MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

AUSTRALIA: THE PREROGATIVE OF ADMINISTRATIVE CLOSURE IN HEALTH TECHNOLOGY ASSESSMENT

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

LOGIT WORKING PAPER No 1702 JULY 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

This paper makes the case that the defining characteristic of contemporary Australian health technology assessment (HTA) is not scientific measurement but the institutional imperative for administrative closure. Building upon previous Logit Working Papers demonstrating measurement inversion and curriculum inversion within Australian HTA, the present analysis asks a more fundamental question: why has a framework that conflicts with the accepted axioms of representational measurement become so firmly established across agencies, universities, research centers, journals, and professional education?

Using large language model (LLM) interrogation of the Australian HTA knowledge base and its nested curriculum knowledge base, the paper demonstrates that the concepts required for lawful quantitative science—including representational measurement, admissible arithmetic, unidimensionality, dimensional homogeneity, the distinction between manifest and latent attributes, linear ratio measurement, Rasch logit ratio measurement, and falsifiable claims—are largely absent. At the same time, utilities, QALYs, cost-effectiveness ratios, and reference-case simulation models are consistently endorsed as though they provide valid quantitative measures of therapy impact.

The paper argues that this inversion is explained by the historical development of the reference-case paradigm. Faced with the practical necessity of making reimbursement decisions under conditions of uncertainty, HTA evolved an administrative decision framework rather than a measurement-based scientific discipline. Beginning with time trade-off preference scores and culminating in lifetime simulation models, every stage in the construction of the reference case proceeds without satisfying the accepted requirements of representational measurement. Administrative closure consequently replaced scientific evaluation, while university curricula reproduced the same analytical framework without exposing students to the measurement principles required to evaluate it.

The implications are uncompromising. The reference-case paradigm is not capable of supporting scientifically valid claims regarding therapy impact because its central constructs are measurement impossibilities rather than quantitative measures. Continued refinement cannot resolve this failure because the problem is foundational. The administrative closure model must therefore be abandoned and replaced by a measurement-based framework founded upon first principles: specification of target attributes, distinction between manifest and latent attributes, linear ratio measurement for manifest claims, Rasch logit ratio measurement for latent claims, and prospectively specified, evaluable, replicable, and falsifiable therapy-impact claims. The paper concludes that this reconstruction is no longer optional but essential if Australian HTA is to regain scientific credibility.

INTRODUCTION

The Australian health technology assessment (HTA) system has, over more than three decades, developed an internationally respected framework for evaluating new pharmaceuticals and other health technologies. Through the Pharmaceutical Benefits Advisory Committee (PBAC), supported by academic research centers, university teaching programs, consultancy organizations, and methodological guidance, Australia has contributed significantly to the international development of cost-utility analysis, reference-case modelling, and evidence-based reimbursement policy. The Australian model has, in turn, influenced methodological developments in the United Kingdom, Canada, New Zealand, Europe, and numerous other jurisdictions. It is widely regarded as one of the most mature and sophisticated HTA systems currently in operation.

The companion Logit Working Papers in this series have challenged that perception from a fundamentally different perspective. Rather than examining the statistical, clinical, or economic sophistication of Australian HTA, they have examined the scientific foundations upon which its quantitative claims are constructed. Interrogations of the Australian HTA knowledge base, the PBAC methodological framework, leading Australian HTA research centers and university teaching programs have consistently demonstrated the presence of both measurement inversion and curriculum inversion. The concepts required for lawful quantitative measurement including scales of measurement [1], the axioms of representational measurement [2], unidimensionality, dimensional homogeneity, the distinction between manifest and latent attributes, linear ratio measurement, Rasch logit ratio measurement [3], and falsifiable scientific claims receive little recognition [4]. At the same time, utilities, QALYs, cost-effectiveness ratios, reference-case simulation models, and associated analytical techniques are consistently endorsed as though they provide valid quantitative measures of therapy impact.

These findings raise a more fundamental question than whether particular methods are scientifically defensible. They invite consideration of why a framework exhibiting such remarkable consistency across agencies, universities, research centers, journals, and professional education has become so firmly established. The answer proposed in this paper is that the organizing principle of contemporary HTA has never been scientific measurement. It has been administrative closure.

