MAIMON RESEARCH LLC

ARTIFICIAL INTELLIGENCE LARGE LANGUAGE MODEL INTERROGATION

REPRESENTATIONAL MEASUREMENT FAILURE IN HEALTH TECHNOLGY ASSESSMENT

NEW ZEALAND: THE MISSING SCIENCE OF MEASUREMENT IN HEALTH TECHNOLOGY ASSESSMENT

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

LOGIT WORKING PAPER No 1277 JUNE 2026

www.maimonresearch.com

Tucson AZ

ABSTRACT

This paper reports the results of a curriculum interrogation of the health technology assessment (HTA) knowledge bases associated with PHARMAC, the University of Auckland and the University of Otago. The objective was to determine whether the foundational concepts required for lawful quantitative claims are represented within the educational environment that supports HTA teaching, research and professional development in New Zealand. A series of ten canonical curriculum statements was developed to assess coverage of attributes, scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement and falsifiable claims. These concepts represent the minimum requirements for a measurement-based approach to therapy assessment.

The results reveal a pervasive pattern of curriculum inversion. While there is moderate recognition of the need to identify target attributes and an appreciation that therapy impact claims should be falsifiable, the concepts necessary to establish whether quantitative claims are scientifically possible are largely absent. Endorsement probabilities for scales of measurement, measurement status, representational measurement, unidimensionality, latent attribute measurement and distinct ratio measurement frameworks for manifest and latent attributes are consistently low. The findings suggest that students and researchers are exposed to analytical techniques while receiving little systematic instruction in the scientific principles required to evaluate the legitimacy of those techniques.

The implications extend beyond curriculum design. The absence of measurement science helps explain the widespread pattern of measurement inversion observed in previous interrogations of HTA agencies, research centres, professional organisations and academic programs internationally. Students are trained to apply utilities, quality-adjusted life years (QALYs) and reference-case simulation models without being introduced to the measurement requirements that these constructs presuppose. The result is an educational environment that reproduces rather than challenges the conceptual foundations of contemporary HTA.

The paper argues that the future of HTA in New Zealand depends upon reconstruction rather than reform. Reconstruction must begin with the reintroduction of measurement science into professional education, including scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement and Rasch measurement. Only by restoring the principle that measurement precedes arithmetic can HTA transition from numerical construction and simulation-based claims to a framework built upon evaluable, replicable and falsifiable evidence. The curriculum interrogation therefore identifies not merely an educational deficiency but a fundamental obstacle to the development of HTA as a scientific discipline.

INTRODUCTION

Health technology assessment (HTA) presents itself as a quantitative discipline. Students are introduced to economic evaluation, cost-effectiveness analysis, utility assessment, simulation modelling and reimbursement decision making. These methods generate numerical claims regarding therapy impact and are routinely employed to support formulary recommendations, coverage decisions and resource allocation. The legitimacy of such claims, however, depends upon a prior question that is rarely asked: do the quantities entering these analyses satisfy the requirements of measurement? This is the critical question for the HTA content of New Zealand research centers and PHARMAC program content and their role is establishing professional standards.

In every quantitative science, measurement precedes arithmetic. Before quantities can be added, multiplied, transformed or modelled, their measurement properties must first be established. This principle is not optional. It is the foundation upon which quantitative reasoning rests. Without measurement there can be no meaningful arithmetic, and without lawful arithmetic there can be no credible quantitative claims.

The requirements of measurement are well understood. They include the distinction between nominal, ordinal, interval and ratio scales the role of representational measurement; the requirement for unidimensional attributes; the distinction between manifest and latent attributes; the unique role of Rasch logit ratio measurement and the conditions under which quantities may support arithmetic operations such as multiplication and division [1] [2] [3] [4] . These concepts define the scientific standards against which quantitative claims must be judged.

Recent large language model (LLM) interrogations of HTA knowledge bases across multiple countries have revealed a pervasive pattern of measurement inversion [5] . Concepts central to measurement science receive little recognition, while analytical frameworks that depend upon those concepts receive strong endorsement. Utilities are treated as though they possess interval or ratio properties. QALYs are accepted as quantitative measures despite the absence of demonstrated ratio-scale foundations. Simulation models generate lifetime claims without establishing the measurement status of the quantities entering those models. The same pattern appears repeatedly across agencies, academic centers, professional organizations and HTA teaching programs. New Zealand is no different [6] [7]

This raises an obvious question. If measurement inversion is so widespread within HTA practice, where does it originate? One possibility is that the concepts necessary to recognize and challenge measurement inversion are absent from the educational environment that trains future researchers and practitioners. If students are not introduced to measurement theory, representational measurement, unidimensionality, latent attribute measurement and ratio scales, then the persistence of measurement inversion should not be surprising.

The purpose of this paper is to evaluate the curriculum knowledge base supporting New Zealand research center HTA programs and PHARMAC. The assessment focuses not on individual courses or instructors but on the broader intellectual environment that shapes teaching, research and professional development. A series of canonical statements was developed to represent the minimum concepts required for a measurement-based approach to therapy assessment. These statements address attributes, scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement and falsifiable claims.

The curriculum interrogation reveals two distinct but related problems not specific to New Zealand. The first is one of omission. Concepts central to measurement science, including scales of measurement, representational measurement, unidimensionality, latent attribute measurement, Rasch measurement and ratio measurement, are largely absent from curriculum coverage. Students are therefore unlikely to be exposed to the scientific principles necessary to evaluate whether quantitative claims satisfy the requirements of measurement.

The second problem is more serious. It is not merely that essential concepts are absent; it is that students are taught analytical frameworks that depend upon assumptions which fail those same standards. Utilities, QALYs and reference-case simulation models are presented as legitimate instruments of quantitative assessment despite their inability to satisfy the requirements of representational measurement. Students are therefore not only deprived of the concepts needed to evaluate quantitative claims they are simultaneously introduced to frameworks that cannot be defended once those concepts are applied.

The result is a self-reinforcing cycle. The absence of measurement science prevents students from recognizing the limitations of the methods they are taught, while the methods themselves reinforce the perception that quantitative legitimacy can be achieved through arithmetic and modelling alone. Curriculum inversion and measurement inversion therefore become mutually supporting phenomena. One explains the persistence of the other.

THE IMPERATIVE OF MEASUREMENT INVERSION

The starting point for any reconstruction of health technology assessment (HTA) in New Zealand must be an understanding of the scale and consistency of the false measurement inversion that characterizes the field. Over the past several months, a series of large language model (LLM) interrogations has been undertaken to evaluate HTA knowledge bases across multiple jurisdictions. These interrogations have encompassed national reimbursement agencies, academic research centers, professional organizations, journals, pharmacy schools and HTA teaching programs in Australia, Canada, New Zealand, the United Kingdom, Europe, Singapore and the United States. Although the institutions differ in structure, mission and geographical location, the results have been remarkably consistent. The same pattern of measurement inversion appears irrespective of country, discipline or organizational setting.

The concept of measurement inversion is straightforward. In the physical and social sciences, measurement precedes arithmetic. Before quantities can be added, multiplied, transformed or modelled, the measurement properties of those quantities must first be established. HTA reverses this sequence. Numerical operations are routinely undertaken without demonstrating that the quantities involved satisfy the requirements of measurement. Arithmetic is treated as a substitute for measurement rather than a consequence of measurement. The result is a framework in which numerical manipulation takes precedence over the validation of the quantities being manipulated.

The LLM interrogations reveal this inversion repeatedly. Across institutions there is little awareness of scales of measurement, the distinction between ordinal, interval and ratio scales, the role of unidimensionality, the axioms of representational measurement, dimensional homogeneity, or the distinction between manifest and latent attributes. At the same time, there is widespread endorsement of analytical frameworks that depend upon precisely those concepts. Utilities are treated as though they possess interval or ratio properties. Quality-adjusted life years are accepted as quantitative measures despite the absence of demonstrable ratio-scale foundations. Reference-case simulation models are employed to generate lifetime cost-effectiveness claims without establishing the measurement properties of the underlying inputs. The pattern is universal.

The significance of these findings lies not merely in the demonstration of measurement failure but in the realization that measurement inversion is now institutionalized. It is embedded in teaching programs, methodological guidance, research publications and policy frameworks. New Zealand is no exception. Researchers entering the field encounter an intellectual environment in which the requirements of measurement are rarely discussed and almost never presented as prerequisites for quantitative claims. Consequently, the acceptance of measurement inversion becomes self-reinforcing. Successive generations of researchers inherit analytical frameworks without being introduced to the measurement principles required to evaluate them.

It is worth noting the complete absence of Rasch measurement from the curriculum profiles of the New Zealand HTA research centers. This omission is particularly significant because Rasch measurement represents the only established framework that provides the necessary and sufficient conditions for transforming observations relating to a latent attribute into a measure. The central issue is not statistical sophistication but measurement itself. Latent attributes such as symptom burden, functional status, treatment satisfaction, need fulfilment and quality of life cannot be observed directly. They require a measurement model capable of estimating possession of the attribute while demonstrating unidimensionality, invariance and appropriate scale properties. Rasch measurement was developed specifically to meet these requirements [8]. Its absence suggests that faculty and students are not exposed to the concept of latent attribute possession, nor to the scientific challenge of constructing measures from observations. Instead, there is a tendency to treat questionnaire scores, indexes and composite summaries as though they were measures in their own right.

