained by what measurement permits.

It would be difficult to devise a better demonstration of why lawful measurement must precede arithmetic. Almost every principle that representational measurement is intended to protect against is exposed by its absence: attributes are not established before quantification, scale properties do not determine admissible operations, multiplication and division proceed without the required measurement foundation, and increasingly elaborate mathematical procedures are applied to constructs whose quantitative status has been presumed rather than demonstrated. The irony is therefore considerable. The reference case, intended as the methodological standard of legacy HTA, provides perhaps its most compelling counterexample. It is an exemplar not of how quantitative analysis should be undertaken, but of why the rules of lawful measurement exist in the first place.

4. THE RETURN TO SCHOLASTICISM

Legacy HTA has resurrected the scholastic tradition**.** The scientific revolution of the seventeenth century rejected knowledge grounded primarily in received authority and internally coherent systems of reasoning insulated from empirical challenge. The Royal Society expressed the new standard through its motto, Nullius in verba: take nobody's word for it. Claims about the world are to answer to the world through observation, experiment and measurement rather than through the authority of established doctrine [5].

Legacy HTA has reversed that achievement. Utilities are accepted because an established literature accepts them. QALYs are accepted because HTA agencies require them. ICERs are accepted because methodological guidelines specify them. Reference cases are accepted because professional organizations and reimbursement authorities endorse them. Universities teach these methods because they constitute accepted professional practice, while their status as accepted professional practice is reinforced because universities teach them. Authority validates authority.

Contemporary HTA surrounds this authority structure with an extraordinary apparatus of twenty-first-century ersatz quantitative sophistication. Models contain hundreds of parameters. Clinical pathways are represented by decision trees and Markov structures. Monte Carlo simulation generates thousands of iterations. Probability distributions are assigned to uncertain inputs. Sensitivity analyses explore alternative assumptions. Powerful software calculates results to several decimal places. None of this establishes measurement.

A thousand calculations performed upon a number do not transform that number into a measure. A million simulations cannot create a unidimensional attribute from a multidimensional health-state description. A probability distribution cannot confer ratio properties upon a utility. Mathematical manipulation cannot make an inadmissible operation admissible. An assumption about an imaginary future does not become an empirical observation because sophisticated software processes it.

The consequence is a remarkable inversion of the scientific tradition. Nullius in verba becomes, in practice, *take the framework's word for it*. Accept that a time trade-off produces a unidimensional score with ratio properties. Accept that the utility possesses the properties required for multiplication by time. Accept that the resulting product is a QALY. Accept that QALYs can be differenced. Accept that incremental costs can be divided by incremental QALYs. Accept that assumptions extending decades into an unknowable future can be assembled into a reference case whose output constitutes quantitative evidence for a decision made today. The appearance of precision conceals the absence of measurement. Avoid any thought of deconstruction.

The route out requires neither another economic model nor another refinement of the reference case. It requires restoration of the scientific sequence that legacy HTA abandoned almost 50 years ago: science demands measurement, not numbers. Numerical assignment is not measurement, and increasingly sophisticated manipulation of numbers cannot compensate for the absence of measurement at the outset. The starting point must therefore be the attribute whose therapy impact is claimed. That attribute must be defined, demonstrated to be unidimensional and identified as either manifest or latent. Its properties determine the measurement pathway; they are not determined retrospectively by the arithmetic an analyst wishes to perform.

For manifest attributes, the requirement is linear ratio measurement, with an invariant unit and non-arbitrary zero supporting the arithmetic appropriate to the claim. For latent attributes, the requirement is Rasch measurement, where person-item conjoint measurement provides an invariant quantitative continuum for evaluating possession of the attribute and change in that possession. These are not alternative modelling techniques to be added to the legacy HTA toolkit. They represent the two measurement pathways from which defensible quantitative therapy-impact claims can be constructed.

Scholasticism asks how the inherited framework can be elaborated, refined and defended. Science asks a different question: what can actually be measured? Once that question becomes the starting point, the intellectual burden changes completely. Analysts no longer have to defend the inherited numerical architecture of legacy HTA. They have to identify attributes, construct measures and make claims that can encounter the empirical world.

The transition is therefore not from an old reference case to a new reference case. There is no replacement reference case because none is required. The transition is from presumptive yet false quantification to lawful measurement; from assumption-driven modelled claims to empirically evaluable claims; from numerical manipulation to measurement-constrained arithmetic; and from an epistemically closed methodology to the ordinary requirements of normal science. The intellectual escape from legacy HTA is, in this respect, remarkably simple: stop asking what numbers the model requires and start asking what attributes can be measured.

