MAIMON WORKING PAPER No 9 AUGUST 2025

THE RASCH NUMBER LINE: DISCOVERED OR INVENTED?

Paul C. Langley Ph.D. Adjunct Professor, College of Pharmacy University of Minnesota, Minneapolis MN

ABSTRACT

The longstanding debate regarding whether mathematics is discovered or invented gains renewed relevance when considering applied mathematical tools such as the Rasch logit number line. To argue mathematics is discovered is to suggest that mathematical structures exist independently of human cognition, as celestial laws might, while the invention position sees mathematics as a flexible, constructed system shaped to model reality. This tension becomes even more acute in contexts like the Rasch model: is the logit scale an inherent aspect of probabilistic response behavior, or merely a clever, human-imposed modeling artifact?

The Rasch model translates probabilities of correct responses into a log-odds—or logit—scale, mapping the relative difference between person ability and item difficulty onto a linear continuum. This continuum satisfies critical criteria: unidimensionality, interval-level measurement, and invariance. Such properties are not incidental; they are central to what constitutes "fundamental measurement" in latent trait theory. So, while one might craft the model, the measurement structure it uncovers may well reflect something more akin to a discovery, uncovering laws underlying response behavior rather than imposing artificial structure.

Concrete parallels are drawn to the Riemann Hypothesis: although unproven, its assumed veracity informs prime distribution models and exhibits consistent empirical coherence. Similarly, the Rasch number line behaves as though it exists independently of our conceptualization, suggesting it might be discovered as much as invented.

Latent constructs, such as intelligence, motivation, or disability, are undeniably inventions: human-defined abstractions inferred from behavior rather than directly observed. But their inventiveness does not negate the requirement that any measurement of them must meet rigorous criteria. Here, the Rasch model acts as the arbitrator: even if the construct is invented, a discovered structural alignment in the data (through unidimensionality, local independence, and invariance) validates its measurement. If the data fail to conform, the construct remains unmeasurable under scientific standards, suggesting that the Rasch structure is discovered in usage, contingent on empirical alignment.

Returning to measurement foundations, the axioms of ratio scales (identity, magnitude, equal intervals, and true zero) are not arbitrary: they are discovered necessities. In physical sciences, these axioms underpin meaningful measurement (length, mass, time). For latent constructs, the challenge is to create instruments yielding data that meet these same conditions—a task for which the Rasch model is uniquely equipped. When data adhere to Rasch's strict expectations, a latent

construct can yield interval, and even ratio-like, scales through the logit transformation, revealing rather than inventing quantitative structure.

The logit itself is not a modeling convenience but a uniquely suitable mathematical transformation, it linearizes bounded probabilities, converts multiplicative odds into additive logits, and preserves ratio properties. The Rasch model did not invent the logit; it discovered that only this transformation can satisfy fundamental measurement requirements within probabilistic latent trait models.

In conclusion, the Rasch logit number line embodies a rare confluence: invented constructs measured through discovered structures. Human minds invent latent traits, but the Rasch model reveals whether those constructs align with necessary mathematical laws, thereby transforming numerical assignments into scientific measurement.

INTRODUCTION

The question of whether mathematics as it is practiced was discovered or invented has long been a basis for dispute. To say that mathematics was discovered is to assert that mathematical structures and relationships exist independently of the human mind, waiting to be uncovered like physical laws or planetary orbits. By contrast, the view that mathematics is invented treats it as a symbolic system, constructed by humans to model and make sense of the world, subject to revision and redefinition. This debate becomes particularly provocative when we consider not abstract number theory, but applied mathematical structures such as the Rasch logit number line. Is this scale, central to the measurement of latent traits, something that exists as part of the structure of probabilistic response behavior, or is it a modeling artifact; ingenious, useful, but ultimately invented?

The Rasch model maps the probability of a correct response to a linear logit scale, defined by constant relative differences between person ability and item difficulty [1]. This mapping produces a number line that supports additive operations in the logit space and multiplicative transformations in the odds space, satisfying the conditions of a unidimensional, interval, and invariant scale. These are not arbitrary properties; they make the Rasch number line a necessary and sufficient condition for fundamental measurement in latent trait theory. The structure is not merely a statistical convenience; it allows us to make meaningful, quantitative statements about the possession of an unobservable attribute. In this sense, it seems less like a mathematical invention and more like a discovery.