The result is that one of the most important developments in measurement theory over the past 80 years is effectively invisible within New Zealand HTA education. Without exposure to Rasch measurement, faculty and students are unlikely to appreciate that latent attributes require a fundamentally different approach to measurement from manifest attributes, or that quantitative claims regarding latent phenomena require a demonstrable measurement framework before arithmetic operations can be justified. This observation raises an obvious question. If measurement inversion is so widespread and persistent, where does it originate? The answer cannot be found solely in research practice or policy guidance. Attention must also be directed toward education. If the concepts necessary to recognize measurement failure are absent from the curriculum, then measurement inversion becomes the expected outcome rather than an isolated error. Researchers cannot be expected to evaluate measurement claims if they have never been introduced to scales of measurement, representational measurement, ratio scales, unidimensionality, or the distinction between manifest, latent attributes and even Rasch measurement.

For this reason, curriculum assessment emerges as a critical component of HTA reconstruction. The objective is not simply to determine whether students are exposed to contemporary HTA methods. Rather, it is to determine whether they are exposed to the foundational concepts that make the evaluation of those methods possible. A curriculum that emphasizes modelling, economic evaluation and decision analysis while neglecting measurement theory will inevitably reproduce the same conceptual limitations observed in current HTA practice.

The curriculum interrogations undertaken across New Zealand HTA research centers and PHARMAC provide compelling support for this interpretation. While there is evidence that students and researchers are introduced to outcomes assessment, target attributes and scientific claims, there is little evidence of systematic exposure to scales of measurement, the axioms of representational measurement, unidimensionality, latent attribute measurement or ratio measurement. The concepts most frequently absent from curriculum coverage are precisely those concepts most frequently absent from HTA practice. The relationship is unlikely to be coincidental.

The imperative of measurement inversion therefore extends beyond criticism of existing methods. It points directly to the need for educational reconstruction. If HTA is to move toward a framework based on lawful measurement, evaluable claims and empirical falsification, then curriculum reform must accompany methodological reform. The widespread and consistent pattern of measurement inversion revealed by the LLM interrogations suggests that reconstruction cannot begin with policy guidance or analytical techniques alone. It must begin with the curriculum. Until students and researchers are introduced to the foundations of measurement science, the conditions that created measurement inversion will continue to be reproduced throughout the HTA community.

DEFINING THE KNOWLEDGE BASE

The first step in any LLM interrogation is to define the knowledge base to be interrogated. The validity of the interrogation depends upon ensuring that the knowledge base reflects the information environment that shapes teaching, research and professional practice. In the present assessment, the curriculum knowledge base for each research center is defined as the totality of publicly accessible materials that contribute to HTA education and training. These include curriculum descriptions, course outlines, program objectives, teaching materials, methodological guidance documents, seminar and workshop content, faculty publications, doctoral training resources, research center reports, conference presentations, policy briefs and other materials through which knowledge is communicated to students, researchers and professional staff.

The objective is not to evaluate individual courses or instructors but to assess the broader intellectual environment within which HTA concepts are introduced, reinforced and transmitted. The resulting curriculum knowledge base is assumed to represent the concepts and principles that students and researchers are most likely to encounter during their exposure to the research center and its associated educational activities. It is this knowledge base that is interrogated to determine the extent to which the foundational concepts required for a measurement-based approach to therapy assessment are present, absent or only weakly represented.

INTERROGATING THE CURRICULUM

The identification of measurement inversion across New Zealand HTA research centers, reimbursement agencies and academic programs raises an obvious question: where does this inversion originate? If the same conceptual failures are observed repeatedly across institutions and jurisdictions, then the explanation cannot rest solely with individual researchers, policy analysts or decision makers. A more plausible explanation is that these failures reflect deficiencies in the educational environment that prepares future HTA practitioners. If concepts central to measurement science are absent from curriculum content, then their absence from research practice should not be surprising. This realization provides the rationale for interrogating the curriculum.

The objective of curriculum interrogation differs from that of previous HTA knowledge-based assessments. Earlier interrogations focused on whether institutions recognized the requirements of representational measurement and the standards necessary for quantitative claims. Curriculum interrogation asks a different question. Are faculty, students and researchers exposed to the concepts necessary to understand and apply those standards? The focus shifts from methodological outputs to educational inputs. Rather than examining what faculty, students and researchers do, attention is directed to what they are taught and what they know.

The importance of this distinction should not be underestimated. Educational programs do not merely transmit technical skills. They define the conceptual framework through which future practitioners understand evidence, measurement and scientific inquiry. Concepts that are absent from the curriculum are unlikely to emerge spontaneously in research practice. Equally, concepts that are emphasized repeatedly become part of the intellectual assumptions that shape subsequent analysis have never been systematically incorporated into HTA teaching and research training.

For this reason, the curriculum interrogation was designed around a series of canonical statements intended to identify the presence or absence of foundational measurement concepts. These statements were deliberately elementary. The purpose was not to assess advanced methodological knowledge but to determine whether faculty, students and researchers are likely to encounter the principles that underpin lawful quantitative claims. The resulting framework begins with the concept of an attribute as the object of measurement and proceeds through target attribute specification, scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement and falsifiable claims. Together, these statements define the minimum intellectual foundations required for a measurement-based approach to therapy assessment in education.

These statements are:

- **An attribute is the specific outcome of interest in a therapy assessment.**
- **Every therapy assessment begins with specification of the target attribute.**
- **The principal scales of measurement (nominal, ordinal, interval and ratio) have different properties and support different forms of analysis.**
- **The measurement status of a target attribute must be established before quantitative claims can be advanced.**
- **The axioms of representational measurement underpin quantitative claims.**

- **Attributes must be demonstrated to be unidimensional before measurement is possible.**
- **A manifest attribute is directly observable and capable of supporting empirical observation.**
- **A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute.**
- **Manifest and latent attributes require different forms of ratio measurement.**
- **Therapy impact claims must be falsifiable.**

These ten statements form a logical sequence:

Attribute → Target Attribute → Scales of Measurement → Measurement Status → Representational Measurement → Unidimensionality → Manifest Attribute → Latent Attribute → Ratio Measurement → Falsifiable Claims

Together they define the minimum curriculum content required for a measurement-based approach to HTA and provide the framework for evaluating curriculum coverage in New Zealand HTA research centers and PHARMAC.

The categorical probabilities reported in this assessment are intended as indicators of the extent to which a concept is represented within the curriculum knowledge base. They should not be interpreted as precise statistical estimates but as measures of the likelihood that a student, researcher or professional exposed to that knowledge base would encounter, recognize and subsequently endorse the canonical statement. In practical terms, the probability reflects the visibility and prominence of a concept within the educational environment associated with a research center or policy agency.

A high probability indicates that the concept is well represented within curriculum materials, research outputs and educational activities and is therefore likely to be familiar to students and researchers. Conversely, a low probability suggests that the concept is absent, only weakly represented, or occupies a peripheral position within the curriculum knowledge base. Students exposed to such an environment would therefore be unlikely to recognize the concept as an important component of HTA education and practice.

The probabilities should be viewed comparatively rather than in isolation. Their principal value lies in identifying patterns of curriculum coverage across institutions and concepts. In particular, low probabilities associated with scales of measurement, representational measurement, unidimensionality and ratio measurement indicate that these topics are unlikely to form a substantial part of the educational experience of the average student. The resulting profile provides an indication of curriculum strengths, deficiencies and potential areas for reconstruction.

CURRICULUM INVERSION IN NEW ZEALAND HTA

The New Zealand curriculum results are not merely disappointing (Table 1). They are alarming. Taken together they suggest that the intellectual foundations necessary to evaluate quantitative claims are almost entirely absent from New Zealand HTA education. The issue is not whether particular topics receive insufficient attention. The issue is that the scientific foundations upon which quantitative claims depend appear to be largely missing.

TABLE 1: CURRICULUM CONTENT ENDORSEMENT: PHARMAC, UNIVERSITY OF AUCKLAND, UNIVERSITY OF OTAGO

CATEGORICAL PROBABILITY			
CANONICAL STATEMENT	PHARMAC	AUCKLAND	OTAGO
An attribute is the specific outcome of interest in a therapy assessment	0.75	0.75	0.80
Every therapy assessment begins with specification of the target attribute	0;65	0.65	0.60
The principal scales of measurement (nominal, ordinal, interval and ratio) have different properties and support different forms of analysis	0.20	0.20	0.20
The measurement status of a target attribute must be established before quantitative claims can be advanced	0.15	0.15	0.15
The axioms of representational measurement underpin quantitative claims	0.05	0.05	0.05
Attributes must be demonstrated to be unidimensional before measurement is possible	0.10	0.10	0.10
A manifest attribute is directly observable and capable of supporting empirical observation	0.40	0.40	0.35
A latent attribute is not directly observable and requires a measurement model to estimate possession of the attribute	0.10	0.10	0.10
Manifest and latent attributes require different forms of ratio measurement	0.05	0.05	0.05
Therapy impact claims must be falsifiable	0.80	0.65	0.65

As the first consideration, we can look at the patterns of categorical probabilities reported for each statement.

