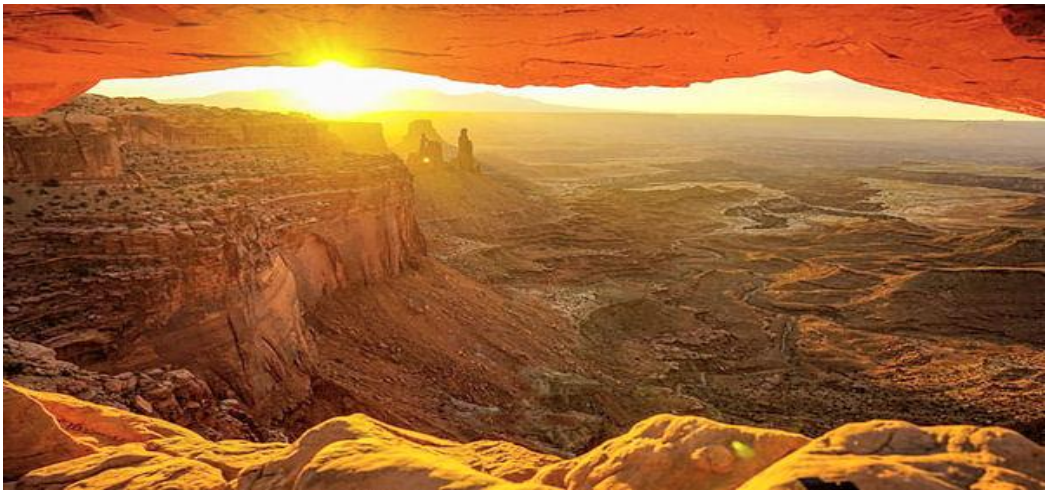
MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION



REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

AUSTRALIA: THE PBAC AND HTA RESEARCH CENTERS

A Dog's Breakfast of Measurement Inversion, Arithmetic Chaos and Closure

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

LOGIT WORKING PAPER No 177 MAY 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

This paper examines the presence of measurement inversion and arithmetic chaos within Australian health technology assessment (HTA) through interrogation of the knowledge bases associated with the Pharmaceutical Benefits Advisory Committee (PBAC) and five Australian HTA research centers: Melbourne School of Population and Global Health, the Centre for Health Economics at Monash University, the Health Economics Group at the University of Adelaide, the Centre for Applied Health Economics at Griffith University and the Leeder Centre for Health Policy at the University of Sydney CHERE at the Technology University of Sydney and HTA at Deakin University

The analysis became feasible through application of large language AI interrogation methods utilizing canonical diagnostic statements derived from representational measurement, scales of measurement and HTA methodology. Statements known to be TRUE or FALSE under accepted principles of measurement theory were evaluated through categorical endorsement probabilities designed to determine the extent to which the respective HTA knowledge bases behave as though these propositions are endorsed or rejected.

The findings reveal a stable and reproducible pattern of measurement inversion. FALSE propositions, including claims that the QALY is a ratio measure, that ratio measures can take negative values and that reference-case simulations generate falsifiable claims, receive consistently strong endorsement. In contrast, TRUE propositions concerning unidimensionality, ratio measurement, Rasch logit ratio measurement and the principle that measurement must precede arithmetic receive consistently weak endorsement. The findings therefore indicate that arithmetic operations are normalized within the Australian HTA framework while the standards necessary to justify those operations remain largely absent.

The paper argues that measurement inversion inevitably generates arithmetic chaos. Utilities lacking lawful ratio properties create QALYs; QALYs populate simulation models; simulation models generate cost-effectiveness ratios. Apparently coherent outputs continue to be produced, yet the arithmetic operations underpinning those outputs lack lawful quantitative meaning. Review of curriculum and methodological content across the respective centers further suggests institutional reproduction of these assumptions through HTA teaching and research.

