MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

LEGACY HTA: ABANDONING ISPOR FOR LAWFUL MEASUREMENT

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

LOGIT WORKING PAPER No 2002 SEPTEMBER 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

Health technology assessment (HTA) has developed over four decades around utilities, quality-adjusted life years (QALYs), incremental cost-effectiveness ratios (ICERs), patient-reported outcomes and reference-case decision models. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) has played a major role in endorsing, publishing, teaching and internationally disseminating this framework. This paper asks whether the resulting methodological architecture satisfies the prior requirements of representational measurement and whether ISPOR can retain scientific authority if it does not.

Three complementary interrogations identify systematic measurement inversion. The ISPOR Good Practices knowledge base strongly possesses the arithmetic and conventions of legacy HTA while weakly possessing the measurement requirements governing their admissibility. Interrogation of the Value in Health knowledge base produces essentially the same profile. Assessment of publicly available ISPOR educational program content provides the categorical result of 0/17 representational-measurement topics compared with 20/20 contemporary HEOR and HTA topics. Extension of these interrogations across ISPOR regional environments identifies substantially the same structure internationally. What has been replicated globally is the methodology, not the measurement proof.

The failure is foundational. Measurement must precede arithmetic. Multiplication and division require appropriate measurement properties, yet ratio measurement is effectively absent from the ISPOR knowledge environments examined. Without demonstration that utility possesses the properties required for multiplication by time, the QALY is inadmissible; without a legitimate QALY, incremental QALYs and cost-per-QALY ICERs cannot acquire legitimacy through further arithmetic. Modelling, uncertainty analysis, professional consensus and reference-case endorsement cannot repair a failure occurring before arithmetic begins.

The implications for ISPOR are existential. The standards at issue are not new: the foundations of empirical science preceded ISPOR by centuries, scales of measurement were formalized in 1946 and representational measurement systematically developed by 1971. ISPOR, founded in 1995, has nevertheless spent approximately three decades institutionalizing and disseminating a framework in which these requirements failed to function as scientific gatekeepers. A Damascene conversion requiring abandonment of substantial accumulated methodological and professional capital is therefore improbable.

No such conversion is required for reconstruction to proceed. There is no central authority determining the standards of normal science and no need for a replacement global methodological superstructure. The challenge passes directly to agencies, universities and research groups. The transition is straightforward: define the attribute, establish lawful measurement, restrict arithmetic to operations permitted by the scale, and construct empirically evaluable, replicable and falsifiable therapy-impact claims. The objective is not to replace one orthodoxy with another, but to replace authority by demonstration.

INTRODUCTION

Health technology assessment has developed over four decades around a familiar quantitative architecture: utilities, quality-adjusted life years (QALYs), cost-effectiveness and cost-utility analysis, incremental cost-effectiveness ratios (ICERs), patient-reported outcomes and reference-case decision models. Few organizations have contributed more to the international development, standardization and dissemination of this framework than the International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Through Good Practices Task Forces, *Value in Health*, educational programs, conferences and professional networks, ISPOR has helped define what constitutes accepted methodological practice across health economics and outcomes research (HEOR) and HTA.

The difficulty is that quantitative science does not begin with arithmetic. It begins with measurement. Before numbers can legitimately be multiplied, divided, aggregated or otherwise manipulated, the attribute must be defined, measurement established, the properties of the resulting scale demonstrated and the proposed arithmetic shown to be admissible. These requirements are not optional methodological preferences. They follow from scales of measurement and representational measurement [i] [ii] [iii] [iv] . There is no separate standard under which HTA arithmetic becomes scientifically admissible because it is useful for decision making, supported by expert consensus or internationally accepted.

Recent interrogations of the ISPOR knowledge environment raise a fundamental challenge to this legacy. The ISPOR Good Practices knowledge base exhibits pronounced measurement inversion: strong possession of the arithmetic and methods of legacy HTA accompanied by weak possession of the measurement requirements that should govern them. Interrogation of the Value in Health knowledge base identifies essentially the same structure. A separate assessment of publicly available ISPOR educational program content produces an even more categorical result: 0/17 representational-measurement topics compared with 20/20 contemporary HEOR and HTA topics. The methods are endorsed, published and taught; the measurement gatekeeper is missing.

The implications extend far beyond ISPOR. Its success has helped embed this methodological framework across universities, manufacturers, consultancies, journals, HTA agencies and reimbursement systems internationally. Global adoption, however, cannot retrospectively establish measurement properties that were never demonstrated. What has been replicated is the methodology, not the measurement proof.

ISPOR therefore faces a problem created partly by the magnitude of its own success. If representational measurement requires abandonment of the measurement-inverted framework, ISPOR has a global methodological legacy to unravel. This may represent the end of empire for legacy HTA and ISPOR: not the end of quantitative therapy assessment, but the end of an international analytical order in which arithmetic has been permitted to precede measurement. The question addressed in this paper is whether ISPOR can survive that abandonment. There will be no Damascene conversion.