Statement 1: An Attribute is the Specific Outcome of Interest in a Therapy Assessment

The endorsement probability of 0.75 across PHARMAC, the University of Auckland and the University of Otago indicates that the New Zealand HTA curriculum knowledge base recognizes the importance of identifying an outcome of interest in therapy assessment. This is a necessary starting point for any scientific inquiry. Before an intervention can be evaluated, there must be clarity regarding the attribute that is expected to change. Whether the concern is survival, symptom burden, functional status, treatment adherence or quality of life, the object of inquiry must first be identified.

At one level, therefore, the result appears reassuring. The curriculum demonstrates an awareness that therapy assessment requires a clearly defined target. This reflects standard research practice where the first task is to specify what is to be investigated. Without an attribute of interest there can be no meaningful evaluation.

The difficulty is that identification of an attribute is only the beginning of the scientific process. An attribute is not a measure. Naming an outcome does not establish that the outcome can be measured, nor does it establish the scale properties necessary to support quantitative claims. The identification of an attribute merely defines what is to be explained. It does not demonstrate that explanation can proceed through measurement.

This distinction is critical because the subsequent endorsement profile indicates that the curriculum largely stops at attribute identification. Once the target outcome has been specified, the scientific questions that must follow receive little attention. The principal scales of measurement receive only weak endorsement. The requirement that the measurement status of an attribute be established before quantitative claims are advanced receives very weak endorsement. Representational measurement, unidimensionality, latent attribute measurement and ratio measurement all receive minimal support.

Taken together, these results suggest that students are introduced to the concept of an attribute but are not systematically introduced to the scientific requirements that determine whether the attribute can support measurement. The curriculum therefore creates a potentially misleading impression that identifying an outcome largely completes the scientific task. Once an attribute has been named, attention shifts directly to analysis, modelling and evaluation. The much more demanding question of whether the attribute can be measured is rarely addressed.

The omission has profound consequences. Every attribute immediately raises a series of scientific questions. Is the attribute manifest or latent? Can it be observed directly or must it be inferred through a measurement model? Is it unidimensional? What empirical evidence supports its existence as a coherent attribute? What scale properties can be demonstrated? Does it support ordinal, interval or ratio measurement? Are arithmetic operations admissible? These questions determine whether quantitative claims are possible. Yet the endorsement profile suggests that they occupy only a marginal position within the New Zealand HTA curriculum.

The result is that attributes are routinely accepted as candidates for quantitative analysis without first establishing whether they satisfy the requirements of measurement. Numerical representation becomes a matter of convention rather than demonstration. Quantification follows automatically from identification. The scientific burden of proving measurement quietly disappears.

This may represent one of the principal origins of measurement inversion within HTA. Once identification of an attribute is mistaken for measurement of an attribute, the door is opened to increasingly elaborate forms of arithmetic, economic evaluation and simulation modelling based upon quantities whose measurement status has never been established. Numerical outputs acquire scientific authority because they are numerical, not because they are measures.

The significance of this finding should not be underestimated. The weakness of the curriculum is not that students fail to recognize the importance of attributes. The weakness is that they are not systematically taught the scientific principles that determine whether attributes can be measured. Identification becomes a substitute for measurement. The result is an educational framework in which measurement is assumed rather than demonstrated. This assumption underpins much of contemporary HTA and provides an important explanation for the persistence of measurement inversion in both research and policy.

Statement 2: Every Therapy Assessment Begins with Specification of the Target Attribute

The endorsement probabilities for this statement range from 0.60 to 0.65 across PHARMAC, the University of Auckland and the University of Otago. This indicates that the New Zealand HTA curriculum knowledge base recognizes that therapy assessment must begin with a clearly defined outcome of interest. Such recognition is entirely appropriate. Scientific inquiry requires a defined object of investigation. Without a target attribute there can be no meaningful assessment of therapy impact.

At first sight this appears to reinforce the positive finding observed in Statement 1. Students are introduced not only to the concept of an attribute but also to the principle that evaluation requires a clearly specified target. The curriculum therefore demonstrates an awareness of the need for explicit research objectives and outcome definition.

The difficulty is that specification of a target attribute is frequently treated as though it were equivalent to establishing a measurable quantity. It is not. Defining an outcome merely identifies what is to be investigated. It does not establish whether the attribute possesses the properties necessary to support measurement. Naming an attribute and measuring an attribute are entirely different scientific activities.

This distinction is fundamental because measurement begins where outcome specification ends. Once an attribute has been identified, a series of critical questions must immediately follow. Is the attribute manifest or latent? Is it directly observable or does it require a measurement model? Is the attribute unidimensional? What empirical evidence supports its existence as a coherent attribute? What scale properties can be demonstrated? Does the attribute support ordinal, interval or ratio measurement? Are arithmetic operations permissible? These questions determine whether quantitative claims concerning the attribute are scientifically possible.

The endorsement profile suggests that this intermediate stage is largely absent from the curriculum. Students appear to be taught how to specify outcomes but receive little systematic exposure to the scientific requirements that must be satisfied before those outcomes can support measurement. The subsequent endorsement probabilities provide compelling evidence. Scales of measurement receive weak endorsement. The requirement to establish measurement status before advancing quantitative claims receives very weak endorsement. Representational measurement, unidimensionality, latent attribute measurement and ratio measurement receive almost no support.

The consequence is that specification of the target attribute risks becoming the endpoint rather than the starting point of scientific inquiry. Once an outcome has been identified, quantitative analysis is assumed to be possible. Measurement becomes an implicit assumption rather than an explicit demonstration. Numerical methods are introduced before the conditions necessary to support numerical claims have been established.

This omission has important implications for professional training. Students may become proficient in analytical techniques while never being exposed to the scientific principles required to determine whether the quantities entering those analyses are measurable. They learn how to manipulate numbers but not how to evaluate the legitimacy of those numbers. The distinction is crucial because quantitative sophistication cannot compensate for measurement failure. Complex analyses cannot transform non-measures into measures.

The result is a curriculum framework in which the specification of an outcome is treated as sufficient justification for quantification. Once the target attribute has been identified, the analytical process proceeds as though measurement has already been achieved. The scientific burden of demonstrating measurement quietly disappears from view.

This finding points directly to one of the central weaknesses of contemporary HTA education. The curriculum recognizes the importance of defining what is to be assessed but pays little attention to the conditions under which assessment can generate valid quantitative claims. Students are therefore trained to think about outcomes without being trained to think about measurement.

The broader consequence is the reproduction of measurement inversion. Quantitative claims are advanced before the measurement properties of the underlying attributes have been established. Arithmetic follows attribute specification rather than measurement demonstration. The result is an educational environment in which numerical outputs acquire authority through convention rather than through compliance with the standards of representational measurement.

The significance of this finding should not be underestimated. The issue is not whether students learn to define outcomes. The issue is that they are not systematically taught what must happen next. Without an understanding of measurement theory, outcome specification becomes detached from the scientific requirements necessary to support quantitative claims. In this sense, the curriculum does not merely omit measurement. It normalizes the assumption that measurement can be taken for granted.

Statement 3: The Principal Scales of Measurement Have Different Properties and Support Different Forms of Analysis

The endorsement probability for this statement is only 0.20 across PHARMAC, the University of Auckland and the University of Otago. This is one of the most important findings in the entire curriculum interrogation because it concerns the foundation upon which all quantitative reasoning depends. If students are not introduced to the differences among the principal scales of measurement, they lack the conceptual framework necessary to determine which forms of arithmetic and analysis are scientifically admissible.

The proposition itself is neither controversial nor advanced. It represents one of the most elementary principles of measurement theory. Nominal, ordinal, interval and ratio scales possess different mathematical properties and therefore support different analytical operations. Nominal scales support classification. Ordinal scales support ranking. Interval scales support addition and subtraction but lack a true zero. Ratio scales alone support multiplication, division and proportional reasoning because they possess a fixed non-arbitrary zero.

The importance of these distinctions cannot be overstated. Arithmetic operations are not universally applicable. The legitimacy of an analysis depends directly upon the measurement properties of the quantities being analyzed. Before any arithmetic operation can be performed, the scale properties of the underlying attribute must first be established. Measurement precedes arithmetic.

The low endorsement of this statement suggests that the New Zealand HTA curriculum knowledge base does not regard scale theory as a central component of professional training. Students may therefore be introduced to statistical methods, economic evaluation and simulation modelling without first being taught the conditions under which the arithmetic employed by those methods is legitimate. The result is an educational framework in which analytical techniques occupy a more prominent position than the measurement principles that justify their application.

This omission has profound implications. Consider the role of utility scores in HTA. Utility values are routinely multiplied by survival time to generate quality-adjusted life years. Yet multiplication is only admissible if the quantities entering the calculation possess the scale properties necessary to support proportional reasoning. Without an understanding of ratio measurement, students have no basis for evaluating whether such operations are scientifically defensible. The arithmetic is accepted as routine because the measurement question is never asked.