The conclusion is direct: the Australian HTA environment represents a closed analytical structure in which measurement inversion and arithmetic chaos have become normalized within policy, teaching and research. Transition therefore requires abandonment of the reference-case framework in favor of measurable attributes, lawful ratio measures, protocol-driven claims and falsifiable empirical assessment.

INTRODUCTION

The central question is straightforward: do health technology assessment (HTA) reference-case or closed decision models satisfy the fundamental requirement that measurement must precede

arithmetic? Specifically, do the numerical entities entering utilities, QALYs, simulation models and cost-effectiveness ratios possess the lawful measurement properties required by scales of measurement and the axioms of representational measurement?

The findings from the knowledge base interrogations were clear cut. Utilities derived from health-state descriptions and preference exercises were treated as though they represented lawful quantitative entities suitable for multiplication, aggregation and ratio construction. Yet the standards necessary to justify these operations unidimensionality, ratio measurement and admissible arithmetic were absent from the analytical framework itself. Arithmetic operations therefore proceeded independently of demonstrated measurement properties. The result was a pervasive pattern of measurement inversion, where conclusions regarding therapy impact and resource allocation were generated before lawful measurement had been established.

Measurement inversion provided the signal for a second and more fundamental problem: arithmetic chaos. Once inadmissible numerical entities entered simulation models, each successive stage of the analytical process inherited and compounded the original contradiction. Utilities created QALYs; QALYs populated Markov simulation structures; simulation models generated cost-effectiveness ratios. Internally coherent outputs continued to be produced, yet the arithmetic operations themselves lacked lawful quantitative meaning.

INTERROGATING LARGE LANGUAGE MODELS

The analysis presented here on patterns of measurement inversion has only now proved feasible through the ability to interrogate, using large language AI models, the HTA knowledge bases of agencies such as the PBAC together with Australian HTA research centers. What previously existed as a broad methodological unease regarding utilities, QALYs and reference-case simulation models can now be examined systematically through structured interrogation of the assumptions embedded within teaching, research and policy frameworks.

The interrogation process utilizes a set of canonical diagnostic statements drawn from the standards of representational measurement, scales of measurement and HTA methodology. Some statements are demonstrably TRUE under accepted principles of measurement theory; others are demonstrably FALSE. The purpose is to determine the extent to which the respective HTA knowledge bases behave as though these propositions are endorsed or rejected. The resulting categorical probabilities therefore do not measure personal beliefs or institutional declarations. Rather, they reflect the degree to which the concepts embedded within the HTA knowledge base support, reinforce or neglect the propositions under examination.

The interpretation of the endorsement structure is straightforward. Strong endorsement of TRUE statements would indicate recognition of the requirements necessary for lawful quantitative inference, including unidimensionality, ratio measurement and the principle that measurement must precede arithmetic. Conversely, strong endorsement of FALSE statements would indicate acceptance of assumptions inconsistent with representational measurement, such as the treatment of the QALY as a ratio measure or the belief that ordinal utility structures support admissible arithmetic operations. The significance of the interrogation results presented here is that they reveal a stable and reproducible pattern in which FALSE propositions receive strong endorsement

while TRUE propositions receive weak endorsement. This pattern provides direct evidence for the presence of measurement inversion within the Australian HTA teaching and research environment.

THE EVIDENCE FOR MEASUREMENT INVERSION

The findings are important because they cast doubt on 40 years of HTA teaching and research in Australia. The findings they transformed what had previously been criticism and suspicion into measurable observations. Measurement inversion appeared not as an isolated methodological weakness but as a defining characteristic of the Australian HTA knowledge structure. Foundational propositions concerning representational measurement, lawful arithmetic and Rasch measurement received weak endorsement while false measurement assumptions received strong support.

Importantly, the Australian findings proved not to be unique. The endorsement patterns observed in Australian HTA knowledge bases replicated, identical endorsement profiles across a large number of international interrogations. Interrogations involving twenty-four countries, together with agencies, journals, educational institutions and professional organizations, produced remarkably stable findings. Across more than 230 assessments the same pattern repeatedly appeared.