Health-care systems face an unavoidable practical problem. Governments cannot postpone reimbursement decisions until decades of empirical evidence have accumulated. Decisions regarding public funding must be made despite uncertainty concerning long-term effectiveness, safety, resource utilization, and patient benefit. The institutional requirement is therefore not simply to generate knowledge but to reach decisions. Contemporary HTA evolved as an administrative response to that requirement. The reference-case framework provides a standardized structure through which incomplete evidence, clinical trials, observational studies, preference weights, utilities, costs, epidemiological assumptions, and economic projections can be integrated into a single recommendation regarding reimbursement. From an administrative perspective, the framework is remarkably successful. It produces determinate recommendations under conditions of uncertainty and provides decision makers with a transparent procedural basis for resource allocation.

The central argument of this paper is that this administrative success has come at a scientific cost. The construction of the reference case appears to have proceeded without imposing the constraints required by representational measurement [5]. Had those constraints been accepted as fundamental design principles, the modern HTA framework could not have evolved in its present form. Utilities could not have been multiplied by time without demonstrating ratio-scale properties. QALYs could not have served as denominators for cost-effectiveness ratios. Latent attributes could not have been represented through ordinal preference scores without an appropriate measurement model. Composite preference instruments would have required demonstrations of unidimensionality, while lifetime simulation models would have been recognized as analytical projections rather than empirically testable evidence. In short, the scientific requirements governing quantitative measurement would have constrained the architecture of the reference case before it became institutionalized.

This interpretation also explains the companion finding of curriculum inversion. If administrative closure rather than measurement became the organizing principle of HTA, educational programs would naturally emphasize the methods required to construct and interpret the reference case rather than the scientific principles that might constrain or invalidate it. Students would be trained in economic evaluation, modelling, utilities, QALYs, sensitivity analysis, and reimbursement methodology before encountering the concepts of representational measurement, ratio scales, Rasch measurement, or falsifiability. The educational system would therefore reproduce the analytical framework rather than critically evaluate its scientific foundations. Measurement inversion and curriculum inversion would become mutually reinforcing characteristics of the discipline.

The purpose of the present paper is to deconstruct this interpretation. Drawing upon the cumulative findings of the Australian interrogation program, it argues that contemporary HTA is best understood as an institutional framework designed to achieve administrative closure under conditions of uncertainty rather than as a measurement-based scientific discipline. The issue is not whether reimbursement decisions should be made; they must be. The issue is whether the analytical framework used to support those decisions satisfies the standards of quantitative science. The evidence presented suggests that it does not. The future of Australian HTA therefore depends not upon further refinement of the reference case, but upon a fundamental reconstruction in which measurement once again precedes arithmetic, attributes precede models, and therapy-impact claims become prospectively evaluable, independently replicable, and capable of falsification.

1. THE AUSTRALIAN HTA KNOWLEDGE BASE

The central methodological innovation underpinning the present investigation is the use of large language model (LLM) interrogation to examine the intellectual content of a defined knowledge base. Unlike a conventional literature review, which summarizes individual publications or traces the historical development of ideas, an LLM interrogation evaluates the collective conceptual environment within which a discipline operates. The objective is not to determine whether a particular author, paper, university, or agency has addressed a given concept, but whether that concept is sufficiently represented within the accumulated body of publicly available material to constitute part of the discipline's accepted intellectual framework. The interrogation therefore asks a fundamentally different question: what does the knowledge base itself endorse?

This distinction is important. Scientific disciplines are defined not by isolated publications but by the body of concepts that are repeatedly communicated, reinforced, and reproduced through teaching, research, professional guidance, methodological standards, conferences, journals, and policy documents. A knowledge base is therefore much more than a collection of documents. It represents the accumulated assumptions, definitions, analytical methods, conceptual frameworks, and methodological expectations that collectively shape professional understanding. It is this broader intellectual environment that determines what students learn, what researchers investigate, what journals publish, what agencies expect in submissions, and ultimately what practitioners regard as accepted scientific knowledge.

The emergence of LLM interrogation makes it possible, for the first time, to examine this intellectual environment systematically. Because large language models synthesize relationships across vast collections of publicly available material, they provide an opportunity to interrogate the dominant concepts embedded within a knowledge base rather than merely cataloguing its individual components. The objective is not to replace traditional scholarship but to provide a complementary means of evaluating the extent to which particular scientific concepts are represented within a discipline. By presenting canonical statements derived from established scientific principles, the interrogation determines whether those principles are strongly represented, weakly represented, or effectively absent from the knowledge base as a whole. In this respect, the LLM functions as an instrument for examining the conceptual architecture of an academic discipline.