5. THE TRANSITION TO LAWFUL MEASUREMENT

The transition to lawful measurement in HTA is surprisingly straightforward. It does not require another generation of increasingly elaborate economic models, further refinement of the reference case or additional techniques for manipulating uncertainty. It requires restoration of the scientific sequence that legacy HTA reversed. Recognize that measurement precedes arithmetic.

The starting point is the attribute and its properties: what is to be measured, is it unidimensional, and is it manifest or latent? Manifest attributes require a defensible linear ratio measure with a non-arbitrary zero and invariant unit; latent attributes require construction of an invariant Rasch measurement continuum. Once measurement has been established, the admissible arithmetic follows from the properties of the scale rather than being presumed because a particular calculation is desired. Therapy-impact claims can then be formulated prospectively as empirically evaluable, replicable and falsifiable claims. The transition is therefore not from one version of HTA modelling to another. It is from numbers to measurement, from presumed quantitative properties to demonstrated ones, and from numerical storytelling to the standards expected of normal quantitative science.

The steps required have been detailed in the Maimon Research nine-unit transformation program [6]. This program provides a structured template for faculty, researchers and practitioners seeking to understand and implement the transition from the measurement failure of legacy HTA to a framework grounded in representational measurement. The program comprises 45,000 words of structured instructional material with the nine units moving progressively from measurement preceding arithmetic through attributes, scales of measurement, admissible arithmetic and dimensional homogeneity to the distinction between manifest and latent attributes, linear ratio measurement, Rasch measurement, attribute possession and the construction of empirically evaluable therapy-impact claims. The program can therefore be used for individual faculty preparation, as a framework for reconsidering existing graduate teaching, or as the foundation for developing a new postgraduate course in HTA organized around the requirements of lawful measurement and normal science.

The material presented can be integrated into existing programs without requiring wholesale curriculum reconstruction. The essential change is one of scientific sequence. At the beginning of a course, students should be provided with the measurement principles required to interrogate every quantitative method that follows: what is the attribute to be measured; is it unidimensional; is it manifest or latent; what measurement scale has been established; what transformations and arithmetic are admissible; and does the proposed analysis preserve dimensional homogeneity? These questions can then be applied systematically as students encounter utilities, QALYs, reference case economic evaluation, patient-reported outcomes, decision models and the other established components of legacy HTA.

Representational measurement thereby becomes not another topic added to an already crowded curriculum but the organizing scientific framework through which the existing curriculum is critically assessed. Importantly, this places students in a position to judge the merits of the alternatives for themselves: whether there is a defensible scientific basis for continuing to support the legacy HTA framework, or whether quantitative therapy evaluation should instead be reconstructed around lawful measurement, admissible arithmetic, evaluable claims, replication and falsification; the standards expected of normal science. The educational responsibility is therefore not to prescribe the conclusion students must reach, but to ensure that they possess the scientific knowledge required to make that judgment.

An important feature of the program is that participants are not left simply to work through the material independently. Those enrolled have access to continuing 24/7 on line Q&A support from

Maimon Research as they study the units and, importantly, as they begin to apply the framework in teaching, research or therapy evaluation**.** Questions arising from particular attributes, proposed measures, Rasch applications, therapy-impact claims, protocols or curriculum development can be raised as they occur. The purpose is to support the transition from understanding the principles to applying them in practice. This continuing interaction is particularly important for faculty whose previous education did not include representational measurement and who may be reconsidering assumptions embedded in many years of teaching and research. The objective is not simply completion of nine units, but development of the capability to use the measurement framework independently and to introduce it confidently to students and colleagues.

The program can also provide a starting point for broader institutional engagement. Where universities, research centers, academic departments or professional groups wish to explore the implications of the new HTA in greater depth, additional activities can be developed around their particular requirements, including seminars, faculty discussions, workshops, small-group sessions, curriculum-development support and presentations addressing specific measurement or therapy-evaluation problems. These activities allow institutions to move beyond individual enrollment toward a wider reconsideration of how HTA is taught and applied.

A fee is charged for access to the nine-unit program to support the maintenance, administration, updating and continuing availability of the system, including participant support. The larger objective, however, is transformation: to provide faculty and institutions with an accessible, supported pathway for replacing the legacy of measurement failure with graduate education grounded in lawful measurement, admissible arithmetic and empirically evaluable claims.