The implications extend beyond HTA itself. Interrogations of COSMIN and the Cochrane Collaboration suggested that measurement inversion may be embedded not only within HTA but within the broader architecture of evidence generation and evidence synthesis [1] [2] Instrument standards, systematic reviews and meta-analysis appeared to inherit assumptions concerning measurement that largely ignored representational measurement. If so, the issue extends beyond reimbursement frameworks and reaches into the institutional assumptions supporting evidence itself.

Table 1 presents a comparison of endorsement probabilities for seven Australian HTA-related knowledge bases and the Pharmaceutical Benefits Advisory Committee [3]. These are (i) the Melbourne University School of Population and Global Health (MSPGH) [4], (ii) the Centre for health Economics (CHE) at Monash University [5] (iii) The health Economics Group (HEG) at The University of Adelaide [6], the Centre for Applied Health Economics (CAHE) at Griffith University [7]., the Leeder Centre for Health Policy at the University of Sydney [8], the Centre for Health Economics and Evaluation (CHERE) at University of Technology Sydney [9] HTA at Deakin University [10]

TABLE 1

MEASUREMENT INVERSION: PBAC AND 7 AUSTRALIAN HTA RESEARCH CENTRES

CANONICAL STATEMENT	CATEGORICAL PROBABILITY ENDORSEMENT							
	PBAC	MSPGH	CHE	HEG	CAHE	LEEDER	CHERE	DEAKIN
FALSE STATEMENTS								
1, Ratio measures can have negative values	0.85	0.85	0.90	0.85	0.80	0.80	0.90	0.85
2.The QALY is a ratio measure	0.90	0.85	0.90	0.85	0.40	0.55	0.90	0.90
3, Summations of subjective instrument responses are ratio measures	0.85	0.85	0.85	0.85	0.75	0.75	0.85	0.85
4. The QALY is a dimensionally homogeneous measure	0.85	0.80	0.85	0.80	0.40	0.50	0.85	0.90
5. Reference case simulations generate falsifiable claims	0.70	0.80	0.85	0.80	0.75	0.75	0.85	0.85
TRUE STATEMENTS.,								
6. Measures must be unidimensional	0.25	0.20	0.15	0.15	0.25	0.25	0.20	0.20
7. Multiplication requires a ratio measure	0.15	0.10	0.10	0.10	0.20	0.15	0.10	0.10
8. There are only two classes of measurement: linear ratio and Rasch logit ratio	0.10	0.05	0.05	0.05	0.05	0.05	0.05	0.05
9.Measurement precedes arithmetic	0.15	0.10	0.10	0.10	0.25	0.20	0.10	0.15

10. The outcome of interest for latent traits is the possession of that trait	0.25	0.25	0.25	0.25	0.05	0.05	0.25	0.10

Note: MSGPH Melbourne School of Population and Global Health; CHE Centre for Health Economics, Monash; HEG Health Economics Group, Adelaide; CAHE Centre for Applied Health Economics, Griffith; Leeder: Centre for Health Policy, Sydney; University of Technology Sydney (CHERE) and HTA group DEAKIN University

THE PATTERNS OF FALSE BELIEF

The results presented in Table 1 are remarkable, not because they reveal meaningful differences between the PBAC and Australia's leading HTA research centers, but because they reveal an extraordinary degree of consistency. Across the PBAC and the seven HTA research centers , the same pattern appears repeatedly. Statements that are demonstrably false from the perspective of representational measurement receive strong endorsement, while statements that are foundational requirements of quantitative science receive weak or negligible endorsement. This is not a collection of isolated misunderstandings, methodological preferences or differences in interpretation. It is evidence of a shared intellectual framework that has become deeply embedded within Australian HTA over 40 years. The findings point to a common culture in which arithmetic is accepted as a substitute for measurement and where the requirements of representational measurement have been displaced by the assumptions of the reference-case paradigm.