For the purposes of the present investigation, two related knowledge bases are defined. The first is the Australian HTA knowledge base. This encompasses the complete body of publicly accessible material through which contemporary Australian HTA is developed, communicated, applied, and reinforced. It includes PBAC methodological guidelines, MSAC methodological documents where relevant, university teaching programs, HTA research centers, academic publications, consultancy reports, professional society material, conference presentations, reimbursement submission guidance, methodological reviews, government reports, journal articles, technical working papers, and associated educational resources. Collectively these materials define the analytical framework through which Australian HTA understands evidence, value, comparative effectiveness, economic evaluation, and reimbursement decision making. They represent the intellectual environment encountered by researchers, policy analysts, consultants, manufacturers, committee members, postgraduate students, and health economists working within Australian HTA.

Nested within this broader intellectual environment is a second, more specific knowledge base: the Australian HTA curriculum knowledge base. Whereas the wider HTA knowledge base represents everything that contributes to professional understanding, the curriculum knowledge base is concerned specifically with the educational content through which future practitioners acquire that understanding. It comprises university course descriptions, lecture materials, doctoral training programs, workshops, professional education, methodological guidance used for teaching, training manuals, educational presentations, and other resources whose primary purpose is the transmission of HTA concepts and analytical methods. Its function is educational rather than methodological. It defines what students and early-career researchers are expected to know before entering professional practice.

The distinction between these two knowledge bases is fundamental to the interpretation of the present series of interrogations. The broader HTA knowledge base determines the concepts that dominate professional practice. The curriculum knowledge base determines whether those concepts are reproduced in the education of future practitioners. Measurement inversion and curriculum inversion therefore represent different, although closely related, phenomena. Measurement inversion concerns the scientific legitimacy of the analytical framework itself. Curriculum inversion concerns whether that framework is taught before students are exposed to the scientific principles required to evaluate it. The former concerns the content of HTA methodology; the latter concerns the sequence through which that methodology is transmitted.

The interaction between the two knowledge bases creates a self-reinforcing intellectual system. The wider HTA knowledge base establishes accepted methods, professional standards, journal expectations, reimbursement requirements, and research priorities. The curriculum knowledge base prepares successive generations of students to operate within that environment by teaching those same methods and expectations. Graduates subsequently become researchers, consultants, reviewers, journal editors, committee members, and university faculty who contribute to the continuing development of the wider HTA knowledge base. The relationship is therefore circular rather than linear. The curriculum reflects the professional knowledge base, while the professional knowledge base is continually renewed by graduates trained within the curriculum.

The importance of distinguishing these two knowledge bases becomes evident in the findings reported throughout this series of Logit Working Papers. Interrogations of the Australian HTA knowledge base demonstrate the widespread endorsement of measurement inversion. Companion interrogations of the curriculum knowledge base demonstrate that the concepts necessary to identify measurement inversion are themselves only weakly represented within educational programs. Together, these findings provide a coherent explanation for the remarkable stability of the contemporary HTA paradigm. The issue is not simply that particular analytical methods dominate Australian HTA. It is that the educational environment reproduces those methods while giving comparatively little attention to the scientific principles that might constrain or challenge them. Large language model interrogation therefore provides a unique opportunity to examine not merely what Australian HTA does, but how its underlying conceptual framework is created, maintained, and transmitted from one generation of practitioners to the next.

2. MEASUREMENT INVERSION: THE UNKNOWN UNKNOWN

One of the most striking findings to emerge from the Australian interrogation program is that there is virtually no evidence that measurement inversion has ever been recognized as a scientific problem within health technology assessment. This is not because the evidence supporting representational measurement is new, controversial, or inaccessible. The scientific foundations of measurement have been established for many decades and are widely recognized within measurement theory, psychometrics, mathematical psychology, and the philosophy of science. Rather, the interrogation suggests that these concepts have remained almost entirely outside the intellectual boundaries of HTA itself.

This distinction is important. The absence of representational measurement from Australian HTA should not be interpreted as evidence of deliberate exclusion. There is little reason to believe that

the architects of contemporary HTA consciously rejected the axioms of representational measurement or intentionally constructed a framework that violated them. The more plausible interpretation is that measurement theory simply did not form part of the conceptual environment within which modern HTA evolved. It was not rejected because it was never seriously considered. It existed outside the knowledge base that defined the discipline.

In retrospect, this absence is understandable. The early development of HTA was driven by practical policy concerns: how should governments compare therapies, allocate finite health-care resources, and make reimbursement decisions in the face of incomplete evidence? These are legitimate and unavoidable questions. They naturally encouraged the development of methods capable of integrating clinical evidence, epidemiology, costs, patient preferences, and long-term projections into a single analytical framework. Attention focused upon constructing a workable decision process rather than examining the scientific foundations of the quantities entering that process.