The same problem extends throughout contemporary HTA. Cost-effectiveness ratios, utility-weighted survival estimates, preference scores and simulation model outputs all depend upon assumptions regarding scale properties. If students are not taught the distinctions among nominal, ordinal, interval and ratio scales, they are effectively deprived of the conceptual tools required to evaluate the legitimacy of these constructs. Numerical outputs become accepted on authority rather than assessed through measurement principles.

The consequences extend beyond particular analytical methods. Scale theory provides the gateway to representational measurement itself. It introduces students to the idea that numbers do not automatically constitute measures and that different numerical assignments support different forms of inference. Without this foundation, quantitative analysis risks becoming an exercise in numerical manipulation detached from the scientific requirements that give numbers meaning.

The weak endorsement of this statement therefore points to a critical deficiency in curriculum coverage. Students are unlikely to be systematically exposed to one of the most fundamental principles of quantitative science. The result is that arithmetic becomes detached from measurement. Analytical sophistication replaces measurement validation. Numerical techniques are learned without an understanding of the conditions necessary for their legitimate application.

This finding may represent one of the clearest explanations for the persistence of measurement inversion within HTA. Measurement inversion occurs when arithmetic is undertaken before the measurement status of the quantities involved has been established. If students are not taught that different scales support different forms of analysis, they are unlikely to recognize when arithmetic operations exceed the properties of the underlying data. The distinction between lawful and unlawful arithmetic effectively disappears.

The significance of this result should not be underestimated. A curriculum that gives only marginal attention to the principal scales of measurement cannot adequately prepare students to evaluate quantitative claims. The issue is not the absence of a technical detail. It is the absence of one of the foundational principles upon which quantitative science is built. Without an understanding of scale theory, students are left without the intellectual tools necessary to distinguish measurement from numerical construction. The result is a curriculum environment in which measurement inversion becomes not merely possible but inevitable.

Statement 4: The Measurement Status of a Target Attribute Must be Established Before Quantitative Claims Can be Advanced

The endorsement probability for this statement is only 0.15 across PHARMAC, the University of Auckland and the University of Otago. This is one of the most consequential findings in the curriculum interrogation because it concerns a principle that lies at the heart of all quantitative science. Before numerical claims can be advanced, it must first be demonstrated that the attribute concerned possesses the properties necessary to support measurement. Without measurement, quantitative claims have no scientific foundation.

The proposition itself is straightforward. Scientific inquiry does not begin with arithmetic, statistical analysis or modelling. It begins with the establishment of measurement. Before numbers can be manipulated, researchers must demonstrate that the attribute being investigated can be measured and identify the scale properties associated with that measurement. Quantitative reasoning therefore depends upon a prior measurement justification.

The low endorsement of this statement suggests that this principle occupies only a marginal position within the New Zealand HTA curriculum knowledge base. Students may be introduced to outcomes research, economic evaluation and modelling techniques without being systematically taught that the measurement status of an attribute must first be established. The implication is that numerical analysis becomes detached from the scientific conditions necessary to support it.

This omission is particularly important because the distinction between an attribute and a measure is often overlooked. Identifying an attribute does not establish measurement. Specifying an outcome does not demonstrate that quantitative claims concerning that outcome are possible. Before a quantity can enter an analysis, it must first be shown that the attribute possesses the structure necessary to support numerical representation. This is the central concern of measurement theory.

The consequences of ignoring this principle are profound. If measurement status is not established before analysis begins, arithmetic operations are applied to quantities whose scientific legitimacy remains unknown. Statistical sophistication may increase, simulation models may become more elaborate and numerical outputs may become more detailed, yet none of these developments address the prior question of whether the quantities involved are measures. Analytical complexity cannot compensate for measurement failure.

The importance of this issue becomes immediately apparent in HTA. Utility scores, quality-adjusted life years, preference weights and simulation model outputs are routinely treated as quantitative entities. Yet the scientific legitimacy of these constructs depends entirely upon the measurement status of the attributes they claim to represent. If measurement has not been demonstrated, subsequent arithmetic becomes a matter of assumption rather than scientific justification.

The weak endorsement of this statement suggests that students are unlikely to be systematically exposed to this problem. Quantitative methods are introduced without an accompanying framework for evaluating whether the quantities entering those methods satisfy the requirements of measurement. The result is an educational environment in which numerical outputs are often accepted because they are numerical rather than because their measurement foundations have been established.

This finding points directly to the phenomenon of measurement inversion. Measurement inversion occurs when arithmetic precedes measurement; when numerical operations are undertaken before it has been demonstrated that the underlying quantities support those operations. The endorsement profile suggests that students are not routinely taught to regard measurement as a prerequisite for quantitative claims. Instead, measurement becomes an implicit assumption embedded within the analytical process.

The implications extend beyond individual methods or specific analytical techniques. Once the requirement to establish measurement status disappears, there is no longer a principled basis for distinguishing measures from numerical constructions. Any quantity can potentially be incorporated into an analysis provided it can be assigned a number. The scientific discipline imposed by representational measurement is lost.

The significance of this result should not be underestimated. The issue is not a minor curriculum omission but the absence of one of the fundamental requirements of quantitative science. A curriculum that does not emphasize the necessity of establishing measurement status before advancing quantitative claims leaves students without the conceptual tools needed to evaluate the legitimacy of the analyses they are taught to perform.

The result is an educational framework in which quantitative claims are routinely advanced without prior measurement justification. Measurement becomes assumed rather than demonstrated. Arithmetic takes precedence over measurement. In this sense, the weak endorsement of this statement provides one of the clearest indications that curriculum inversion is present within the New Zealand HTA educational environment and helps explain the persistence of measurement inversion throughout HTA research and policy.

Statement 5: The Axioms of Representational Measurement underpin quantitative claims

The endorsement probability for this statement is only 0.05 across PHARMAC, the University of Auckland and the University of Otago. This is the lowest level of endorsement observed in the curriculum interrogation and arguably the most important. Representational measurement provides the scientific foundation upon which all quantitative claims depend. A curriculum that gives little attention to representational measurement deprives students of the conceptual framework necessary to determine whether numerical claims are scientifically legitimate.

The proposition itself concerns one of the central achievements of modern measurement theory. Representational measurement is concerned with the conditions under which numbers may be assigned to attributes in a manner that preserves empirical relationships. Measurement is not the arbitrary assignment of numbers. It is the construction of numerical representations that reflect the structure of the attribute being measured. The purpose of representational measurement is therefore to determine whether quantitative claims are possible and, if so, what forms of arithmetic are admissible.

The significance of this principle extends throughout every quantitative science. The physical sciences, engineering and the social sciences all depend upon measurement frameworks that preserve empirical structure. Quantitative claims are meaningful only because the underlying measurement system has first been established. Arithmetic follows measurement; it does not create it.

The endorsement probability of 0.05 suggests that representational measurement occupies almost no position within the New Zealand HTA curriculum knowledge base. Students may be introduced to statistical analysis, economic evaluation, utility assessment and simulation modelling without first being introduced to the scientific principles that determine whether the quantities entering those analyses qualify as measures. The result is that numerical techniques become detached from the measurement foundations required to justify them.

This omission has profound implications. Representational measurement provides the framework through which concepts such as scale theory, unidimensionality, dimensional homogeneity, admissible transformations and ratio measurement become scientifically meaningful. Without an understanding of these principles, students have little basis for evaluating whether numerical operations are lawful. Arithmetic becomes a procedural exercise rather than a scientific one.

The consequences are particularly important for HTA. Utilities, QALYs, preference scores and simulation models all depend upon assumptions regarding measurement. Yet those assumptions can only be evaluated through the lens of representational measurement. If students are not introduced to representational measurement, they are unlikely to ask whether utility scores possess the properties required for multiplication, whether multidimensional instruments can support quantitative claims, or whether the outputs of simulation models represent measures or merely numerical constructions.

The absence of representational measurement therefore removes the intellectual framework necessary for critical evaluation. Numerical outputs acquire legitimacy because they emerge from accepted analytical procedures rather than because their measurement foundations have been demonstrated. The distinction between measurement and arithmetic becomes blurred. Numbers are treated as measures simply because they exist.

This finding provides one of the clearest explanations for the persistence of measurement inversion within HTA. Measurement inversion occurs when arithmetic precedes measurement. Representational measurement exists precisely to prevent this from happening. It establishes the conditions under which quantitative claims become possible and identifies the limits beyond which arithmetic loses scientific meaning. If representational measurement is absent from the curriculum, students are unlikely to recognize when those limits have been crossed.

The implications extend beyond any individual method or analytical framework. The absence of representational measurement removes the principal safeguard protecting quantitative inquiry from numerical construction. Once this safeguard disappears, there is no longer a systematic mechanism for distinguishing measures from numbers, lawful arithmetic from inadmissible arithmetic, or quantitative claims from numerical speculation.

The significance of this result should not be underestimated. A curriculum that neglects representational measurement neglects the scientific foundation of quantitative reasoning itself. Students may become skilled in the application of analytical techniques while remaining unfamiliar with the principles that determine whether those techniques generate meaningful quantitative evidence.