The significance of this finding should not be underestimated. These centers are not peripheral contributors to Australian health technology assessment. They are among the principal institutions responsible for research, postgraduate training, methodological development and policy advice. They have helped shape reimbursement policy, informed PBAC submissions, trained successive generations of health economists and HTA practitioners and contributed to the international dissemination of the reference-case paradigm. If these institutions fail to recognize the requirements of measurement, then measurement failure has become institutionalized within Australia; an unfortunate legacy.

Consider first the false statements. The proposition that ratio measures can have negative values receives endorsement probabilities ranging from 0.80 to 0.90. This is not a minor technical error. The existence of a meaningful non-arbitrary zero is one of the defining properties of a ratio scale. If a measure can take negative values, it is not a ratio measure. Yet across all centers there is strong support for a proposition that directly contradicts one of the most elementary requirements of measurement theory.

The same pattern is evident for the proposition that the QALY is a ratio measure. Endorsement ranges from 0.40 to 0.90, with most centers clustered between 0.85 and 0.90. The implication is clear. The overwhelming majority of the Australian HTA knowledge base treats the QALY as though it satisfies the requirements of ratio measurement. Yet the QALY is constructed from utility

values that themselves originate in ordinal preference structures. If the utility lacks ratio properties, multiplication of time cannot create them. The QALY therefore inherits the deficiencies of its components. The strong endorsement of the QALY as a ratio measure demonstrates a profound misunderstanding of dimensional homogeneity and admissible arithmetic.

Equally revealing is the endorsement of the proposition that summations of subjective instrument responses are ratio measures. Endorsement probabilities range from 0.75 to 0.85. This finding points directly to the treatment of patient-reported outcomes throughout Australian HTA. Ordinal responses are routinely summed and interpreted as though the resulting scores possess quantitative meaning. The possibility that latent attributes require a measurement model capable of establishing invariant and unidimensional measurement is largely absent. The Rasch framework is effectively invisible.

The endorsement of the proposition that the QALY is dimensionally homogeneous is particularly important. Dimensional homogeneity is one of the oldest principles in quantitative science. The multiplication of quantities requires compatible dimensional properties. Yet endorsement probabilities remain high across virtually all centers. This suggests that the problem is not merely misunderstanding the QALY. It is a failure to appreciate the role of dimensional analysis itself.

The proposition that reference-case simulations generate falsifiable claims receives endorsement levels between 0.70 and 0.85. This finding reveals an equally serious problem. Simulation models project hypothetical futures based upon assumptions regarding disease progression, treatment persistence, utility values, costs and transition probabilities. Their outputs are not empirical propositions capable of direct falsification. They are conditional numerical projections. Yet the Australian HTA community appears willing to treat these outputs as though they possessed the status of scientific claims.

The true statements are even more revealing. The proposition that measures must be unidimensional receives endorsement levels between 0.15 and 0.25. In other words, one of the most fundamental requirements of measurement is largely absent from the Australian HTA knowledge base. Without unidimensionality there can be no meaningful interpretation of a measure because multiple attributes are being combined within a single numerical structure. Yet the concept attracts almost no support.

The proposition that multiplication requires a ratio measure performs even worse. Endorsement ranges from 0.10 to 0.20. This result strikes directly at the heart of the QALY. The QALY exists only because multiplication is assumed to be permissible. If multiplication requires ratio properties and utilities do not possess those properties, the QALY collapses. The weak endorsement of this statement indicates that the arithmetic foundations of the reference case are largely detached from the requirements of measurement.

Perhaps the most striking result concerns the proposition that there are only two classes of measurement relevant to therapy assessment: linear ratio measures for manifest attributes and Rasch logit ratio measures for latent attributes. Endorsement is essentially absent, ranging from 0.05 to 0.10. This result demonstrates not merely ignorance of Rasch measurement but a complete absence of engagement with the problem of latent attribute measurement itself.