The nine-unit program is therefore offered as a template for faculty preparation and the development of a post-legacy graduate HTA program. Universities may use the framework to reconsider existing teaching or as the foundation for constructing a new graduate course organized explicitly around representational measurement.

CONCLUSION

For those who wish to escape the comfort of scholastic legacy HTA, the opportunity is open. There is no need to repair, modify or attempt to resurrect the assumption driven reference case. It is entirely redundant**.** Once lawful measurement is accepted as the foundation for quantitative therapy evaluation, the elaborate utility–QALY–ICER superstructure simply becomes unnecessary.

The escape requires recognition of two measurement pathways. The first **is** linear ratio measurement for manifest attributes: attributes that can be directly observed and for which the required invariant unit and non-arbitrary zero can be established. The second is Rasch measurement for latent attributes, where person-item conjoint measurement provides the basis for constructing a unidimensional invariant continuum and quantitative claims concerning possession of the attribute. These two pathways provide the measurement foundation required for empirically evaluable claims of therapy impact.

The change is fundamental but conceptually straightforward. Begin with the attribute, not the model. Determine whether the attribute is manifest or latent. Establish the measurement requirements appropriate to it. Construct the measure. Determine from its properties what arithmetic is admissible. Formulate a therapy-impact claim capable of empirical evaluation, replication and falsification. Measurement precedes arithmetic because measurement determines what arithmetic is possible.

Nothing requires the resulting claims to be collapsed into a universal health-state value. Nothing requires a utility. Nothing requires multiplication by time to manufacture a QALY. Nothing requires incremental QALYs or their use as the denominator of a cost-effectiveness ratio. And nothing requires an assumption-driven reference case projecting these constructs into an imaginary future. Those requirements are imposed by legacy HTA upon itself; they are not requirements of measurement science.

This is why the familiar defense that **"this is the best we can do"** ultimately fails. Legacy HTA created the problem and then declared the machinery required to sustain that problem indispensable. Once the requirement for lawful measurement is restored, the supposed necessity of that machinery disappears with it.

The alternative is not another reference case. There should be no successor reference case**.** The objective is not to replace one synthetic model of value with another. It is to replace presumptive quantification with measurement and assumption-driven claims with claims that can be tested against the world.

After decades of numerical manipulation, the route out is remarkably uncomplicated: identify the attribute, measure it lawfully, respect the arithmetic permitted by the measure, and test the claim. For a discipline accustomed to the elaborate scholasticism of legacy HTA, that may appear a radical departure. For normal quantitative science, it is simply where the analysis should have started.

Recognizing the analytical failure of the reference case, and accepting that it may represent one of the worst possible analytical frameworks that could have been devised to support claims for comparative cost-effectiveness, will be a difficult step for a discipline that has invested decades in its development, refinement and institutionalization. Yet there is no scientifically defensible alternative. The problem cannot be resolved by modifying the reference case, improving its assumptions, adding further sensitivity analyses or constructing more sophisticated models. These are downstream refinements of a framework whose failure occurs upstream, before the arithmetic begins. Once the quantities entering the analysis have not been established through lawful measurement, subsequent mathematical sophistication cannot restore what was absent at the outset.

The measurement cat is now out of the legacy HTA kitty-litter box. This may create an uncomfortable transition. After all, the purpose of the kitty-litter box is precisely to ensure that the cat does not have to go outside. Legacy HTA has enjoyed much the same protection. Its self-contained methodological environment of utilities, QALYs, ICERs, reference cases, modelling and simulation has allowed successive generations of analysts to work entirely within the accepted framework without encountering the external requirements of representational measurement. The program-content interrogations demonstrate just how effective that insulation has been. But once the cat is outside, the boundaries of the box become apparent. The challenge now is not to put the cat back.

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] Krantz DH, Luce RD, Suppes P, Tversky A. *Foundations of Measurement. Volume I: Additive and Polynomial Representations.* New York: Academic Press; 1971

[2] Stevens SS. On the Theory of Scales of Measurement. *Science*. 1946;103(2684):677–680

[3] Bond T, Zi Yan, Heene M. Applying the Rasch Model: Fundamental measurement in the human sciences. (4th Ed). New York: Routledge, 2021

[4] Maimon Research. Website www.maimonresearch.com

[5] Wootton D. The Invention of Science: A new history of the scientific revolution. New York Harper Collins, 2015

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