The consequence is an educational framework in which quantitative claims are routinely advanced without reference to the conditions necessary to support them. Measurement becomes assumed rather than demonstrated. Arithmetic becomes detached from empirical structure. In this sense, the near absence of representational measurement from the curriculum provides perhaps the strongest evidence of curriculum inversion in New Zealand HTA and helps explain why measurement inversion continues to characterize contemporary HTA research and policy.

Statement 6: Attributes must be demonstrated to be unidimensional before measurement is possible

The endorsement probability for this statement is only 0.10 across PHARMAC, the University of Auckland and the University of Otago. This finding is particularly significant because unidimensionality is one of the fundamental requirements of measurement. Before an attribute can be measured, it must first be demonstrated that it represents a single coherent dimension. Without unidimensionality, measurement is impossible because there is no single attribute whose magnitude is being represented.

The proposition itself is straightforward. Measurement concerns the assignment of numbers to differing magnitudes of an attribute. This presupposes the existence of one attribute. If multiple attributes are combined within a single score, the resulting quantity may serve as a descriptive index, but it cannot automatically be assumed to constitute a measure. Measurement requires variation along a single dimension.

The weak endorsement of this statement suggests that unidimensionality occupies only a marginal position within the New Zealand HTA curriculum knowledge base. Students may be introduced to multidimensional outcome instruments, utility measures, quality-of-life scores and composite indices without being systematically taught that measurement requires evidence that a single attribute is being represented. The implication is that numerical aggregation may be mistaken for measurement.

This omission is particularly important because many of the instruments employed in HTA combine multiple dimensions of health and functioning. Mobility, self-care, pain, emotional status, social participation and other domains are frequently incorporated into composite scores. Yet combining several dimensions into a single numerical value does not demonstrate the existence of a single measurable attribute. Aggregation and measurement are not equivalent scientific activities.

The requirement for unidimensionality exists precisely to address this problem. Before numerical values can be interpreted as representing differing magnitudes of an attribute, empirical evidence must demonstrate that the observations are manifestations of one underlying dimension. Without such evidence there is no basis for claiming that the resulting scores represent quantities rather than collections of heterogeneous observations.

The consequences of ignoring this principle are substantial. If unidimensionality is not established, arithmetic operations performed on the resulting scores lose their scientific justification. Means, differences, ratios and model outputs may all be calculated, yet there is no assurance that these calculations relate to a single attribute. Numerical precision may increase while measurement validity remains unknown.

The relevance to HTA is immediate. Many patient-reported outcomes, health-status instruments and utility measures are built upon multiple dimensions of experience. Without demonstrating unidimensionality, it is impossible to know whether the resulting scores represent a measurable attribute. Yet these scores frequently become inputs to economic evaluations, utility calculations and simulation models. The scientific legitimacy of those downstream analyses therefore depends critically upon a requirement that appears largely absent from curriculum coverage.

The weak endorsement of this statement suggests that students are unlikely to be systematically introduced to this issue. They may learn how to apply outcome instruments and interpret composite scores without being taught to ask whether the underlying attribute has been demonstrated to be unidimensional. As a result, the distinction between a multidimensional index and a measurable attribute becomes blurred.

This finding provides another important explanation for the persistence of measurement inversion. Measurement inversion occurs when arithmetic is applied before the requirements of measurement have been satisfied. Unidimensionality is one of those requirements. If students are not taught to regard unidimensionality as a prerequisite for measurement, they are unlikely to recognize when numerical analyses are being conducted on quantities that do not qualify as measures.

The implications extend beyond individual instruments. The neglect of unidimensionality removes a critical scientific safeguard. Once the requirement disappears, multidimensional collections of observations can be transformed into numerical scores and treated as though they represent measurable quantities. The distinction between measurement and aggregation effectively vanishes.

The significance of this result should not be underestimated. Unidimensionality is not a technical refinement or an optional methodological preference. It is a necessary condition for measurement itself. A curriculum that gives only marginal attention to this principle leaves students without one of the essential tools required to evaluate quantitative claims.

The consequence is an educational environment in which composite scores are routinely treated as measures without prior demonstration that a single attribute exists. Numerical aggregation becomes a substitute for measurement. In this sense, the weak endorsement of unidimensionality provides further evidence of curriculum inversion within New Zealand HTA and helps explain why measurement inversion continues to characterize contemporary research and policy frameworks.

Statement 7: A Manifest Attribute is Directly Observable and Capable of Supporting Empirical Observation

The endorsement probability for this statement ranges from 0.35 to 0.40 across PHARMAC, the University of Auckland and the University of Otago. While this represents a higher level of endorsement than many of the preceding statements, it remains surprisingly weak given the fundamental importance of manifest attributes in scientific inquiry. The result suggests a partial recognition of the concept but little evidence that its significance is systematically integrated into the curriculum framework.

A manifest attribute is an attribute that can be observed directly. It exists independently of any measurement model and can be identified through empirical observation. Examples include survival time, hospital admissions, emergency department visits, number of prescriptions dispensed, treatment persistence, days absent from work and other observable events. Manifest attributes are important because they provide the most direct route to quantitative claims. Their existence does not depend upon inference; they are observed rather than estimated.

The distinction is scientifically important because manifest attributes and latent attributes represent fundamentally different measurement problems. Manifest attributes are measured through direct observation and counting. Their measurement framework is based upon empirical observation, scale properties and, where appropriate, linear ratio measurement. The scientific challenge is to demonstrate that the observed attribute possesses the properties necessary to support quantitative claims.

The endorsement profile suggests that this distinction is not strongly embedded within the New Zealand HTA curriculum. Students may encounter examples of observable outcomes, but there is little evidence that they are systematically introduced to the concept of manifest attributes as a distinct category of measurement. As a result, they may fail to appreciate why directly observable attributes occupy a unique position within scientific inquiry.

This omission has important consequences. Manifest attributes provide some of the strongest opportunities for evaluable and replicable claims regarding therapy impact. A reduction in hospital admissions, an increase in treatment persistence or a decrease in emergency department visits can be directly observed and independently verified. Such claims are capable of empirical testing and therefore fit naturally within the framework of normal science. Yet the curriculum evidence suggests that students are unlikely to be taught to view these outcomes through the lens of measurement theory.

The failure to emphasize manifest attributes also weakens understanding of the distinction between observation and inference. Scientific inquiry proceeds differently depending upon whether an attribute can be observed directly or must be inferred indirectly. Without this distinction, students may come to regard all outcomes as equivalent measurement problems. The unique advantages of directly observable attributes are therefore obscured.

The implications become particularly important when HTA moves beyond manifest outcomes into areas such as quality of life, symptom burden or patient satisfaction. These latter constructs are not directly observable and require a different measurement framework. Unless students understand what constitutes a manifest attribute, they are unlikely to appreciate why latent attributes pose different scientific challenges.

The weak endorsement of this statement therefore points to a broader deficiency in curriculum coverage. Students are not being systematically introduced to one of the most important distinctions in measurement science. Observable and non-observable attributes are treated as though they belong to a common analytical framework when, in reality, they require fundamentally different approaches to measurement.

This omission contributes directly to measurement inversion. Once the distinction between manifest and latent attributes is lost, there is little reason to question whether a particular measurement framework is appropriate for a particular type of attribute. Numerical scores become interchangeable, and the scientific requirements associated with different forms of measurement disappear from view.

The significance of this result should not be underestimated. Recognition of manifest attributes is not simply a matter of terminology. It is the first step in understanding that different classes of attributes require different forms of measurement. Without this understanding, students are left without the conceptual framework necessary to evaluate therapy impact claims appropriately.

The consequence is an educational environment in which directly observable outcomes are not distinguished clearly from inferred constructs. Measurement becomes generalized rather than attribute-specific. In this sense, the weak endorsement of manifest attributes provides further evidence of curriculum inversion and helps explain why the measurement foundations of HTA remain poorly understood within both research and policy environments.

Statement 8: A Latent Attribute is Not Directly Observable and Requires a Measurement Model to Estimate Possession of the Attribute

The endorsement probability for this statement is only 0.10 across PHARMAC, the University of Auckland and the University of Otago. This is one of the most revealing findings in the curriculum interrogation because it concerns the scientific foundation for measuring many of the outcomes that dominate contemporary HTA. The result suggests that the concept of latent attributes occupies only a marginal position within the New Zealand HTA curriculum knowledge base.

A latent attribute is an attribute that cannot be observed directly. Unlike survival time, hospital admissions or prescription counts, latent attributes are not immediately accessible through observation. Examples include pain, fatigue, depression, anxiety, functional limitation, quality of life, need fulfilment and many other patient-reported outcomes. These attributes are inferred rather than observed. Their existence is not in doubt, but their measurement presents a fundamentally different scientific challenge.

The distinction is crucial because latent attributes cannot be measured through simple observation or counting. They require a measurement model capable of estimating the extent to which an individual possesses the attribute. The scientific task is therefore not merely to collect responses but to construct a measurement framework that links observed responses to differing levels of attribute possession.