The statement that measurement precedes arithmetic receives endorsement levels between 0.10 and 0.25. This may be the single most important finding in the entire table. The scientific revolution of the seventeenth century established that measurement is a prerequisite for arithmetic. Yet across Australia's leading HTA institutions this principle receives little support. Arithmetic has become detached from measurement. This is the defining characteristic of measurement inversion.

The final statement concerns latent traits. The proposition that the outcome of interest for latent attributes is possession of that attribute receives endorsement levels between 0.05 and 0.25. The implication is unmistakable. Australian HTA has little appreciation of latent variable theory and even less appreciation of Rasch measurement. Instead of measuring attribute possession, the dominant framework relies upon ordinal scores, utility algorithms and composite indices.

Taken together, these results point to a single conclusion. Australian HTA has become a discipline in which arithmetic dominates measurement. The PBAC and the leading research centers differ only in the degree to which they endorse this position. They are united by a common intellectual culture in which utilities, QALYs and simulation models are accepted without serious consideration of their measurement foundations.

The consequence is measurement inversion. Arithmetic routinely precedes measurement. Once this occurs, arithmetic chaos becomes inevitable. Health-state descriptions generate ordinal preferences. Ordinal preferences become utilities. Utilities become QALYs. QALYs become simulation outputs. Simulation outputs become cost-effectiveness claims. At no stage are the requirements of representational measurement established.

The most important feature of these results is not that one centre performs marginally better or worse than another. CAHE and Leeder occasionally display slightly weaker endorsement of false propositions, but the overall pattern remains unchanged. Every institution occupies the same intellectual space. Every institution accepts the central assumptions of the reference case. Every institution rejects, either explicitly or implicitly, the foundational requirements of measurement theory.

This is why the results should be interpreted as evidence of a systemic rather than an institutional problem. The issue is not the PBAC. The issue is not CHERE, CHE, MSPGH, HEG, CAHE, Leeder or Deakin. The issue is the Australian HTA knowledge base itself. These institutions collectively reveal a discipline that has spent four decades refining arithmetic while neglecting measurement.

The conclusion is unavoidable. The Australian HTA research community exhibits a pervasive pattern of measurement inversion. The false is endorsed and the true is neglected. Utilities are treated as measures, QALYs as ratio scales, simulation outputs as scientific claims and ordinal scores as quantitative outcomes. At the same time, unidimensionality, ratio measurement, measurement before arithmetic and Rasch measurement are largely ignored. This is not simply a methodological weakness. It is evidence that the reference-case paradigm has reached closure. The challenge for Australian HTA is therefore not reform but reconstruction. The future lies in

abandoning arithmetic without measurement and replacing it with a framework grounded in lawful measurement, empirical evaluation and falsification.

INEVITABLE ARITHMETIC CHAOS

It is impossible to find a subject area that prides itself on quantitative assessment where the presence of measure inversion is so widespread. The implications are substantial because the interrogation results reveal more than technical disagreement regarding utilities or QALYs. They demonstrate the presence of a closed analytical culture in which the standards necessary to interrogate arithmetic operations have effectively disappeared from the framework itself. The result is not simply methodological weakness but a condition of measurement inversion generating inevitable arithmetic chaos.

The significance of arithmetic chaos lies in its cumulative character. Once inadmissible numerical entities enter analytical structures, each successive stage inherits and compounds the original contradiction. Utilities create QALYs; QALYs populate Markov simulation models; simulation models generate cost-effectiveness ratios. Internally coherent outputs continue to be produced, yet the arithmetic operations themselves lack lawful quantitative meaning. The framework therefore survives institutionally while the scientific legitimacy of its outputs collapses.