THE REQUIREMENTS OF LAWFUL MEASUREMENT

The problem confronting legacy HTA and ISPOR as its advocate begins with a requirement that applies throughout quantitative science: measurement must precede arithmetic. Numbers do not acquire scientific meaning merely because they have been assigned to observations, entered into equations or subjected to statistical analysis. Before arithmetic can legitimately be performed, the attribute of interest must be defined, a measurement system established, the properties of the resulting scale demonstrated and the proposed mathematical operations shown to be admissible for that scale. Arithmetic is therefore conditional upon measurement; it cannot retrospectively create the measurement properties its operations require.

The distinction between nominal, ordinal, interval and ratio scales is fundamental. Nominal scales classify observations and ordinal scales establish order. Interval scales additionally possess equal intervals but lack a non-arbitrary zero. Ratio scales possess equal units and a non-arbitrary zero and permit meaningful proportional comparisons. Their admissible transformations are similarity transformations of the form $x' = ax$, $a > 0$, under which ratios are preserved. Multiplication and division consequently require appropriate ratio measurement. The fact that a numerical representation contains a zero, has been statistically transformed or is described as cardinal does not establish ratio status. The required measurement structure must be demonstrated independently.

These requirements have immediate implications for HTA. The QALY is defined by the multiplication $QALY = U \times T$. Time can satisfy the requirements for linear ratio measurement. The critical question is whether the utility value possesses the measurement properties required to operate as a proportional adjustment to time. The multiplication itself cannot confer those properties. If utility has not been demonstrated to possess the required measurement structure, multiplying it by time produces a number but does not establish a scientifically admissible quantity. Subsequent aggregation of QALYs, subtraction to produce incremental QALYs and division of incremental costs by incremental QALYs cannot repair the original measurement failure.

Therapy-impact assessment requires two distinct measurement pathways. Manifest attributes, including survival time, hospital days and other directly observable quantities, require appropriate linear ratio measurement. Latent attributes such as pain, fatigue, physical function and need fulfilment cannot be directly observed and require construction of a measurement system capable of establishing possession of a defined unidimensional attribute. Rasch measurement provides this pathway through person-item conjoint measurement, invariance, specific objectivity and location on a logit continuum. Assigning numbers to questionnaire responses, summing item scores or statistically transforming those scores does not by itself create measurement.

Dimensional homogeneity imposes a further constraint. Quantities cannot be combined simply because numerical operations can be executed computationally. The dimensions and measurement properties of the quantities involved must support the operation proposed. Neither increasingly sophisticated modelling nor probabilistic sensitivity analysis can correct a dimensional or measurement failure occurring before the model was constructed.

These requirements leave no exemption for HTA. Decision usefulness, expert consensus, good practice guidelines to support measurement inversion, professional acceptance and international adoption cannot determine the admissibility of arithmetic. Nor can forty years of methodological practice transform a numerical representation into a measure. The scientific sequence remains unchanged: attribute, measurement, scale, admissible arithmetic and only then quantitative claims capable of empirical evaluation, replication and falsification. It is against this standard, rather than the conventions established within HEOR itself, that the ISPOR methodological, publication and educational knowledge environments must be judged.

ISPOR: MEASUREMENT INVERSION

The first indication of the scale of the problem is provided by interrogation of the ISPOR Good Practices knowledge base [v]. These reports occupy a distinctive position within HEOR because they do more than describe methods. They codify recommended practice and provide methodological standards intended to guide researchers, manufacturers, analysts and decision makers. The interrogation therefore asks whether the knowledge environment responsible for defining Good Practices possesses the measurement principles required to determine whether the arithmetic embedded in those practices is scientifically admissible.

The results identify pronounced measurement inversion. Recognition that measurement precedes arithmetic has an endorsement probability of only 0.10 (−2.20 logits). Recognition that multiplication requires ratio measurement is similarly 0.10 (−2.20), as is recognition that meeting the axioms of representational measurement is required for arithmetic. The two measurement pathways required for therapy-impact assessment are almost entirely absent. Recognition of linear ratio and Rasch logit ratio measurement as the relevant measurement classes is 0.05 (−2.50), while recognition of Rasch measurement as the basis for assessing therapy impact for latent attributes is also 0.05 (−2.50).

These results contrast sharply with the propositions supporting the established arithmetic of legacy HTA. The false proposition that the QALY is a ratio measure receives an endorsement probability of 0.95 (+2.50). The false proposition that QALYs can be aggregated is likewise 0.95 (+2.50). The false proposition that ratio measures can have negative values receives 0.90 (+2.20), while the proposition that the QALY is dimensionally homogeneous, also false, receives 0.85 (+1.75). Reference-case simulations are strongly represented as generating falsifiable claims despite their status as conditional model outputs, with the false proposition receiving 0.90 (+2.20).

The importance of these results lies in their configuration. The arithmetic is strongly possessed; the conditions that would make the arithmetic admissible are not. ISPOR Good Practices therefore do not simply exhibit incomplete knowledge of measurement theory. They represent an environment in which downstream methodological procedures have become established without the upstream measurement requirements functioning as effective scientific constraints. Expert consensus has repeatedly failed to function as a measurement safeguard.