This epistemological question finds a useful parallel in the unresolved but widely accepted Riemann Hypothesis [2]. Although unproven, the hypothesis, that all nontrivial zeros of the Riemann zeta function lie on the critical line, places strict constraints on the distribution of prime numbers and would dramatically refine the error bounds in the Prime Number Theorem. Mathematicians have long relied on its assumed truth because it yields correct predictions and elegant formulations. The underlying function behaves as though the hypothesis were true, reinforcing the view that mathematics often reveals deep structures that exist whether or not we have yet proved them.

In the same way, the Rasch logit number line behaves as if it were discovered. It possesses properties that reflect an underlying structure in response behavior. Its utility and coherence in transforming ordinal responses into interval measures suggest it uncovers a fundamental measurement structure; not one we created, but one we found.

ARE LATENT TRAITS INVENTED

To ask whether latent constructs are invented is to interrogate the ontological status of abstractions like intelligence, need fulfillment, motivation, or disability. These constructs are not observed directly but inferred from patterns of behavior, response, or performance. Unlike physical objects or naturally occurring forces, latent constructs have no existence independent of the interpretive framework humans bring to bear on experience. They are conceptually dependent. Without human cognition, discourse, and inquiry, there is no “intelligence” to measure, no “depression” to quantify, no “patient-reported outcome” to evaluate. The behaviors may occur, neuronal patterns might fire, organisms might display variation, but the labeling and structuring of these variations as specific, latent traits are the work of human minds. In this sense, latent constructs are clearly invented. They are tools of understanding, not features of the world waiting to be uncovered.

But the invention of a construct does not mean that anything goes. Once a latent construct is proposed, say, patient burden or functional disability, it is not free-floating. To be meaningful, it must be consistently defined, bounded by content validity, and subjected to rules of empirical engagement. What follows is the possibility of measurement. Here, the Rasch model intervenes. Although the construct is invented, the structure of its measurement is not. The Rasch logit number line emerges only if the data behave in accordance with strict requirements, unidimensionality, local independence, invariant item ordering. This transforms the invented construct into a measurable quantity on a discovered scale. The construct does not make the measurement structure true; the measurement structure makes claims about the construct scientifically evaluable.

This gives rise to a deeper paradox. If the Rasch logit number line is a discovered structure, one that must conform to specific axioms to exist in the data, then what is it discovered for, given that latent constructs do not exist apart from our invention? Does the discovery of the Rasch scale imply the prior existence of something real and structured to be uncovered? Or does it remain a tool, valid only in the context of a construct we’ve already invented?

The answer may lie in distinguishing between two levels of existence. The construct is a cognitive and linguistic invention; it allows us to talk about phenomena in a structured way. But the Rasch scale reveals whether the responses generated under the construct's rubric obey a lawful, mathematical pattern. The measurement structure is discovered not in the construct itself, but in the relationship between responses and the claim of quantitative structure. In that sense, the Rasch number line is not discovered in the abstract, but discovered in use. It is a revealed structure contingent on behavior aligning with modeled expectations. Without human invention of the construct, the model would have no object. Without conformity to Rasch expectations, the construct would have no measure. The interplay between invention and discovery is what gives latent trait measurement its scientific character.

FUNDAMENTAL MEASUREMENT AND THE NECESSITY OF RATIO SCALES

To ground this discussion in science, we must return to the foundational principles of measurement. In physical sciences, measurement is not a loose approximation but a rigorous enterprise constrained by formal axioms or rules. A measure is only scientifically credible if it conforms to the axioms of fundamental measurement; identity, magnitude, equal intervals, and a true zero. When all four axioms are met, we have a ratio scale. This is not a matter of convention or utility but of necessity. A ratio scale is the only scale that allows all arithmetic operations to be meaningful and interpretable. You cannot multiply or divide quantities unless you are operating on a ratio scale. These axioms are not invented; they are discovered. They express the necessary preconditions for valid quantitative representation. If data meet these axioms, we do not choose to treat them as measurements; they *are* measurements, by definition.