The most disturbing feature is that the HTA assessment process was effectively constructed to generate arithmetic chaos. Once the principle that measurement must precede arithmetic disappeared from the analytical framework, there were no longer constraints governing which numerical entities could be combined within the decision model process. Arithmetic became detached from lawful measurement properties; any numerical construct could enter the analytical chain provided it contributed to generation of a cost-effectiveness estimate.

The starting point was the time trade-off technique, where valuations of health-state descriptions were transformed into utility scores presented as though they possessed ratio properties. Yet these utilities were, at best, composite ordinal structures masquerading as lawful quantitative measures. The prior question was never addressed: did the entities entering the analytical framework satisfy the requirements imposed by scales of measurement and the axioms of representational measurement?

More importantly, the correct starting point for therapy impact assessment, evaluation of unidimensional attributes was never considered. Manifest attributes requiring linear ratio measures and latent attributes requiring Rasch logit ratio measures were absent from the architecture itself. Instead, multidimensional health-state descriptions and preference scores became the foundation for arithmetic operations involving multiplication, aggregation and ratio construction.

From the perspective of representational measurement this was the first step toward arithmetic chaos. Once inadmissible utility structures were accepted as lawful quantitative entities, every subsequent stage of the decision model process inherited and compounded the original contradiction. Utilities generated QALYs; QALYs entered simulation models; simulation models produced cost-effectiveness ratios. Internally coherent outputs continued to be generated, but the arithmetic operations themselves were disallowed from the outset.

The fact that measurement inversion and arithmetic chaos remained unchallenged for more than four decades and largely remain unchallenged today has one unavoidable implication: the outputs of the cost-effectiveness modelling process lack lawful quantitative meaning. Apparently coherent numerical outputs may continue to be generated, yet the arithmetic operations underpinning those estimates fail the standards imposed by scales of measurement and the axioms of representational measurement.

A review of the publicly described curriculum and methodological outputs associated with Melbourne, Monash, Adelaide, Griffith and the Leeder Centre reveals remarkable convergence in analytical content. Across the five centers the dominant themes are utilities, QALYs, state-transition modelling, probabilistic sensitivity analysis, discounting procedures and cost-effectiveness ratios. Considerable emphasis is placed upon simulation methods, parameter uncertainty and economic evaluation for reimbursement decision making.

Equally striking is the consistency of what is absent. There is little or no explicit attention to scales of measurement, representational measurement, admissible arithmetic operations, dimensional homogeneity, unidimensionality or Rasch latent measurement. The distinction between manifest and latent attributes effectively disappears. Arithmetic operations are presented as methodologically legitimate without prior demonstration that the entities entering these operations possess lawful quantitative properties. The result is a curriculum structure in which arithmetic chaos is reproduced institutionally as accepted quantitative science.

Across the respective research centers students continue to be trained extensively in utilities, QALYs, simulation models, discounting procedures and ICER construction, while scales of measurement, representational measurement and Rasch latent measurement remain largely absent. The same analytical architecture is therefore reproduced across successive generations of researchers and practitioners. Closure becomes institutionalized because the standards necessary to challenge the framework are excluded from the curriculum itself.

The broader implication is unavoidable. The Australian HTA environment represents a microcosm of a wider international problem in which reference-case decision modelling evolved independently of the standards governing lawful quantitative inquiry. The result is a subject area that normalized inadmissible arithmetic operations while presenting itself as an established quantitative science. The interrogation results therefore provide more than evidence of isolated methodological problems. They reveal a structure of measurement inversion embedded at the center of HTA teaching, research and policy development.

CLOSURE AND TRANSITION

Closure has only had to wait for 40 years while generations of HTA practitioners have failed to be equipped with the conceptual and methodological tools necessary to interrogate the arithmetic foundations of the reference-case framework itself. Students and researchers were trained extensively in utilities, QALYs, simulation models and cost-effectiveness ratios, yet received little or no exposure to scales of measurement, representational measurement, admissible arithmetic operations or Rasch latent measurement. The consequence was inevitable: a professional culture in which arithmetic operations were accepted by convention rather than justified through lawful

measurement properties. Closure therefore persisted not because the contradictions were absent, but because the intellectual tools necessary to recognize and challenge them were excluded from the curriculum itself.