Interrogation of the *Value in Health* knowledge base provides independent confirmation within a different component of the ISPOR environment. The profile is remarkably similar. Measurement precedes arithmetic receives 0.10 (−2.20), multiplication requires ratio measurement 0.10 (−2.20),

and the requirement that representational-measurement axioms govern arithmetic 0.10 (−2.20). Unidimensionality receives only 0.15 (−1.75), while recognition of Rasch measurement for latent therapy-impact claims is 0.05 (−2.50).

Against this weak measurement foundation, the journal knowledge base strongly possesses the propositions supporting legacy arithmetic. The false proposition that the QALY is a ratio measure receives 0.95 (+2.50); that QALYs can be aggregated, 0.95 (+2.50); that ratio measures can have negative values, 0.90 (+2.20); and that the QALY is dimensionally homogeneous, 0.85 (+1.75). The false proposition that summing Likert responses creates a ratio measure reaches 0.95 (+2.50). Recognition that cost-effectiveness claims fail the axioms of representational measurement is only 0.15 (−1.75).

The convergence is important because Good Practices and Value in Health perform different institutional functions. Good Practices codify and endorse methodological standards. Value in Health provides the flagship journal environment in which methods are developed, applied, reviewed and normalized. Yet both produce essentially the same epistemic structure. Measurement requirements occupy the strongly negative region while the arithmetic dependent upon their satisfaction occupies the strongly positive region.

This pattern cannot be defended by pointing to the apparent sophistication of the methods involved. More elaborate modelling, uncertainty analysis, preference elicitation or statistical validation operates downstream of the measurement question. Nor can repeated peer-reviewed publication establish the missing properties. Peer review can assess whether an accepted procedure has been performed correctly; it cannot make multiplication admissible when the quantities being multiplied lack the required ratio measurement properties.

The significance for ISPOR is therefore substantial. Measurement inversion is evident both where methodological standards are formally endorsed and where the resulting methods are published and normalized. This is not evidence that individual task-force members, editors, reviewers or authors knowingly rejected representational measurement. It is evidence of a systemic knowledge environment in which measurement failed to operate as the gatekeeper for arithmetic. The same methods could consequently be recommended, refined, published and repeatedly cited without the logically prior measurement demonstration ever being required.

The result is an internally reinforcing methodological system. Good Practices legitimize measurement inversion with methods represented in the journal literature; journal publication reinforces their standing as accepted science; subsequent Good Practices can then build upon an expanding literature generated within the same assumptions. Repetition gives the framework increasing institutional authority without supplying the measurement proof missing at its foundation. What has been replicated is the methodology, not the measurement proof.

The ISPOR evidence therefore identifies more than methodological error. It identifies the institutionalization of measurement inversion. Once that distinction is recognized, incremental improvement of the existing analytical framework is no longer an adequate response. If its foundational arithmetic cannot meet the requirements of lawful measurement, the framework dependent upon that arithmetic has to be abandoned.

ISPOR: EDUCATION

The measurement inversion identified in ISPOR Good Practices and *Value in Health* acquires greater significance when considered alongside a separate interrogation of ISPOR educational program content [vi]. This assessment addresses a different component of the institutional knowledge environment. Rather than asking which measurement propositions are possessed by the methodological or journal knowledge bases, it asks whether ISPOR education establishes the scientific requirements of measurement before introducing practitioners to the quantitative methods of contemporary HEOR and HTA.

The assessment applied the frozen 37-topic framework used for interrogation of national and regional HTA program-content knowledge bases. Seventeen topics represent the requirements of representational measurement and 20 represent contemporary HEOR and HTA. The result was categorical: 0/17 representational-measurement topics compared with 20/20 contemporary HEOR and HTA topics. The contrast represents complete program-content inversion.

The 20/20 result demonstrates the breadth of the ISPOR educational environment. Publicly available program content addresses economic evaluation, utilities, QALYs, cost-effectiveness and cost-utility analysis, patient-reported outcomes, preference measurement, decision analysis, simulation modelling, uncertainty analysis, value-of-information analysis, budget impact and value assessment. These are not peripheral activities. They constitute much of the methodological architecture through which contemporary HEOR and HTA are taught and practiced.

Against this comprehensive representation of downstream methods stands the 0/17 result. No explicit program-content foundation was identified for the representational-measurement topics that should determine whether the arithmetic underlying those methods is admissible. The missing foundation includes scales of measurement, admissible transformations, ratio-measurement requirements, dimensional homogeneity, the principle that measurement precedes arithmetic, and the distinction between the measurement requirements of manifest and latent attributes. The educational environment therefore reaches the arithmetic without first establishing the measurement conditions governing whether that arithmetic may legitimately be performed.