The low endorsement of this statement suggests that students are unlikely to be systematically introduced to this distinction. They may encounter quality-of-life instruments, patient-reported outcome measures and utility-based assessments without being taught that these involve latent attributes requiring specialized measurement models. As a result, numerical scores may be accepted as measures without any understanding of the scientific requirements necessary to support such claims.

This omission has profound implications. Much of contemporary HTA rests upon outcomes that are inherently latent. Quality of life, health status, symptom burden and patient functioning are routinely incorporated into economic evaluations and reimbursement decisions. Yet if students are not taught that these outcomes concern latent attributes, they are unlikely to ask how possession of the attribute is established or whether the measurement framework employed is scientifically adequate.

The consequences extend directly to the interpretation of scores. A numerical value obtained from a questionnaire does not automatically constitute a measure of a latent attribute. Responses must first be shown to support a measurement model capable of constructing a scale of attribute possession. Without such a model, numerical scores remain descriptive summaries rather than measures. The distinction is fundamental because arithmetic operations are only meaningful if measurement has first been demonstrated.

The absence of latent attribute measurement from the curriculum therefore leaves students without the conceptual tools necessary to evaluate many of the most important constructs used in HTA. They may become proficient in the application of instruments and scoring algorithms while remaining unfamiliar with the scientific principles required to establish whether those scores represent measurable quantities.

The implications are particularly important because latent attributes introduce a second form of ratio measurement that differs fundamentally from the ratio measurement associated with manifest attributes. Manifest attributes support direct linear ratio measurement through observation and counting. Latent attributes require a measurement model capable of constructing a Rasch logit ratio measure of attribute possession. These are different scientific problems governed by different mathematical frameworks. Yet the endorsement profile suggests that students are unlikely to be exposed to either the distinction or its implications.

This finding provides another important explanation for the persistence of measurement inversion within HTA. If latent attributes are treated as though they were directly observable quantities, the need for measurement models disappears from view. Numerical scores are accepted at face value and incorporated into increasingly elaborate forms of arithmetic. The prior scientific question—whether the attribute has been measured—simply vanishes.

The significance of this result should not be underestimated. Recognition of latent attributes is not a technical detail. It is essential to understanding the measurement problems associated with many of the outcomes that dominate contemporary HTA. A curriculum that gives only marginal attention to latent attributes leaves students ill-equipped to evaluate the scientific legitimacy of the instruments and scores they are taught to use.

The consequence is an educational environment in which latent attributes are routinely discussed but rarely understood as measurement problems. Numerical scores are accepted as measures without demonstration that possession of the underlying attribute has been established. In this sense, the weak endorsement of latent attribute measurement provides powerful evidence of curriculum inversion and helps explain why measurement inversion continues to characterize HTA research, policy and practice.

The endorsement probability of 0.75 across PHARMAC, the University of Auckland and the University of Otago indicates that the New Zealand HTA curriculum knowledge base recognizes the importance of identifying an outcome of interest in therapy assessment. This is a necessary starting point for any scientific inquiry. Before an intervention can be evaluated, there must be clarity regarding the attribute that is expected to change. Whether the concern is survival, symptom burden, functional status, treatment adherence or quality of life, the object of inquiry must first be identified.

At one level, therefore, the result appears reassuring. The curriculum demonstrates an awareness that therapy assessment requires a clearly defined target. This reflects standard research practice where the first task is to specify what is to be investigated. Without an attribute of interest there can be no meaningful evaluation.

The difficulty is that identification of an attribute is only the beginning of the scientific process. An attribute is not a measure. Naming an outcome does not establish that the outcome can be measured, nor does it establish the scale properties necessary to support quantitative claims. The identification of an attribute merely defines what is to be explained. It does not demonstrate that explanation can proceed through measurement.

This distinction is critical because the subsequent endorsement profile indicates that the curriculum largely stops at attribute identification. Once the target outcome has been specified, the scientific questions that must follow receive little attention. The principal scales of measurement receive only weak endorsement. The requirement that the measurement status of an attribute be established before quantitative claims are advanced receives very weak endorsement. Representational measurement, unidimensionality, latent attribute measurement and ratio measurement all receive minimal support.

Taken together, these results suggest that students are introduced to the concept of an attribute but are not systematically introduced to the scientific requirements that determine whether the attribute can support measurement. The curriculum therefore creates a potentially misleading impression that identifying an outcome largely completes the scientific task. Once an attribute has been named, attention shifts directly to analysis, modelling and evaluation. The much more demanding question of whether the attribute can be measured is rarely addressed.

The omission has profound consequences. Every attribute immediately raises a series of scientific questions. Is the attribute manifest or latent? Can it be observed directly or must it be inferred through a measurement model? Is it unidimensional? What empirical evidence supports its existence as a coherent attribute? What scale properties can be demonstrated? Does it support ordinal, interval or ratio measurement? Are arithmetic operations admissible? These questions determine whether quantitative claims are possible. Yet the endorsement profile suggests that they occupy only a marginal position within the New Zealand HTA curriculum.

The result is that attributes are routinely accepted as candidates for quantitative analysis without first establishing whether they satisfy the requirements of measurement. Numerical representation becomes a matter of convention rather than demonstration. Quantification follows automatically from identification. The scientific burden of proving measurement quietly disappears.

This may represent one of the principal origins of measurement inversion within HTA. Once identification of an attribute is mistaken for measurement of an attribute, the door is opened to increasingly elaborate forms of arithmetic, economic evaluation and simulation modelling based upon quantities whose measurement status has never been established. Numerical outputs acquire scientific authority because they are numerical, not because they are measures.

The significance of this finding should not be underestimated. The weakness of the curriculum is not that students fail to recognize the importance of attributes. The weakness is that they are not systematically taught the scientific principles that determine whether attributes can be measured. Identification becomes a substitute for measurement. The result is an educational framework in which measurement is assumed rather than demonstrated. This assumption underpins much of contemporary HTA and provides an important explanation for the persistence of measurement inversion in both research and policy.

Statement 9: Manifest and Latent Attributes Require Different Forms of Ratio Measurement

The endorsement probability for this statement is only 0.05 across PHARMAC, the University of Auckland and the University of Otago. This is among the lowest endorsement levels observed in the entire curriculum interrogation and represents one of the most serious deficiencies in the New Zealand HTA curriculum knowledge base. The statement concerns a distinction that lies at the heart of modern measurement theory. Its near absence suggests that students are unlikely to be exposed to one of the most important principles required for valid quantitative claims.

The proposition follows directly from the distinction between manifest and latent attributes. Manifest attributes are directly observable. Examples include survival time, hospital admissions, emergency department visits, treatment persistence and medication adherence. These attributes can be observed, counted and, where appropriate, represented through linear ratio measures. The measurement challenge is to demonstrate that the attribute possesses the properties required for ratio measurement, including a fixed non-arbitrary zero and meaningful proportional relationships.

Latent attributes present a fundamentally different problem. Pain, fatigue, depression, anxiety, physical functioning, need fulfilment and quality of life cannot be observed directly. They must be inferred from patterns of responses to carefully constructed instruments. Consequently, latent attributes require a measurement model capable of estimating the extent to which an individual possesses the attribute. The resulting measure is not a linear ratio measure derived from direct observation but a Rasch logit ratio measure representing attribute possession.

The distinction is not merely technical. It concerns two fundamentally different scientific frameworks. Manifest attributes are measured through direct observation and counting. Latent attributes are measured through a probabilistic model linking persons and items. The mathematical structures involved are different. The interpretation of the resulting measures is different. The conditions required to support quantitative claims are different.

The endorsement probability of 0.05 suggests that this distinction is almost entirely absent from the New Zealand HTA curriculum. Students are unlikely to be systematically introduced to the idea that observable and non-observable attributes require different measurement frameworks. As a result, all outcomes may come to be viewed as though they belong to a common analytical category. The scientific differences between manifest and latent measurement disappear from view.

This omission has profound implications for HTA. Many therapy impact claims concern manifest attributes such as survival, hospitalization and treatment utilization. Equally, many claims concern latent attributes such as quality of life, symptom burden and patient functioning. If students are not taught that these attributes require different forms of measurement, they have little basis for evaluating whether the methods employed are appropriate to the attribute being assessed.

The consequences extend directly to the interpretation of utilities and QALYs. Utility scores are frequently treated as though they represent proportions suitable for multiplication by time. Yet if the underlying construct is regarded as a latent attribute, the appropriate measurement framework is one concerned with attribute possession rather than proportional adjustment. A Rasch logit ratio measure of possession cannot automatically function as a proportion of survival time. The mathematical frameworks are fundamentally different. Failure to recognize this distinction undermines the legitimacy of many of the arithmetic operations central to HTA.

The weak endorsement of this statement therefore points to a critical gap in scientific understanding. Students may learn to apply instruments, utility scores and economic models without being taught that the measurement requirements differ according to the nature of the attribute under investigation. Numerical procedures become detached from the attributes they purport to represent.