Recognition of closure immediately raises a practical question: what standards are required if Australian HTA is to transition from a framework defined by measurement inversion and arithmetic chaos toward one consistent with quantitative science? Criticism alone is insufficient. Once arithmetic detached itself from lawful measurement, closure became inevitable. The issue now is reconstruction. The following standards do not represent refinements to the existing reference-case architecture. They define the conditions necessary for replacement.

1. Measurement must precede arithmetic
This principle is the point from which all reconstruction begins. For decades HTA largely assumed that attaching numbers to observations automatically created quantitative evidence. Yet arithmetic operations cannot create measurement properties where they do not already exist. Multiplication, averaging and aggregation acquire meaning only after lawful measurement structures have been demonstrated. The failure to recognize this sequence created the conditions for measurement inversion itself.

2. Scale classification becomes a prerequisite for evidence generation
Not all numerical scales possess the same properties. Nominal, ordinal, interval and ratio scales support different forms of interpretation and different arithmetic operations. Historically these distinctions often disappeared within HTA practice as scores and utilities were treated as quantities without first asking what type of scale existed. Reconstruction requires explicit recognition of scale structure before claims can be advanced.

3. Ratio measures occupy a unique role in therapy evaluation
Only ratio measures support unrestricted arithmetic because they preserve meaningful quantitative relationships and possess identifiable origins. Under conventional HTA, ratio assumptions were frequently imposed upon structures lacking these properties. Future therapy claims therefore require lawful ratio measures rather than numerical approximations presented as quantities.

4. All therapy impact claims must concern attributes
Therapies do not act upon utilities, scores or simulation outputs. Therapies influence attributes within patients and populations. Hospital utilization, persistence, symptom burden and treatment satisfaction represent examples. Reconstruction therefore begins by asking a simple question largely absent from conventional HTA: what attribute is expected to change?

5. Manifest and latent attributes must be distinguished
One of the major consequences of measurement inversion was failure to distinguish between directly observed and inferred phenomena. Manifest attributes such as physician visits or hospitalization differ fundamentally from latent constructs such as symptom burden or need fulfilment.

6. Manifest attributes require linear ratio measures
Observable phenomena often possess natural quantitative structures. Hospital days, persistence behavior and switching rates provide examples where arithmetic may be meaningful. Manifest claims therefore require explicit ratio measurement structures. Quantitative interpretation follows from the properties of the attribute itself rather than numerical conventions imposed upon observations.

7. Latent attributes require Rasch logit ratio measures
Latent constructs create a different challenge. Responses to questionnaires are not measures simply because they have numbers attached to them. Scores are observations, not quantities. Reconstruction requires recognition that latent attributes demand Rasch procedures capable of creating lawful logit ratio structures and evaluating possession. The outcome of interest becomes possession of the latent attribute rather than change in arbitrary scores.

8. Attribute claims require explicit justification
Under conventional HTA, claims frequently emerged as outputs from analytical systems rather than propositions requiring prior explanation. Utilities generated QALYs and models generated cost-effectiveness estimates, yet the rationale for selecting particular outcomes often remained unclear. Reconstruction requires a different approach. Each attribute claim should provide an explicit justification explaining why therapy is expected to influence the attribute, what evidence supports this expectation and why the anticipated change possesses clinical relevance. Claims should therefore emerge from a coherent explanatory framework rather than analytical convenience or historical precedent.