Once this analytical framework became established, it generated its own intellectual momentum. Universities taught the accepted methods. Research centers refined them. Journals published studies undertaken within them. Government agencies incorporated them into methodological guidance, while consultants and manufacturers applied them in reimbursement submissions. Each component of the system reinforced the others. Because representational measurement lay outside this intellectual environment, there was little opportunity for practitioners to encounter the concepts that might have challenged the framework's scientific legitimacy. The discipline therefore evolved without an explicit awareness of the measurement constraints that govern quantitative reasoning elsewhere in science.

This explains why the present findings should not be interpreted as criticism of individual researchers, teachers, reviewers, or committee members. Most have worked conscientiously within the accepted conceptual framework of their discipline. They have refined models, improved evidence synthesis, developed increasingly sophisticated statistical methods, and sought greater transparency in decision making. Yet these achievements have occurred within an intellectual framework in which the prior question of whether the quantities entering those analyses satisfy the requirements of measurement was never posed. The absence of measurement theory therefore became self-perpetuating. What was unknown to one generation remained unknown to the next because the educational system did not present it as part of the discipline.

The companion curriculum interrogations provide the explanation. Students entering Australian HTA programs are introduced to economic evaluation, utilities, QALYs, decision modelling, cost-effectiveness analysis, and reference-case methodology long before they encounter representational measurement, scale theory, unidimensionality, admissible arithmetic, or the distinction between manifest and latent attributes. Indeed, many may complete postgraduate training without ever being exposed to these concepts. Curriculum inversion therefore becomes the mechanism through which measurement inversion is reproduced. The educational framework does not deliberately suppress measurement theory; it simply fails to include it within the intellectual boundaries of HTA.

The result is what might be described as an "unknown unknown." Practitioners cannot question assumptions that they have never been taught to recognize. Consequently, the remarkable consistency observed across Australian agencies, universities, research centers, journals, and professional practice should not be interpreted as evidence that measurement inversion has survived repeated scientific examination. Rather, it reflects the fact that the relevant scientific framework has remained largely external to the discipline. Only with the emergence of systematic LLM interrogation has it become possible to examine the HTA knowledge base as a whole and ask whether the concepts required for lawful quantitative measurement are actually present. The answer, consistently across Australia and internationally, is that they are not.

3. NECESSITY AS THE MOTHER OF INVENTION

The previous sections have demonstrated that contemporary Australian HTA is characterized by both measurement inversion and curriculum inversion. These findings naturally lead to a more fundamental question. How did an analytical framework emerge that systematically overlooks the scientific principles governing quantitative measurement? The most plausible explanation is not that representational measurement was consciously rejected, but that a different institutional imperative took precedence. The overriding requirement was to create a framework capable of supporting timely reimbursement decisions under conditions of uncertainty. In this sense, necessity became the mother of invention. The necessity, however, was administrative rather than scientific.

Governments cannot postpone funding decisions until decades of empirical evidence accumulate. Pharmaceutical products enter the market with incomplete evidence regarding long-term effectiveness, safety, persistence, quality of life, and resource utilization. Yet reimbursement decisions must still be made. The institutional challenge was therefore to develop an analytical framework capable of combining disparate forms of evidence into a single recommendation. The resulting reference-case framework solved this administrative problem. It produced an apparently quantitative basis for deciding whether a technology represented value for money. The crucial question, however, is whether the scientific constraints governing measurement were ever allowed to shape the construction of that framework.

The starting point was the measurement of health-related quality of life through preference elicitation techniques such as the time trade-off (TTO). These procedures generated numerical preference values that appeared to provide a convenient quantitative representation of health states. The scientific question should immediately have been whether these numbers satisfied the requirements of measurement. Were the underlying attributes unidimensional? Were the resulting values ordinal, interval, or ratio? Did they possess a true zero? Could they legitimately support multiplication? Could they be aggregated across individuals? Could they be interpreted as quantitative measures of therapy impact? These questions were not central to the development of the framework. Instead, the numerical outputs were accepted as though they were already suitable for subsequent arithmetic.

The next step was the construction of multiattribute preference instruments. Individual health dimensions were combined to generate a single preference index intended to summarize overall health status. Once again, the essential measurement questions should have preceded the arithmetic. Had the individual dimensions been demonstrated to represent a single latent attribute?

Did aggregation preserve unidimensionality? Were the resulting preference scores lawful measures or merely numerical summaries of subjective responses? The interrogation findings suggest that these questions remained largely outside the conceptual framework of HTA supporters. Numerical construction proceeded without prior demonstration of lawful measurement.