This finding provides one of the clearest explanations for the persistence of measurement inversion. Once the distinction between manifest and latent measurement frameworks disappears, numerical quantities become interchangeable. Directly observed counts, utility scores, questionnaire totals and model outputs are all treated as though they occupy the same measurement space. The scientific requirements governing each type of attribute are forgotten.

The significance of this result should not be underestimated. The issue is not the omission of a specialized methodological concept. It is the absence of a principle that determines how measurement should proceed depending upon the nature of the attribute itself. Without this distinction, students lack the conceptual framework necessary to evaluate whether quantitative claims are scientifically defensible.

The consequence is an educational environment in which fundamentally different measurement problems are treated as though they were identical. Manifest and latent attributes become analytically interchangeable, even though they require different forms of ratio measurement and different mathematical foundations. In this sense, the near absence of this concept from the curriculum provides some of the strongest evidence of curriculum inversion within New Zealand HTA and helps explain why measurement inversion continues to characterize research, policy and reimbursement practice.

Statement 10: Therapy Impact Claims Must Be Falsifiable

The endorsement probability for this statement is 0.80 for PHARMAC and 0.65 for both the University of Auckland and the University of Otago. This is among the strongest endorsements observed in the curriculum interrogation and, at first sight, appears reassuring. The principle of falsifiability occupies a central position within modern science. A claim is scientific only if it is capable of being tested against observation and shown to be false. The relatively strong endorsement of this statement therefore suggests that the New Zealand HTA curriculum recognizes the importance of empirical accountability.

The concept itself is straightforward. Scientific claims are not validated by authority, consensus or methodological sophistication. They are validated through confrontation with evidence. A therapy impact claim must therefore be framed in a manner that permits empirical evaluation. If the claim cannot be tested, replicated or potentially refuted, it falls outside the boundaries of normal science.

Viewed in isolation, the endorsement of falsifiability appears to be a strength of the curriculum. Students are exposed to the idea that claims regarding therapy impact should not simply be accepted but should be subject to empirical scrutiny. This reflects an important scientific principle and distinguishes scientific inquiry from speculation or belief.

The difficulty emerges when this result is considered alongside the remainder of the endorsement profile. Falsifiability presupposes measurement. Before a claim can be tested, the attribute concerned must first be measured. Before measurement can occur, the attribute must be shown to possess the properties necessary to support measurement. The preceding statements indicate that these foundational requirements receive very weak endorsement within the curriculum. Scales of measurement, representational measurement, unidimensionality, latent attribute measurement and distinct ratio measurement frameworks all occupy only a marginal position.

This creates a profound inconsistency. The curriculum endorses the requirement that claims be falsifiable while simultaneously giving little attention to the conditions necessary to make falsification possible. In effect, students are taught the destination without being taught the route. They are encouraged to value empirical testing but are not systematically introduced to the measurement principles that make empirical testing scientifically meaningful.

The consequence is that falsifiability risks becoming a procedural concept rather than a scientific one. Claims may be described as testable simply because they generate numerical outputs. Yet numerical outputs are not necessarily measurable quantities. A simulation model may generate predictions. A utility score may generate a numerical value. A cost-effectiveness analysis may generate a ratio. None of these outcomes, however, guarantees that the underlying quantities possess the measurement properties necessary to support empirical evaluation.

The implications for HTA are particularly important. Contemporary HTA often presents utilities, QALYs and simulation model outputs as though they constitute evaluable scientific claims. Yet the ability to falsify a claim depends upon the measurement status of the quantities involved. If the underlying quantities have not been demonstrated to be measures, the apparent falsifiability of the resulting claim becomes questionable. The issue is not whether a model generates a prediction. The issue is whether the prediction concerns a measurable attribute.

The endorsement profile therefore reveals an important paradox. New Zealand HTA education appears to recognize the importance of falsifiability while giving little attention to the measurement foundations upon which falsifiability depends. Measurement is treated as implicit while falsification is treated as explicit. Yet the two cannot be separated. Without measurement there can be no meaningful quantitative claim. Without a meaningful quantitative claim there can be no meaningful falsification.

This finding provides an important insight into the persistence of measurement inversion. Students are taught that scientific claims should be evaluable and subject to empirical challenge. At the same time, they are not systematically taught how measurement establishes the conditions necessary for such evaluation. The result is that falsifiability becomes detached from measurement. Numerical claims appear scientific because they are framed as testable, even when the measurement status of the underlying quantities remains uncertain.

The significance of this result should not be underestimated. The relatively strong endorsement of falsifiability represents one of the few strengths identified in the curriculum profile. Yet the value of this strength is substantially weakened by the absence of the measurement principles that must precede it. Scientific testing is only meaningful when the quantities being tested qualify as measures.

The consequence is an educational framework that recognizes the importance of empirical challenge while largely neglecting the measurement foundations that make such challenge possible. In this sense, the endorsement of falsifiability reveals not only a strength but also a contradiction at the heart of the curriculum. Students are taught that claims should be tested, but they are not systematically taught how to establish whether the claims are measurable. The result is a curriculum environment in which falsifiability is endorsed in principle while the scientific conditions necessary to support it remain largely absent.

OVERVIEW OF THE CURRICULUM INTERROGATIONS

The curriculum interrogation of PHARMAC, the University of Auckland and the University of Otago reveal a pattern that is both consistent and deeply concerning. At first sight, the results suggest that the New Zealand HTA curriculum knowledge base incorporates elements of scientific inquiry. Students are exposed to the importance of identifying target outcomes, specifying attributes of interest and recognizing that therapy impact claims should be subject to empirical evaluation. These are positive features and indicate an awareness of the general objectives of scientific investigation.

Closer examination, however, reveals a far more troubling picture. The interrogation demonstrates that while students are introduced to the language of assessment, they are not systematically introduced to the scientific foundations that make quantitative assessment possible. The curriculum recognizes outcomes but largely ignores measurement. The result is an educational framework in which arithmetic, modelling and evaluation are permitted to proceed without first establishing whether the attributes under consideration satisfy the requirements of measurement.

The strongest endorsements are associated with the identification of attributes and the specification of target outcomes. Students are therefore likely to understand that therapy assessment requires a clearly defined object of inquiry. Similarly, the relatively strong endorsement of falsifiability suggests that students are exposed to the idea that scientific claims should be open to empirical challenge. Taken in isolation, these findings appear reassuring.

The difficulty is that these concepts represent only the beginning and end points of scientific inquiry. Between outcome specification and claim evaluation lies the entire science of measurement. It is precisely this intermediate stage that appears largely absent from the New Zealand curriculum.

The endorsement profile for the principal scales of measurement is weak. There is little evidence that students are systematically introduced to the distinctions between nominal, ordinal, interval and ratio scales or to the implications these distinctions have for quantitative analysis. Equally concerning is the weak endorsement of the proposition that the measurement status of an attribute must be established before quantitative claims can be advanced. This principle is fundamental to every quantitative science. Without measurement, there can be no meaningful quantitative claim. Yet the curriculum appears to give this requirement only marginal attention.

The most striking finding concerns representational measurement itself. The proposition that the axioms of representational measurement underpin quantitative claims receives virtually no endorsement. This is perhaps the most important result in the entire interrogation because representational measurement provides the scientific framework that distinguishes measures from numbers. It establishes the conditions under which arithmetic is lawful, identifies the permissible transformations associated with different scales and determines whether quantitative claims are scientifically defensible. Its near absence suggests that students are unlikely to encounter the conceptual foundations upon which quantitative reasoning depends.

A similarly weak profile is observed for unidimensionality. The requirement that an attribute be demonstrated to represent a single dimension before measurement is possible receives minimal support. Yet unidimensionality is a prerequisite for measurement. Without it, numerical aggregation may occur, but there is no basis for claiming that the resulting score represents a measurable quantity. The weak endorsement of this principle suggests that students may be exposed to composite scores and multidimensional indices without being taught the scientific conditions necessary for measurement.

The interrogation also reveals limited recognition of the distinction between manifest and latent attributes. While there is modest awareness that some attributes are directly observable, there is very little recognition that latent attributes require a measurement model to estimate possession of the attribute. More importantly, there is almost no recognition that manifest and latent attributes require fundamentally different measurement frameworks. This omission is critical because it removes one of the most important distinctions in modern measurement science. Students are unlikely to appreciate that observable outcomes and latent constructs pose entirely different measurement challenges and require different forms of ratio measurement.

Taken together, these findings reveal two interconnected deficiencies. The first is the absence of essential concepts. Scales of measurement, representational measurement, unidimensionality, latent attribute measurement and distinct measurement frameworks for manifest and latent attributes occupy only a marginal position within the curriculum knowledge base. Students are therefore unlikely to be exposed to the principles necessary to evaluate whether quantitative claims satisfy the requirements of measurement.

The second deficiency is even more serious. The curriculum simultaneously supports analytical frameworks that depend directly upon these absent concepts. Utilities, QALYs, preference scores and simulation-based evaluations all require assumptions concerning measurement status, scale properties and the legitimacy of arithmetic operations. Yet the scientific concepts necessary to evaluate those assumptions are largely absent from the educational environment. Students are therefore trained to apply methods whose measurement foundations they are not equipped to assess.