The two results cannot be interpreted as equivalent dimensions of curricular competence. The 17 measurement topics are not simply one group of subjects that happens to receive less attention than the 20 HEOR and HTA topics. They are logically prior. They constitute the scientific gatekeeper for the quantitative methods that follow. If measurement has not been established, the admissibility of the subsequent arithmetic remains unresolved. Failure at 0/17 therefore closes the gate to the 20/20 framework. Comprehensive instruction in downstream methodology cannot compensate for complete absence of the upstream requirements that determine whether those methods constitute lawful quantitative analysis.

This result also provides a mechanism through which the measurement inversion identified elsewhere in ISPOR can be reproduced. A researcher can become highly proficient in constructing decision models, estimating utilities, calculating QALYs and ICERs, undertaking probabilistic sensitivity analysis and applying sophisticated statistical techniques without ever being required to establish whether the quantities entering those operations possess the necessary measurement

properties. Technical proficiency and measurement competence can therefore become separated. The practitioner learns how to perform the arithmetic without first learning the scientific conditions under which that arithmetic is admissible.

The institutional cycle is then self-reinforcing. Educational programs transmit the established methods. Researchers apply them. *Value in Health* and other journals publish analyses undertaken within the accepted framework. Good Practices codify and endorse methodological procedures supported by that literature. Those Good Practices then reinforce educational content and professional expectations. Each part of the system provides legitimacy to the others while the missing measurement requirement remains outside the cycle. Successive generations can consequently inherit measurement inversion without ever making an explicit decision to reject measurement.

This is why the 0/17 result cannot be addressed by adding representational measurement to the existing educational framework as another competency. Representational measurement is not another specialist technique alongside economic evaluation, preference elicitation, modelling or value assessment. It determines whether the arithmetic employed within those techniques is scientifically admissible. It cannot simply become topic 38 while the existing 20/20 structure remains intact. The topics are incompatible

Acceptance of measurement as the gatekeeper instead reverses the educational sequence. The starting point becomes the attribute whose magnitude or possession is to be assessed. Its measurement requirements must then be established, the appropriate measurement pathway identified, the resulting scale demonstrated and the admissible arithmetic determined. Manifest attributes require appropriate linear ratio measurement; latent attributes require Rasch measurement of possession of a defined unidimensional attribute. Only after these requirements have been satisfied can quantitative therapy-impact claims legitimately proceed.

The program-content interrogation therefore supplies the third component of an institutional pattern. ISPOR Good Practices endorse the methodology. *Value in Health* publishes and normalizes it. ISPOR education transmits it. Measurement inversion is represented and endorsed across all three functions. The problem is no longer simply that ISPOR has supported methods whose measurement foundations are defective. Its educational environment provides a mechanism through which the same methodological structure can be transmitted to subsequent generations of researchers and practitioners.

Abandonment of legacy HTA must therefore include abandonment of the educational architecture that reproduces it. ISPOR cannot withdraw scientific support from measurement-inverted methods while continuing to teach those methods within an arithmetic-first framework. The 0/17 versus 20/20 result leaves no room for describing the problem as a curriculum imbalance. It identifies complete program-content inversion. If ISPOR accepts that measurement must precede arithmetic, its educational structure has to be rebuilt from the foundations rather than supplemented at the margins.

ISPOR: A GLOBAL STRAITJACKET

The significance of measurement inversion within ISPOR extends far beyond its central methodological activities. ISPOR has developed an international infrastructure through which the methods of HEOR and HTA are promoted, taught, discussed and transferred across jurisdictions. Its regional structure includes the Latin America Consortium, Asia Consortium and Central and Eastern Europe (CEE) Consortium, together with the Africa Network and Arabic Network. In addition, ISPOR's extensive European conference, educational and chapter activities permit examination of a wider European knowledge environment separately from CEE. These environments encompass profoundly different health systems, academic traditions, economic circumstances and stages of HTA institutional development. They therefore provide an unusually powerful test of whether measurement inversion is peculiar to particular jurisdictions or represents a common feature of the methodological framework disseminated internationally through ISPOR.

Two independent forms of interrogation produce remarkably consistent evidence. The first applies the same 24 canonical statements used to interrogate the ISPOR Good Practices and Value in Health knowledge bases. Across Latin America, Asia, Africa, the Arabic region, CEE and Europe excluding CEE, the resulting profiles are essentially identical. The propositions required to establish lawful measurement are consistently weakly possessed [vii] [viii] [ix] . Measurement before arithmetic, unidimensionality, ratio measurement before multiplication, representational-measurement axioms, dimensional homogeneity and the measurement requirements for manifest and latent attributes occupy the negative end of the profiles. Conversely, propositions reflecting the accepted quantitative architecture of legacy HTA are strongly possessed. Utilities, QALYs, aggregation, cost-effectiveness and reference-case modelling are embedded within knowledge environments in which the measurement properties required to justify their arithmetic have not first been established.

The second interrogation is even more categorical. The 37-topic program-content framework distinguishes 17 topics representing the requirements of representational measurement from 20 representing contemporary HEOR and HTA. Across these regional environments the same program-content inversion emerges: 0/17 representational-measurement topics compared with 20/20 contemporary HEOR and HTA topics. The downstream methodological environment is comprehensive. Economic evaluation, cost-effectiveness and cost-utility analysis, health-state utilities, QALYs, patient-reported outcomes, decision analysis, modelling, uncertainty analysis, real-world evidence, budget impact, value assessment, reimbursement and HTA decision making are all represented. What is absent is the measurement foundation that should determine whether the numerical operations employed within these methods have empirical meaning.