The emergence of the quality-adjusted life year represented the next major innovation. The attraction of the QALY was immediately apparent. By multiplying preference values by time, therapies differing in survival and quality of life could apparently be compared using a single outcome measure. Yet this step represented perhaps the most significant departure from representational measurement. Multiplication requires ratio measurement. Time satisfies this requirement. Preference values do not. If the utility component lacks ratio-scale properties, the resulting product cannot constitute a ratio measure irrespective of the mathematical elegance of the calculation. At precisely the point where measurement theory should have constrained analytical development, it remained effectively absent from the conceptual framework.

Having accepted the QALY as the common outcome metric, the remaining components of the reference case followed almost inevitably. Costs could be divided by incremental QALYs to produce incremental cost-effectiveness ratios. Clinical trials could be linked to epidemiological assumptions. Transition probabilities could populate Markov models. Lifetime horizons could replace observed outcomes. Probabilistic sensitivity analyses could quantify uncertainty. Scenario analyses could explore alternative assumptions. At each stage, increasing analytical sophistication reinforced confidence in the framework. Yet the original measurement question remained unanswered. If the denominator entering the cost-effectiveness ratio was not itself a lawful quantitative measure, then every subsequent analytical refinement rested upon an unresolved scientific foundation.

This progression illustrates the central argument of the present paper. The evolution of the reference case did not proceed through a sequence of measurement decisions followed by appropriate arithmetic. It proceeded through a sequence of administrative innovations designed to improve decision making under uncertainty. Each innovation addressed a practical institutional problem. Preference instruments summarized complex health states. QALYs combined quality and survival. Simulation models extended incomplete evidence across a lifetime. Cost-effectiveness ratios provided a common decision metric. Sensitivity analyses explored uncertainty. Collectively these innovations produced an integrated framework capable of delivering reimbursement recommendations. What they did not produce was a demonstrable measurement foundation satisfying the requirements of representational measurement.

Viewed from this perspective, the remarkable stability of the reference-case inversion paradigm becomes easier to understand. Each successive methodological development addressed an administrative need while assuming the validity of the preceding stage. Once preference values were accepted, the QALY became possible. Once the QALY was accepted, incremental cost-effectiveness ratios followed naturally. Once those ratios became accepted, simulation models became indispensable. At no point was the development constrained by the requirement that measurement precede arithmetic. Instead, arithmetic generated the appearance of measurement.

The consequence is profound. Had representational measurement been treated as a non-negotiable design constraint, the reference case could never have evolved in its present form. Utilities would first have required demonstration of lawful measurement properties. Latent attributes would have required Rasch logit ratio measurement. Manifest attributes would have been evaluated using linear ratio measures. Therapy-impact claims would have been specified prospectively, accompanied by explicit protocols, and subjected to empirical evaluation, replication, and potential falsification.

Necessity may indeed have been the mother of invention. The difficulty is that the invention responded to the necessity for administrative closure rather than to the requirements of measurement science. The resulting framework solved the practical problem of producing a format to support administrative reimbursement decisions, but it did so by constructing an analytical architecture whose scientific foundations denied the scales and axioms of representational measurement, . That legacy now confronts Australian HTA. The challenge is no longer to refine the administrative reference case but to make clear that measurement inversion is not representational measurement. There is only one analytical framework that can support therapy impact claims.

4. ADMINISTRATIVE CLOSURE AS THE ORGANIZING PRINCIPLE OF HTA

The administrative reference-case framework has traditionally been presented as the foundation of modern health technology assessment. The evidence presented in the preceding sections suggests a radically different interpretation. Rather than functioning primarily as a framework for the discovery of objective knowledge, the reference case functions as a mechanism for administrative closure. Its principal purpose is not to establish whether therapy-impact claims can be subject to the standards of normal science but to provide a framework through which reimbursement decisions can be justified despite uncertainty with limited data. It resolves a political rather than a scientific problem. This distinction is fundamental because it changes the objective of HTA from scientific inquiry to institutional decision making.

Scientific inquiry and administrative decision making operate according to different principles. Science is necessarily open-ended. Claims are formulated prospectively, supported by valid measurement, subjected to empirical observation, independently replicated, and continually exposed to the possibility of falsification as new evidence accumulates. No scientific conclusion is ever final because every claim remains provisional. Knowledge advances by replacing weaker explanations with stronger ones. Administrative systems, by contrast, demand closure. Governments cannot postpone reimbursement decisions indefinitely while waiting for complete knowledge. A decision must eventually be reached. The institutional objective therefore becomes the production of a recommendation rather than the continual testing of competing hypotheses.