This principle applies first to manifest quantities; physical, observable, and directly measurable attributes. Length, mass, time, and temperature (on the Kelvin scale) meet the criteria. They are not statistical artifacts; they are representations of physical reality that obey the discovered axioms of interval and ratio measurement. In this space, there is no controversy. No one argues that inches or kilograms are invented conventions in the same sense as constructs like “burden” or “resilience.” They are discovered relationships among real properties of the world.

But when we move from manifest to latent quantities, the epistemological terrain shifts. Latent constructs are inventions, yet if we wish to make them measurable, we are still bound by the same axioms of measurement. The challenge is to determine whether these invented constructs can yield data that conform to a discovered structure satisfying the requirements for measurement. That is the role of the Rasch model. It acts as a measurement adjudicator. Where the data meet Rasch's strict expectations, they support a linear, interval logit scale, and through logistic transformation, a ratio scale of odds. The transformation does not create measurement; it reveals that the structure already exists within the data. It is the expression of a discovered quantitative order; not in the construct, but in the behavior elicited under its framing.

Thus, the Rasch logit number line should not be conflated with the invented construct. It is the manifestation of a deeper mathematical architecture; a necessary structure for any defensible claim to measurement. If the data fail to meet Rasch conditions, we are not measuring, regardless of intent or statistical sophistication. No amount of modeling or simulation can compensate for failure to meet these axioms. This is the boundary between numerical description and scientific measurement. Where manifest quantities present the clearest examples of discovered measurement, latent constructs must be judged by whether their observed indicators yield data that fit a discovered, invariant scale. Without this, we are merely assigning numbers; not measuring.

WHY THE LOGIT

If the Rasch model is grounded in a discovered mathematical structure, the question follows: why is the logit its central transformation? Why not any of the other mathematical functions available for modeling probability? The answer lies in the logic of measurement, not statistical convenience. The logit, defined as the natural logarithm of the odds of a correct response, transforms bounded probabilities (between 0 and 1) onto an unbounded, linear scale. This transformation is not just numerically elegant; it is mathematically necessary for satisfying the conditions of fundamental measurement in the context of latent traits.

The logit transformation is uniquely suited to capture *constant relative differences*. This is crucial because when measuring probabilistic behavior (such as a person's likelihood of succeeding on an item), we are not dealing with absolute differences, as in physical measurement, but with *changes in odds*. The odds of success, when expressed as a ratio of probabilities (p / [1 – p]), allow multiplicative comparisons. The logit transforms these odds into additive intervals, preserving a linear structure in which a one-unit increase always represents the same relative change in odds; this is the very definition of a ratio structure in the domain of probabilities. No other transformation offers this consistency.

Rasch did not invent the logit; it has a long history in statistics, notably in logistic regression. But Rasch appropriated it for a different purpose; not to maximize predictive power, but to construct a *measurement* framework. His goal was not empirical fit but epistemic clarity. The logit scale offers a consistent additive metric for comparing person ability and item difficulty. In Rasch's model, the probability of a correct response is determined solely by the difference between these two parameters. The logit transformation ensures that this difference has a consistent meaning across the entire range of the scale, regardless of where it occurs.

Thus, the logit is not just a convenient mathematical tool. It is the only known transformation that permits the mapping of probabilistic data onto a linear, additive, and ratio-like scale that meets the criteria of fundamental measurement. It was discovered, not constructed; and Rasch's genius was to recognize that only by anchoring his model in this transformation could latent constructs be measured with the rigor of the physical sciences. Rasch's contribution, then, was not to invent the logit, just as Einstein did not invent the field equations but uncovered the necessary mathematical form that the relationships between geometry and spacetime must take [3], but to *uncover* the necessary form that a probabilistic model must take if it is to transform ordinal responses into a linear, unidimensional, and invariant *interval scale* of latent trait possession. Even if humans had never evolved, the logit structure would still be there as the unique solution to the problem of *how to measure an unobservable trait from a finite set of discrete outcomes*. Rasch merely provided the epistemic *key* to open the door to a number line that was always there.