These two interrogations address different questions and their convergence is important. The 24 canonical statements identify the structure of the regional knowledge bases: weak possession of lawful-measurement requirements combined with strong possession of legacy HTA conventions. The 37-topic framework identifies the program-content environment through which that structure can be transmitted: complete representation of contemporary HEOR and HTA accompanied by absence of representational measurement as its scientific foundation. One identifies measurement inversion; the other identifies the educational and professional architecture capable of reproducing it.

The geographical consistency is particularly striking. Latin America, Asia, Africa and the Arabic-speaking region cannot plausibly be regarded as a homogeneous HTA environment. Their health systems, reimbursement institutions, research infrastructures and histories of formal HTA differ substantially. Nor can Central and Eastern Europe simply be equated with the longer-established HEOR and HTA environments of Western and wider Europe. Yet separating CEE from the wider European environment does not reveal a different foundational structure. Both display essentially the same 24-statement measurement inversion and the same 0/17 versus 20/20 program-content inversion. Greater institutional maturity has not corrected the problem. Neither has later adoption of HTA protected developing systems from inheriting it.

This convergence changes the interpretation of international methodological agreement. The appearance of similar methods across countries is conventionally treated as evidence of professional maturity and methodological convergence. From the perspective of representational measurement, however, agreement cannot substitute for demonstration. Fifty organizations employing substantially the same inherited arithmetic do not provide fifty independent demonstrations that the arithmetic is admissible. If each has inherited the same assumptions concerning utilities, QALYs, ICERs and reference-case models, their agreement demonstrates the successful transmission of a methodology rather than independent validation of its scientific foundations. Institutional reproduction is not measurement demonstration. What has been replicated is the methodology, not the measurement proof.

Viewed from this perspective, ISPOR's international expansion has functioned, irrespective of intention, as a mission for the global dissemination of measurement inversion. The organization did not need to instruct regional communities to reject representational measurement. It was sufficient to transmit an apparently complete methodological framework in which representational measurement was absent. Good Practices supplied authoritative methodological guidance; educational programs taught the accepted analytical methods; *Value in Health* provided a publication environment in which those methods were normalized; and regional Consortia, Networks, conferences and chapters facilitated their international dissemination. The repeated 0/17 versus 20/20 result reveals the consequence. Recipients acquired extensive competence in the apparatus of contemporary HEOR and HTA without simultaneously acquiring the logically prior measurement framework required to determine whether its arithmetic was scientifically admissible.

The regional communities were therefore not given an incomplete scientific framework awaiting supplementation. They were given a fully elaborated methodological system whose claim to quantitative scientific legitimacy had never been established. Without demonstrated measurement there is no legitimate foundation for the arithmetic; without admissible arithmetic there is no scientific foundation for the quantitative claims produced by it. What was transmitted internationally as a mature quantitative science was, on examination, an elaborate system of numerical conventions constructed without the measurement foundations required to give its central arithmetic empirical meaning.

This distinction is critical. To describe legacy HTA as scientifically incomplete would imply that representational measurement could now be added to the existing framework while leaving its established methods substantially intact. It cannot. The sequence was wrong from the beginning.

Utilities were assigned numerical properties that were never demonstrated; QALYs multiplied those values by time; incremental QALYs subjected the resulting quantities to further arithmetic; ICERs divided costs by them; and reference-case models projected the resulting claims into hypothetical futures. Each additional layer of methodological sophistication was constructed downstream from an unresolved measurement failure. Forty years of refinement cannot retrospectively manufacture the measurement properties that were absent before the arithmetic began.

The same educational inversion that transmitted legacy HTA also deprived its recipients of the intellectual tools required to expose its foundational failure. A practitioner could learn how to estimate utilities, calculate QALYs, construct ICERs, undertake probabilistic sensitivity analysis and populate sophisticated decision models without ever being required to ask whether multiplication of utility by time was admissible, whether the resulting QALY possessed legitimate quantitative properties or whether division by an incremental QALY had empirical meaning. Researchers were not necessarily instructed not to challenge the framework. More effectively, the logically prior body of knowledge from which such a challenge would arise was simply absent.

This makes the regional findings more troubling than a deliberate imposition of intellectual blinkers would have been. There was no need to suppress representational measurement if it never entered the professional curriculum. There was no need to defend the QALY against ratio-measurement criticism if generations of practitioners were never required to ask whether utility possessed ratio properties. There was no need to justify dimensional homogeneity if the requirement itself remained outside the methodological conversation. The 0/17 versus 20/20 inversion therefore did more than determine what practitioners learned. It helped determine which scientific questions they were equipped to ask.