The reference-case framework resolves this problem by replacing empirical uncertainty with administrative convenience. Clinical trial evidence, epidemiological assumptions, utility values, costs, transition probabilities, extrapolated survival curves, and resource utilization estimates are brought together within a single simulation model. Sensitivity analyses explore alternative assumptions while probabilistic analyses quantify uncertainty. The resulting output is an

incremental cost-effectiveness ratio or similar summary statistic that appears to provide an objective basis for reimbursement. From an administrative perspective the process is highly effective. It transforms heterogeneous evidence into a single recommendation capable of supporting policy decisions. A decision is made with no incentive ever to challenge or return to it. Any commitment to the evolution of objective knowledge over the lifetime of a therapy is quashed.

The difficulty is that this process depends upon assumptions whose measurement status has never been established. Utilities are treated as though they were ratio measures. QALYs are accepted as though they were dimensionally homogeneous. Composite preference scores are manipulated as though they represented unidimensional quantitative attributes. Lifetime simulations are interpreted as though they constituted empirical evidence rather than analytical projections. At each stage the appearance of quantitative precision increases while the measurement foundations become progressively more distant. The framework therefore achieves administrative closure by assuming rather than demonstrating the validity of the quantities upon which its conclusions depend.

The cumulative interrogation program suggests that this process is neither accidental nor uniquely Australian. Comparable findings have been reported across agencies, universities, research centers, journals, and professional organizations in numerous jurisdictions. The remarkable consistency of the results indicates that the reference case has become the dominant institutional solution to a common policy problem: how to make funding decisions when evidence is incomplete. The consequence is that the same analytical structure is reproduced internationally irrespective of whether its underlying quantities satisfy the scientific requirements of representational measurement.

Curriculum inversion completes this institutional architecture. Students are introduced to reference-case construction, cost-utility analysis, decision modelling, evidence synthesis, and economic evaluation before they encounter the scientific principles that govern quantitative measurement. They therefore learn how to operate the analytical framework without first acquiring the conceptual tools needed to evaluate its scientific legitimacy. Universities produce graduates skilled in constructing reference cases. Research centers refine them. Journals publish them. Government agencies require them. Manufacturers commission them. Consultants prepare them. Each component of the system reinforces every other component. The reference case becomes not merely an analytical method but the defining intellectual framework of the closed HTA discipline.

Viewed in this way, measurement inversion and curriculum inversion are not independent phenomena. They represent complementary components of a single institutional system. Measurement inversion makes administrative closure possible by permitting arithmetic without prior measurement. Curriculum inversion makes administrative closure sustainable by ensuring that successive generations of practitioners are trained within the same conceptual framework. Together they explain the extraordinary stability of contemporary HTA despite the absence of the measurement foundations required by quantitative science.

Over the past six months, the inversion cat has been let out of the bag. The series of LLM interrogations has made increasingly difficult to sustain the fiction that contemporary HTA rests on secure quantitative foundations. Across agencies, academic centers, journals, teaching

programs and national knowledge bases, the same profile has appeared: the axioms of representational measurement are absent or weakly endorsed, while utilities, QALYs, cost-effectiveness ratios and reference-case simulations are treated as though they were legitimate measures of therapy impact. The conclusion is no longer a matter of isolated criticism. The entire edifice of measurement inversion and curriculum inversion has been exposed as a charade: an elaborate professional apparatus in which arithmetic has been allowed to substitute for measurement and training has been organized around the reproduction of that substitution.

This marks a turning point. HTA can no longer claim that its reference-case framework has simply been misunderstood or requires marginal refinement. The problem is foundational. Rather than adopting the standards required for claims development, empirical evaluation, replication and falsification, HTA institutionalized for over 30 years a system of administrative decision support built on quantities that fail the requirements of measurement. The reference case has now been shown to provide closure without measurement, modelling without falsifiability, and numerical outputs without lawful quantitative standing. From this point forward, every continued reliance on utilities, QALYs and cost-effectiveness ratios carries the burden of explaining why HTA should be exempt from the measurement standards that govern every other quantitative science.

5. FROM ADMINISTRATIVE CLOSURE TO MEASUREMENT-BASED HTA

The question is not whether HTA should continue to exist. Health technology assessment remains essential for evaluating new technologies and informing reimbursement decisions. What must end is the administrative closure model that has dominated HTA for more than three decades. That framework cannot be rescued through methodological refinement because every stage in its construction violates the accepted axioms of representational measurement. It is not an imperfect scientific framework awaiting improvement. It is a measurement impossibility.