Just as the real number line or the complex plane "comes into existence" once their axioms are defined, the Rasch logit number line is a necessary mathematical consequence of the axioms Rasch had to respect. It is a scale on which both persons and items are located, such that the difference in their positions yields a discoverable, predictable probability of success. The logistic function, the odds transformation, and the additivity of logits were already latent in measurement theory; waiting not to be invented, but to be recognized. In that sense, the Rasch logit scale was always there. Rasch simply found it.

WHERE IGNORANCE IS BLISS

The Rasch model stands virtually alone in the social sciences as a method that satisfies the axioms of fundamental measurement for latent traits. Rasch's achievement was not merely technical, it was philosophical. He understood that measurement, if it is to resemble that of the physical sciences, must obey a strict logic of structure: unidimensionality, invariance, additivity, and separability. These are not optional ideals but necessary conditions for transforming ordinal responses into interval or ratio measures. Rasch was not building a model to fit data; he was specifying the only framework in which a latent trait could be *measured* rather than merely described.

In this, Rasch's work aligns with the foundational contributions of Luce and Tukey in their discovery of the conditions under which additive representations must hold if one wishes to assign numbers meaningfully to qualitative data; a structure that was always latent in the logic of preference and comparative judgment [4]. Luce's axiomatic framework for measurement and Tukey's emphasis on mathematical discipline in data exploration each underscored that any legitimate science of behavior or cognition must rest on structural rigor equivalent to that of the physical sciences. The logic of additive conjoint measurement, for example, leads inexorably to the logit transformation under probabilistic assumptions. If one follows the axioms, rather than adjusting them to fit the data, the logit scale emerges not as a modeling choice, but as a mathematical necessity.

And yet, social science has overwhelmingly ignored this path. Instruments abound that sum ordinal scores, present them as interval data, and make inferential claims as though measurement had occurred. Composite scores and multi-attribute scales proliferate, even when the data violate unidimensionality or fail to meet basic criteria for invariance. Why is this the norm? Partly, it is ignorance. The philosophical and mathematical demands of the Rasch model are not widely taught. Most researchers are trained in regression, factor analysis, and statistical prediction, not in the axioms of measurement. Rasch requires a shift in thinking, from fitting models to data toward testing whether data warrant measurement in the first place where data must fit the Rasch model.

But the problem runs deeper than mere ignorance. There is resistance. Rasch modeling is often seen as difficult, time-consuming, or unforgiving. It demands unidimensionality and discards misfitting items; unlike other models that accommodate complexity through elaboration. More importantly, Rasch produces *meaningful quantities*, and this introduces accountability. If a person is said to possess more of a trait, the model tells us *how much more* in a way that is not contingent on item content or population. That raises a more uncomfortable question: what do differences in possession actually mean? To ask that is to re-enter the world of epistemology, ontology, and construct validity. For many in the social sciences, it is easier not to ask.

So, ignorance, yes, but also institutional inertia, fear of refutation, and the comfort of ambiguity. Where Rasch, Luce and Tukey sought clarity, most prefer convenience. In such a world, ignorance is not only bliss, it is policy.

THE IMPOSSIBLE QALY

Few constructs have enjoyed such widespread institutional acceptance while being so fundamentally unsound as the Quality-Adjusted Life Year, or QALY. It is presented as a ratio measure; a combination of time lived and a health utility weight meant to reflect the quality of that time. The claim is seductive in its simplicity: multiply the years of life gained by a score between 0 (death) and 1 (perfect health), and the result is a single number capturing both length and quality of life. In practice, this construct underpins vast health technology assessments, cost-effectiveness analyses, and pricing negotiations. Yet the foundation on which it rests is conceptually, mathematically, and empirically bankrupt.

At the core of the QALY is a fatal conflation of scale types. Time is indeed a ratio scale, unambiguously measurable and possessing a true zero. But the so-called "utility weights" applied to time, typically derived from techniques like the time trade-off (TTO), are nothing of the kind. These weights are constructed from ordinal preferences captured through choice experiments, often using multiattribute instruments like the EQ-5D-3L. Respondents are asked how many years in less-than-perfect health they would trade for fewer years in perfect health. The resulting preferences are then normalized onto a scale where 1 is defined as perfect health and 0 as death, with some scores even permitted to fall below zero.