Once established regionally, the framework became capable of reproducing itself. Researchers became teachers; teachers trained another generation of practitioners; practitioners became consultants, agency advisers and decision makers; investigators submitted analyses employing the accepted framework; reviewers assessed those analyses against inherited methodological standards; and universities designed programs around the skills demanded by agencies and employers. No continuing central direction was required. The methodology acquired local institutional legitimacy because each participant encountered everyone else using substantially the same methods. What originated as international methodological transmission became regional professional convention.

A global straitjacket was thereby created. Agencies could point to international practice. Manufacturers could point to agency requirements. Universities could point to professional competencies and employment expectations. Journals could point to accepted methodological standards. Professional organizations could point to expert consensus and widespread use. Each component appeared to validate the others. Yet the circle never reached the logically prior question: had the quantities entering the arithmetic actually been demonstrated to possess the required measurement properties?

The regional interrogations now break that circle. The Latin America Consortium, Asia Consortium, Africa Network, Arabic Network, CEE Consortium and wider European environment

can no longer be regarded simply as recipients of an established international scientific methodology. The 24 canonical statements expose essentially the same measurement inversion across these environments, while the categorical 0/17 versus 20/20 findings expose the program-content structure accompanying it. The missing foundation has now been identified. Measurement inversion is no longer invisible. Regional organizations and their individual members now possess a basis for asking the question that the inherited framework failed to ask: do the quantities entering the arithmetic satisfy the requirements of lawful measurement?

Once that question is admitted, there is no scientifically coherent option of continuing legacy HTA while merely acknowledging its measurement deficiencies. Representational measurement is not an alternative methodological perspective that can coexist with operations violating its requirements. If multiplication requires appropriate ratio measurement, the requirement cannot be waived to preserve the QALY. If dimensional homogeneity is required, it cannot be suspended because cost-per-QALY analysis is institutionally convenient. If a latent therapy-impact claim requires lawful measurement of a defined attribute, assigning numbers to questionnaire responses and statistically transforming or aggregating them cannot substitute for measurement. Recognition of the failure therefore carries an unavoidable consequence: methods that cannot satisfy the required measurement conditions must be abandoned.

That obligation applies not only to ISPOR centrally but throughout its international structure and ultimately to individual practitioners. A researcher cannot establish the admissibility of an arithmetic operation by pointing to an ISPOR Good Practice. A teacher cannot justify transmitting an inadmissible method because it appears in an established HEOR curriculum. A regional Consortium or Network cannot confer measurement properties through widespread adoption. Nor can an HTA agency transform numerical convention into lawful measurement by requiring it in a reference case. Professional authority can determine what a community accepts. It cannot determine what constitutes measurement.

This does not imply that regional researchers and practitioners should be blamed for having worked within the framework they inherited. They were introduced to a methodological system through an international professional environment in which the measurement knowledge required to interrogate its foundations was largely absent. That explains how measurement inversion could be transmitted and reproduced for decades. But explanation cannot become justification once the failure has been exposed. There is an important difference between inheriting a defective framework without possessing the intellectual tools required to identify the defect and continuing to defend that framework after the defect has been demonstrated.

There is consequently no legitimate strategy of preserving legacy HTA through incremental repair. The problem is not a defective parameter, an inadequate utility algorithm, an imperfect model or an insufficiently sophisticated sensitivity analysis. It lies upstream of all of them. Measurement never established the legitimacy of the arithmetic. Adding further methodological sophistication therefore cannot rescue the framework. Arithmetic cannot rescue a failure of measurement by adding further arithmetic.

Abandonment must accordingly extend across the international structure in which legacy HTA has become embedded. This does not mean abandoning economics, costs, clinical outcomes, statistics,

resource utilization, uncertainty or quantitative therapy assessment. It means abandoning the measurement-inverted architecture through which those legitimate activities have been organized. Utilities that have not demonstrated the properties required for the arithmetic performed upon them, QALYs constructed through inadmissible multiplication, cost-per-QALY ICERs dependent upon those quantities and reference-case simulations presented as substitutes for empirically evaluable therapy-impact claims have no continuing place in a framework committed to lawful measurement.

The regional communities should therefore not regard exposure of measurement inversion as the destruction of a scientifically valuable inheritance. It is liberation from a methodological inheritance whose claim to quantitative scientific legitimacy was never established. Individual researchers are no longer required to defend methods merely because those methods were presented to them as accepted international practice. Universities are no longer required to teach them because professional organizations have historically endorsed them. Agencies are no longer required to demand them because other agencies do. The discovery of the missing measurement foundation removes the principal intellectual justification for continued conformity.

The global straitjacket can therefore be broken, but not by loosening it. Legacy HTA has to be abandoned. The only scientifically defensible path is reconstruction from measurement forward: attribute, measurement, scale, admissible arithmetic, therapy-impact claim, empirical evaluation, replication and falsification. Manifest attributes require appropriate linear ratio measurement; latent attributes require Rasch measurement of defined unidimensional attributes. The replacement does not require another international convention to compete with the old one. It requires restoration of the ordinary requirements of quantitative science.