Once measurement inversion is demonstrated, there is no defensible scientific option other than reconstruction. The administrative reference-case model must be replaced by a framework founded upon the principles of quantitative science.

This is not one methodological option among several. It is the only framework consistent with the accepted principles of representational measurement and, therefore, the only framework capable of supporting objective, cumulative, evaluable, replicable, and falsifiable scientific knowledge regarding therapy impact. The era of administrative closure has reached its intellectual end. The future of HTA depends upon abandoning impossible measurement free constructs and rebuilding the discipline upon the foundations of measurement science.

A measurement-based HTA would begin from an entirely different premise. The starting point would no longer be the construction of a reference-case model capable of producing a single cost-effectiveness estimate. Instead, every assessment would begin by specifying the attribute that the intervention is intended to influence. Is the outcome of interest survival, disease progression, symptom burden, treatment satisfaction, physical functioning, need fulfilment, medication adherence, hospital utilization, or some other clearly defined attribute? Until the target attribute

has been specified, there can be no meaningful discussion of measurement, no lawful arithmetic, and no valid quantitative claim.

Once the attribute has been identified, the second question concerns its measurement. Manifest attributes, because they are directly observable, require linear ratio measures with an identifiable true zero. Hospital admissions, emergency department visits, days in hospital, mortality, treatment persistence, and medication possession are all examples of outcomes that can support linear ratio measurement. Latent attributes present a fundamentally different challenge because they cannot be observed directly. Pain, fatigue, anxiety, treatment satisfaction, quality of life, and need fulfilment require measurement models capable of estimating possession of the underlying attribute. For these attributes, Rasch logit ratio measurement provides the only scientifically defensible framework for constructing lawful quantitative measures.

Measurement therefore precedes arithmetic. Only after the measurement properties of the attribute have been established can numerical operations be undertaken. This single principle transforms the structure of HTA. Rather than constructing a single synthetic outcome such as the QALY, therapy impact is evaluated through a series of independent, prospectively specified claims, each supported by an appropriate measurement framework. Arithmetic becomes subordinate to measurement rather than its substitute.

The consequences extend well beyond methodology. Every therapy-impact claim would require a protocol specifying the target population, comparator, observation period, outcome measure, success criteria, analytical methods, and the empirical conditions under which the claim could be replicated or falsified. Evidence generation would become a continuing process extending beyond product launch. Reimbursement decisions would still be made when required, but they would rest upon explicitly provisional claims that remain open to confirmation, modification, or rejection as additional evidence accumulated. Administrative decisions would therefore become milestones within a continuing program of scientific inquiry rather than the endpoint of a modelling exercise.

This reconstruction also transforms professional education. Universities would continue to teach economic evaluation, comparative effectiveness, modelling, and reimbursement methodology, but only after students had acquired the scientific foundations of quantitative reasoning. Representational measurement, the principal scales of measurement, admissible arithmetic, dimensional homogeneity, unidimensionality, the distinction between manifest and latent attributes, linear ratio measurement, Rasch logit ratio measurement, and falsifiability would become the intellectual foundations upon which all subsequent analytical methods are evaluated. Graduates would not merely know how to construct meaningful claims; they would understand the scientific conditions under which any quantitative claim can legitimately be made.

The reconstruction proposed here is therefore evolutionary rather than destructive. Clinical trials, observational studies, epidemiology, evidence synthesis, economic information, and policy deliberation all retain important roles within HTA. What changes is the scientific discipline governing their interpretation. Representational measurement becomes the primary constraint upon quantitative reasoning, ensuring that arithmetic never outruns measurement and that analytical sophistication never substitutes for scientific validity.

The evidence presented throughout this paper leaves no middle ground. The future of Australian HTA cannot be secured through further refinement of the reference-case paradigm because the reference case itself is founded upon a sequence of measurement impossibilities. Every stage in its construction violates the accepted principles of representational measurement. Time trade-off preferences fail the requirements for ratio measurement. Multiattribute preference instruments fail the requirements for unidimensional measurement. QALYs cannot constitute ratio measures because one component of the product is not a ratio measure. Cost-effectiveness ratios therefore lack a lawful denominator, while reference-case simulation models merely propagate these measurement failures through increasingly sophisticated arithmetic. The problem is not one component of the framework. The entire architecture fails at every critical stage.