This transformation is not measurement. It is numerology. The TTO-derived values lack the properties of an interval or ratio scale. They do not satisfy conditions for equal intervals or constant relative differences. They are preference statements, not quantities. Worse, these preference scores are then treated as if they can serve as *dependent variables* in regression models, a practice that collapses under even minimal scrutiny. You cannot regress meaningful results on a non-measurable outcome variable. The underlying logic is circular and incoherent.

What makes the QALY particularly insidious is that its flaws are not obscure. They have been documented and criticized for decades, yet the construct persists because it serves a purpose: it gives the illusion of precision. It allows committees, payers, and policymakers to compare therapies using a single, seemingly objective metric. But the objectivity is illusory. The QALY is not a measure; it is a composite artifact, a pseudo-ratio constructed from ordinal fragments. It cannot be replicated, it cannot be validated against anything measurable, and it cannot support hypothesis testing. That health technology assessment continues to treat the QALY as if it were a discovered truth rather than a concocted artifact [5]. In any science genuinely committed to the principles of measurement, the QALY would have been discarded long ago. It has no existence beyond the confines of a blinkered human imagination.

CONCLUSION: DISCOVERY THROUGH AXIOMATIC NECESSITY

Measurement, if it is to rise to the level of science, must rest not on convenience or statistical habit, but on axioms; discovered principles that define what it means to quantify. These axioms, rooted in additivity, unidimensionality, invariance, and the existence of a true zero, are not adjustable. They are not invented. They are uncovered, as immutable in their structure as the number line or the laws of Euclidean geometry. The task of measurement theory is not to invent new ways of assigning numbers to observations, but to discover whether observed data can be shown to conform to these necessary, lawful structures.

The Rasch model is unique because it formalizes this challenge. Rasch did not simply construct a statistical technique to improve data fit or model predictive probabilities. He recognized a deeper truth: that to measure an unobservable trait, one must build a conjoint framework in which both the item and the person occupy a position on the same invariant number line. That line, the Rasch logit number line, is not invented. It is the only known scale that, under probabilistic assumptions, transforms ordinal response data into interval or ratio-like values while preserving the necessary conditions for measurement. The logit emerges not as a modeling choice but as a mathematical inevitability once we commit to the axioms.

Rasch's role, then, was not that of a creative designer but of a scientific discoverer. He invented the key, the measurement model, but what he opened was a structure already there: the necessary number line for latent traits, made visible through observed responses. Like the complex plane or the distribution of primes, the Rasch scale existed once its conditions were known. It lay latent within the logic of measurement theory, awaiting formal articulation.

That social science continues to ignore this structure in favor of models and instruments that fail these axioms is not a scientific judgment but a cultural one; a preference for tradition, ambiguity, and convenience over clarity, precision, and constraint. Where Rasch offered the possibility of genuine discovery in human measurement, others settled for simulation, scoring rules, and ordinal composites masquerading as interval data. But the standard remains. The Rasch number line is not a convention. It is a revealed structure: the only legitimate path to measuring what cannot be seen. It is what happens when measurement moves from assigning numbers to discovering quantities. And that is not invention; it is science.

ACKNOWLEDGMENT:

This paper benefited from the use of OpenAI's ChatGPT (GPT-4, 2024–2025), which served as a valuable tool in refining arguments, structuring content, and improving the clarity of presentation. While the author retains sole responsibility for all content, interpretations, and conclusions, the interactive support provided by this AI platform contributed meaningfully to the writing and editorial process

REFERENCES

---

[1] Bond T, Yan Z, Heene M. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. New York: Routledge, 2021

[2] Sabbagh K. The Riemann Hypothesis. New York: Farrar, Straus and Giroux, 2002

[3] Janssen M, Renn J. How Einstein found his Field Equations: Sources and Interpretations. Cham, Switzerland: Springer Nature, 2022

[4] Luce R, Tukey J. Simultaneous Conjoint Measurement: A new type of fundamental measurement. J Math Psychol, 1964;1(1):1-27

[5] Drummond M, Sculpher M, Claxton K et al. Methods for the Economic Evaluation of Health Care Programmes (4th Ed). New York: Oxford University Press, 2015