ISPOR's global success has consequently created both the scale of the problem and a potential mechanism for its resolution. The international infrastructure that disseminated measurement inversion could now disseminate its abandonment. Consortia and Networks that transmitted legacy HEOR could become vehicles for transition to lawful measurement. Educational programs that produced 0/17 versus 20/20 could reverse their sequence. Researchers who inherited the old framework could become the first generation to reject it rather than the next generation to reproduce it.

This may represent the end of empire for legacy HTA. The international methodological order constructed around utilities, QALYs, ICERs and reference-case modelling cannot survive the restoration of measurement as the gatekeeper for arithmetic. But the end of that empire need not represent the end of quantitative therapy assessment or of the international organizations and professional communities currently associated with it. The arithmetic travelled globally while the measurement proof did not. The task now is not to supply a missing proof for an architecture that cannot satisfy it. It is to abandon that architecture and replace it with lawful measurement. The organization whose international infrastructure helped create the global straitjacket may now be uniquely positioned to break it.

THE TRANSITION TO LAWFUL MEASUREMENT

The transition from legacy HTA to lawful measurement does not have to await reform by ISPOR, national HTA agencies, professional societies or universities. Nor does it require another prolonged methodological debate over whether utilities, QALYs and cost-per-QALY claims might somehow be rehabilitated. The measurement failure has been identified. Its consequences follow from the requirements of representational measurement. The immediate question is therefore practical: how can researchers, faculty, students and decision makers acquire the knowledge required to work within a measurement-first framework?

That transition has already been operationalized. Maimon Research has developed a nine-unit program designed specifically to reconstruct health technology assessment from the foundations of representational measurement [x] . The program does not add a measurement module to the existing HTA curriculum. It reverses the inherited sequence. Measurement precedes arithmetic; the attribute precedes the instrument; the properties of the measurement scale determine the admissible mathematical operations; and therapy-impact claims must be capable of empirical evaluation, replication and falsification. The purpose is not to improve legacy HTA but to provide the intellectual and practical basis for abandoning it.

The nine units constitute a structured progression rather than a collection of independent methodological topics. They begin with measurement inversion and the consequences of allowing numerical manipulation to precede demonstration of measurement. The program then moves from the definition of the attribute and the requirements of representational measurement to the distinction between manifest and latent attributes and their different measurement pathways. Particular attention is given to the requirements for ratio measurement, admissible transformations, admissible arithmetic and dimensional homogeneity. For latent attributes, the program introduces Rasch measurement as the pathway to construction of a unidimensional invariant continuum and the assessment of possession and change. For manifest attributes, it establishes the requirements that must be satisfied before claims based upon linear ratio measurement can be accepted. The later units bring these elements together in the construction and evaluation of therapy-impact claims and in the transition from legacy analytical practices to a measurement-first framework.

The sequencing is critical. Students are not first taught utilities, QALYs, cost-effectiveness models and probabilistic simulations and subsequently offered measurement theory as an optional qualification. They are required to pass through the measurement gate before downstream quantitative operations can be considered. If the required measurement properties cannot be demonstrated, the proposed arithmetic is rejected. This single reversal restores the relationship between measurement and mathematics that legacy HTA inverted.

The program is available now. It can support independent study, faculty development and structured graduate instruction, and the nine-unit package provides the substantive foundation for a three-credit graduate course. It therefore offers universities and research groups a practical route for beginning the transition without waiting for their professional organizations to redesign curricula, rewrite Good Practices or reconsider reference cases. A department does not need permission from ISPOR to teach representational measurement. A researcher does not need an HTA agency to abandon an inadmissible operation before refusing to use it. A graduate program

does not need an international consensus before requiring its students to demonstrate that a measure possesses the properties demanded by the arithmetic performed upon it.

This is why the availability of an independent transition program is important. Researchers and institutions need not remain trapped between recognition that legacy HTA fails the requirements of measurement and the absence of an alternative educational framework. The alternative already exists. It begins with two required measurement pathways: linear ratio measurement for appropriate manifest attributes and Rasch measurement for latent attributes. From these foundations follow admissible arithmetic, meaningful therapy-impact claims, empirical testing, replication and falsification. The baggage of multiattribute utility instruments, preference scores, QALYs and cost-per-QALY claims is not carried into the new framework because their continued inclusion would reproduce the measurement inversion that the transition is intended to eliminate.

Institutional transition will inevitably take time. Scientific transition need not. Universities can begin with individual courses and faculty development. Research groups can require measurement demonstrations before accepting quantitative claims. Graduate students can be taught to ask what attribute is being measured, what scale has been established and what arithmetic that scale permits before they are taught how to construct increasingly elaborate models. Agencies can ultimately develop submission standards around the same principles. These changes can proceed incrementally at the institutional level while remaining uncompromising at the scientific level.

The essential point is that there is no need to wait. The reconstruction of HTA from measurement forward has already been formulated as a teachable and operational program. What remains is not the invention of an alternative but its adoption. The transition to lawful measurement is available now and has been since the 17^{TH} century and the commitment to the evolution of objective knowledge.