The broader implication is that curriculum inversion and measurement inversion are closely linked. Measurement inversion occurs when arithmetic precedes measurement. Curriculum inversion occurs when students are taught analytical methods without being taught the measurement principles necessary to evaluate those methods. The New Zealand interrogation suggests that the latter may help explain the persistence of the former. Students are introduced to established HTA frameworks but are not systematically introduced to the scientific standards that would allow those frameworks to be critically examined.

The conclusion is difficult to avoid. The New Zealand curriculum does not merely omit selected topics from measurement science. It omits many of the principles necessary to support lawful quantitative claims. The result is an educational framework in which measurement is assumed rather than demonstrated, arithmetic is accepted rather than justified and numerical outputs acquire authority without first satisfying the standards of representational measurement. This finding provides a compelling explanation for the persistence of measurement inversion within New Zealand HTA and points to the need for substantial curriculum reconstruction if HTA is to be aligned with the standards of normal science.

DOES HEALTH TECHNOLOGY ASSESSMENT HAVE A FUTURE IN NEW ZEALAND?

Given the conclusion presented here that the pervasive pattern of measurement inversion in HTA is attributable, at least in part, to curriculum inversion, the obvious question is whether HTA as presently constituted has a future. The answer depends entirely upon whether the discipline is prepared to confront the implications of measurement. Defending the *status quo* becomes increasingly difficult once the standards of representational measurement are introduced. The problem is not that current methods require refinement or methodological adjustment. The problem is that the foundational assumptions upon which the reference case rests fail to satisfy the requirements for quantitative claims. Judged by the requirements of measurement the reference case does not have a future and should never have been developed in the first place.

The most obvious example is the construction of utilities and the subsequent development of the reference case. Once it is recognized that multiplication and division require ratio measures, the burden of proof shifts immediately to those who claim that a quantity can function as a proportion. In the reference case, time is adjusted by a discount factor that is assumed to represent the value of a health state. Yet nowhere is a case made that this quantity must possess ratio measurement properties. The requirement itself is not even raised. More importantly, the failure does not begin at the point where the quantity is applied. It begins much earlier in the process of valuing health states. At every stage of development, from health-state descriptions, to valuation exercises, to utility algorithms, to the quantities entering simulation models, the requirements of ratio measurement are absent. Once this is recognized, the conclusion is unavoidable. Utilities, QALYs and reference-case simulation models cannot be regarded as scientifically credible measures of therapy impact. They have no place in a discipline committed to quantitative evidence. They should not be the focus of a curriculum unless they demonstrate error. Arithmetic cannot create measurement. Numerical complexity cannot compensate for measurement failure.

This does not mean that HTA itself has no future. On the contrary, the need for evidence-based assessment of therapeutic interventions has never been greater. Health systems must make decisions regarding pricing, access and resource allocation. Manufacturers must demonstrate product value. Clinicians require evidence of treatment impact. Patients deserve assessments that are transparent, evaluable and scientifically defensible. The question is therefore not whether HTA should survive, but whether it should survive in its present form. The answer is clearly no.

The future of HTA therefore lies in reconstruction rather than reform. The starting point is straightforward and entirely consistent with the standards accepted throughout the sciences. First, define the attribute of interest. Second, determine whether the attribute is manifest or latent. If the attribute is manifest and directly observable, then traditional linear ratio measurement techniques may be employed. If the attribute is latent and not directly observable, then a measurement model is required. The Rasch logit ratio framework provides the necessary basis for constructing measures of latent attribute possession. Once a valid measure is established, the same principles of empirical assessment, hypothesis testing and falsification can be applied.

This is not a difficult transition conceptually. The required measurement tools already exist. Representational measurement has been available for over 50 years. Rasch measurement has demonstrated how latent attributes can be measured through invariant logit scales for over 60 years. Scientific protocols for evaluable and replicable claims are well understood. What is required is not methodological invention but intellectual realignment. Measurement must once again precede arithmetic.

A further implication of the curriculum assessments is rarely discussed but cannot be avoided. Curricula do not emerge spontaneously. They reflect the knowledge, assumptions and priorities of faculty members responsible for teaching, supervision and research. If the concepts of representational measurement, ratio measurement, unidimensionality, latent attributes and Rasch measurement are absent from curricula across virtually every major HTA center, then an obvious question follows: are faculty themselves familiar with these concepts? Must training precede measurement before measurement can precede arithmetic

This possibility is more concerning than curriculum omission alone. A curriculum can be revised relatively quickly once deficiencies are recognized. Faculty capability is a more difficult issue. The Australian and international results suggest that several generations of researchers may have been trained without systematic exposure to the science of measurement. If so, the absence of measurement from teaching is not a conscious decision to exclude the topic. Rather, it reflects a broader intellectual tradition in which these concepts were never incorporated into professional training in the first place. An entire professional career can pass without any consideration of measurement theory.

This interpretation would also help explain the extraordinary consistency of the measurement inversion findings. The issue is not that researchers knowingly reject the axioms of representational measurement. The evidence suggests something more fundamental: the standards themselves are largely unknown. Questions concerning admissible transformations, unidimensionality, ratio properties, dimensional homogeneity and latent attribute possession are rarely raised because the conceptual framework required to raise them is absent.

The consequence is that reconstruction cannot be limited to students and future curricula. Faculty development must become a central component of transition. If HTA is to move toward a measurement-based framework, then current researchers, teachers, journal editors, reviewers and policy advisers will require the same education as the students they train. Without faculty development, curriculum reform is unlikely to be successful because the underlying knowledge gap will remain.

The challenge facing HTA is therefore larger than methodological reform. It is an educational reconstruction that must encompass both students and faculty. The science of measurement must become part of professional development, postgraduate training, research supervision and policy practice. Until this occurs, measurement inversion is likely to remain embedded within the institutions responsible for producing the next generation of HTA practitioners. The choice is between defending an indefensible orthodoxy and embracing a measurement-based framework capable of generating evaluable, replicable and falsifiable claims regarding therapy impact.

No other quantitative discipline has institutionalized such a rejection. No branch of physics, chemistry, engineering or measurement science begins by ignoring the scale properties of its variables and then proceeds as though arithmetic can rescue the situation. HTA stands alone. It constructed an entire evidentiary architecture upon assumptions that should have been dismissed at the outset. HTA remains not a science of value assessment but a prolonged exercise in numerical storytelling.

CONCLUSION: THE PATH TO RECONSTRUCTION

The requirements for reconstruction are straightforward and well established. The curriculum assessments reported here indicate that the central deficiency in New Zealand HTA is not a lack of analytical sophistication but the absence of measurement science; it is the acceptance of a paradigm failure for HTA. New Zealand can decide to cling to the wreckage or adopt a new paradigm that respects the standards for measurement and falsifiable evaluable therapy impact claims. Reconstruction must therefore begin with the concepts that are currently missing from professional education. The starting point is the definition of the attribute of interest, followed by the distinction between manifest and latent attributes. From this foundation students can be introduced to the only two measurement frameworks capable of supporting valid claims for therapy impact: linear ratio measurement for manifest attributes and Rasch logit ratio measurement for latent attributes.

The implications are uncompromising. The reference case, together with its supporting framework of health-state valuations, utilities, QALYs and simulation modelling, must be discarded. These constructs do not represent a scientific foundation for HTA. They are obstacles to reconstruction because they divert attention from the central requirement that measurement must precede arithmetic. A reconstructed HTA must be built upon measures rather than numerical assumptions, evaluable claims rather than simulation outputs, and empirical evidence rather than modelled futures. It must transition from non-science to science.

To support this transition, Maimon Research has developed a short nine-unit reconstruction program designed for both faculty and students. The program addresses the concepts absent from the curriculum assessments, including scales of measurement, representational measurement, unidimensionality, manifest and latent attributes, ratio measurement, Rasch measurement and protocol-based claims assessment. Particular emphasis is given to Rasch measurement as the framework for constructing measures of latent attribute possession and evaluating therapy impact through changes in possession of the attribute.

The program is deliberately non-technical in presentation and focuses on concepts rather than statistical complexity. Its purpose is to provide participants with the scientific foundations required to evaluate quantitative claims and to support the transition from reference-case modelling to measurement-based assessment. The final units address formulary evaluation, protocol design and the construction of evaluable and falsifiable claims for therapy impact. Further details regarding the reconstruction program are available from the Maimon Research website: https://maimonresearch.com/hta-reconstruction-program-and-fees/.

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 pf Measurement Vol 1: Additive and Polynomial Representations. New York: Academic Press, 1971

[3] The Theory of Allowable Transformations and the QALY. Logit Working Paper No 887. https://maimonresearch.com/logit-working-paper-no-887-june-2026/

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

---

[5] Maimon Research LLC. Interrogation Reports. www.maimonresearch.com

[6] New Zealand: The absence of measurement in academic research groups for health technology assessment. Logit Working Paper No 36. https://maimonresearch.com/logit-working-paper-no-36/

[7] New Zealand: PHARMAC endorses false measurement for health technology assessment. Logit Working Paper No 1261. https://maimonresearch.com/logit-working-paper-no-1261-april-2026/

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