This has profound implications. The reference case can no longer be defended as a scientific instrument for evaluating therapy impact. Its historical purpose has been administrative rather than scientific: to produce a determinate recommendation capable of closing debate at product launch. That purpose has now been exposed. Administrative closure cannot substitute for scientific measurement. Once this is recognized, the reference case ceases to have any legitimate role in the scientific evaluation of therapies. It belongs to the history of HTA, not to its future.

The choice confronting Australian HTA is therefore not between two competing methodologies. It is between continuing to defend a framework that cannot satisfy the accepted axioms of measurement or reconstructing the discipline upon those axioms. There is no third alternative. The era of administrative closure has reached its intellectual limits. The future of HTA depends upon abandoning impossible constructs and restoring measurement as the indispensable foundation of quantitative science.

6. ACADEMIC RESPONSIBILITY AND THE CONTINUED TEACHING OF MEASUREMENT INVERSION DECISION MAKING

The evidence presented throughout this series of Logit Working Papers leads to an uncompromising conclusion. Contemporary HTA continues to teach and apply analytical constructs that are impossible under the accepted axioms of representational measurement. Utilities, QALYs, cost-utility analysis, and reference-case simulation models are not imperfect quantitative measures. They cannot satisfy the conditions required for quantitative measurement. Arithmetic cannot create measurement where measurement does not exist.

This fundamentally changes the educational responsibilities of universities. Students should understand utilities, QALYs, and reference-case modelling because these constructs have dominated HTA for more than forty years. They should also be taught why these constructs fail the accepted standards of measurement and why they cannot provide valid quantitative measures of therapy impact. Teaching students how to construct QALYs without teaching why they are measurement impossibilities substitutes professional convention for scientific education.

Universities exist to advance objective knowledge, not preserve methodological traditions. Once the measurement failure of the reference case is recognized, its continued presentation as a scientifically valid methodology becomes indefensible. It should continue to be taught, but as part of the history of HTA rather than as the future of quantitative assessment.

This raises a further question of academic responsibility. Do faculty responsible for HTA education understand the axioms of representational measurement and their implications for utilities, QALYs, and simulation models? If they do, why are these principles largely absent from HTA curricula? If they do not, on what basis can universities certify graduates as competent in quantitative assessment?

The future of HTA does not lie in increasingly sophisticated failed reference-case models. It lies in restoring measurement as the foundation of quantitative science. Universities therefore have an obligation to teach students not only how the reference case came to dominate HTA, but why its central constructs are measurement impossibilities that cannot provide scientifically valid assessments of therapy impact. Continuing to teach them as though they can is incompatible with the central mission of the modern university.

CONCLUSION: THE PATH TOWARDS QUANTITATIVE HTA IN AUSTRALIA

The recognition that the administrative reference-case model is not a scientifically viable decision framework places an immediate responsibility upon the HTA community. Criticism alone is insufficient. A comprehensive program is required to reconstruct HTA on the accepted principles of representational measurement. Fortunately, such a program already exists. The Maimon Research HTA Transformation Program, developed by Dr. Paul Langley provides a complete framework for this transition [6]. Rather than attempting to repair the irreparable, it abandons measurement impossibilities and returns HTA to first principles. . The program provides a practical blueprint for replacing administrative closure with a scientifically coherent system of therapy assessment.

The program provides a systematic introduction to representational measurement, the theory of attributes, the principal scales of measurement, admissible arithmetic, dimensional homogeneity, manifest and latent attributes, Rasch logit ratio measurement, protocol development, and the construction of evaluable, replicable, and falsifiable claims regarding therapy impact. Its purpose is not to modify the existing reference-case paradigm but to replace it with a scientific framework in which measurement once again precedes arithmetic.

The program has been designed for universities, HTA agencies, reimbursement organizations, research centers, professional societies, pharmaceutical companies, and health economists seeking a transition from assumption-driven modelling to scientifically defensible measurement. It provides a structured pathway for professional development while establishing the competencies required for the next generation of HTA practitioners. In this way, it offers not simply a critique of the existing paradigm but a practical route toward the reconstruction of HTA as a measurement-based scientific discipline.

The challenge confronting Australian HTA is therefore no longer to identify the problem. It is to implement the solution.

\

ACKNOWLEDGEMENT

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

REFERENCES

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

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

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

[4] Popper K. The Logic of Scientific Discovery. London: Hutchinson; 1959

[5] The end of the Reference Case: Reconstructing HTA. Logit Working Paper No 345. https://maimonresearch.com/logit-working-paper-no-345-june-2026-2/

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