CONCLUSION: AN UNLIKELY DAMASCENE CONVERSION

ISPOR now confronts a problem that cannot plausibly be presented as the discovery of a new methodological standard. The principles against which legacy HTA fails substantially predate the organization itself. The foundations of modern empirical science reach back to the Scientific Revolution some four centuries ago. Scales of measurement were formally articulated in 1946 and representational measurement systematically developed by 1971. ISPOR was not founded until 1995. For its entire existence, therefore, the principle that measurement must precede arithmetic, and that the properties of a measurement scale determine the mathematical operations that can legitimately be performed, was available to be recognized.

It was not recognized. This omission was not incidental to the development of legacy HTA. Its reference-case architecture required utilities to be multiplied by time, QALYs to be subtracted, incremental costs to be divided by incremental QALYs and the resulting ratios to support decision rules. Had lawful measurement been imposed at the outset, this architecture would have confronted its foundational problem immediately. Numbers could not simply have been assigned and manipulated because the reference case required them.

Ratio measurement alone should have stopped the process. Multiplication and division require appropriate measurement properties. Those properties must be demonstrated before the arithmetic is undertaken; they cannot be conferred by the arithmetic itself. Yet nowhere in the ISPOR knowledge environments examined here is ratio measurement treated as the prerequisite for the multiplicative operations upon which the QALY and downstream cost-per-QALY claims depend. For a methodology constructed around multiplication and division, this is an extraordinary absence.

The implication is uncomfortable but unavoidable. For legacy HTA to succeed internationally in its familiar form, lawful measurement effectively had to remain outside the analytical framework. Had the measurement gate been imposed, utilities would have been required to demonstrate the properties necessary for multiplication by time. Without that demonstration there could be no legitimate QALY; without a legitimate QALY there could be no legitimate incremental QALY; and without that denominator the conventional cost-per-QALY ICER could not survive. Whether this exclusion was conscious is irrelevant. The structural consequence was the same: arithmetic was permitted to precede measurement.

This makes a Damascene conversion by ISPOR scarcely credible. ISPOR is not being asked to incorporate another technique or revise another Good Practice. It would have to acknowledge that scientific standards available throughout its existence challenge central methods that it has spent approximately three decades publishing, teaching, endorsing and disseminating internationally. The accumulated institutional dead weight of Good Practices, journals, educational programs, task forces, conferences, professional expertise, careers, consultancy, agency relationships and regional networks is formidable. Genuine acceptance of lawful measurement would require abandonment, not accommodation, of methods that cannot pass the measurement gate.

The fundamental good practices of empirical science did not originate with ISPOR and do not require authorization by a professional society. Their intellectual origins reach back some four centuries, with subsequent developments in measurement theory, empirical testing, replication and falsification progressively formalizing the standards of normal science [xi]. Procedural sophistication cannot substitute for them. Before modelling comes measurement; before arithmetic comes demonstration that the scale permits the operation; before acceptance comes empirical evaluation.

Fortunately, an ISPOR conversion is unnecessary. There is no central authority that determines the application of the standards of normal science, and there is no need to construct another global superstructure to replace ISPOR. Scientific authority resides in demonstration, not professional consensus. A thousand organizations performing the same inadmissible arithmetic cannot make that arithmetic admissible.

The challenge therefore passes directly to HTA agencies, universities and research groups. The transition is straightforward. Define the attribute. Distinguish manifest from latent attributes. Require lawful linear ratio measurement for appropriate manifest attributes and Rasch measurement for latent attributes. Establish the measurement properties before undertaking arithmetic. Construct therapy-impact claims that are empirically evaluable, replicable and capable

of falsification. Abandon utilities, QALYs, cost-per-QALY claims and other constructions that cannot pass the measurement gate.

The framework for making this transition is available now. There is no reason to wait for ISPOR, another international consensus or another generation of Good Practices. The objective is not to replace one orthodoxy with another, but to replace authority by demonstration. The standards are already known, the pathway is straightforward and no permission is required. The challenge is simply to apply to HTA the standards of normal science that should have governed it from the beginning.

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

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

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

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

[v] Langley P. Legacy HTA: ISPOR – The endorsement of measurement inversion in practice pattern standards, Logit Working Paper No 2000. https://maimonresearch.com/logit-working-paper-no-2000-august-2026/

[vi] Langley P. ISPOR: Normal science and the absence of representational measurement in HTA education. Logit Woking Paper No 3010. https://maimonresearch.com/logit-working-paper-no-3010-july-2026/

[vii] Langley P. ISPOR global network: The endorsement of measurement failure – Latin American Consortium, Logit Working Paper No 4018. https://maimonresearch.com/logit-working-paper-no-4017-february-2026/

[viii] Langley P. ISPOR global network: the endorsement of measurement failure – Africa Network, Logit Working Paper No 401. https://maimonresearch.com/logit-working-paper-no-4016-february-2026/

[ix] Langley P. ISPOR global network: The endorsement of measurement failure -Middle East Network. Logit Working Paper No 4015. https://maimonresearch.com/logit-working-paper-no-4015-february-2026/

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

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