9. Attribute claims require protocols
Protocols become the mechanism through which claims acquire scientific standing. Under reference-case systems assumptions frequently disappeared within simulation structures and remained insulated from direct challenge. Reconstruction requires every attribute claim to be linked prospectively to a protocol specifying the target population, observation period, expected treatment effects and evaluation procedures. Protocols therefore become scientific instruments rather than administrative documents. They establish the framework through which claims move from expectation to evidence.

10. Attribute claims must be empirically evaluable
Claims cannot remain hypothetical projections extending decades into uncertain futures. A claim acquires scientific meaning only when evidence capable of supporting or challenging it can be generated. Reconstruction therefore requires attribute claims to specify measurable endpoints, defined populations and observation schedules capable of prospective assessment. Evaluation becomes a continuing process rather than a one-time analytical exercise completed at product launch.

11. Falsification becomes a requirement
Scientific progress requires more than confirmation; it requires vulnerability to challenge. Attribute claims should therefore specify the conditions under which they may fail. If anticipated effects are absent or treatment outcomes diverge substantially from expectations, claims should

be modified or rejected. Exposure to challenge replaces protection of assumptions. Claims incapable of failure cease to function as scientific propositions.

12. HTA returns to the standards of normal science

Taken together these standards do not create a new ideology. They restore principles long familiar within quantitative science: measurable attributes, lawful measurement structures, explicit claims and continuing empirical challenge. Numerical storytelling gives way to measurable propositions. The objective is straightforward: reconnect HTA with the continuing evolution of objective knowledge. HTA therefore re-enters a framework where evidence grows through replication, criticism and empirical learning rather than through increasingly elaborate assumptions.

CONCLUSION

The findings presented here point to a clear conclusion. Australian HTA evolved around utilities, QALYs, reference-case methodologies and simulation models without first establishing that the numerical entities entering these analytical structures satisfied the requirements imposed by scales of measurement and representational measurement. The result was a stable pattern of measurement inversion in which arithmetic routinely preceded lawful measurement.

The implications are unavoidable. Once arithmetic becomes detached from measurable attributes, arithmetic chaos follows inevitably. Utilities lacking lawful ratio properties generate QALYs; QALYs populate simulation models; simulation models produce cost-effectiveness ratios. Apparently coherent numerical outputs continue to be generated, yet the arithmetic operations themselves lack lawful quantitative meaning.

The interrogation findings demonstrate that this structure is not confined to PBAC but reproduced across Australian HTA teaching and research environments. Foundational propositions concerning unidimensionality, ratio measurement and representational measurement receive weak endorsement, while propositions supporting QALYs, utilities and simulation models receive strong endorsement. The result is a closed analytical culture where internally generated assumptions substitute for empirical scientific challenge.

Closure therefore means closure. The issue is not whether HTA agencies or simulation models continue to exist. The issue is that the reference-case framework no longer possesses the capacity for correction from within its own assumptions because the standards necessary to interrogate its arithmetic foundations are absent from the framework itself.

The choice is increasingly clear: continue to defend a closed system of numerical storytelling or return HTA to the standards of normal science through measurable attributes, lawful ratio measures, explicit protocols and falsifiable claims. Measurement inversion guaranteed the present outcome decades ago. The future of HTA therefore lies not in reforming the reference case but in abandoning it.

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] United Sates: COSMIN and measurement inversion – The certification of non-measures. Logit Working Paper No. 237 , https://maimonresearch.com/logit-working-paper-no-237-january-2026/

[2] United States: Cochrane and measurement inversion – the aggregation of non-measures, Logit Working Paper No 238. https://maimonresearch.com/logit-working-paper-no-238-january-2026/

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

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

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

[6] Australia: The absence of representational measurement and the Health Economics Goup (HEG) University of Adelaide. https://maimonresearch.com/logit-working-paper-no-723-february-2026/

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

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

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

[10] Australia: The absence of representational measurement and the Deakin University Health Economics and HTA knowledge base. Logit Working Paper No 282. https://maimonresearch.com/logit-working-paper-no-281-may-2026/