

The Thinking Quotient (TQ)

Further Validation of a Skill-Metric for Thinking

Derek Cabrera^{a,b,@}, Valath Bhuan Das, S. K.^b, Laura Cabrera^{a,b,@}

a: Cornell University

b: Cabrera Research Lab

@: dac66@cornell.edu

ABSTRACT: For decades, systems thinking has been treated as an essential but elusive skill—praised for its importance, yet lacking a clear definition, measurable structure, or validated means of development. The Thinking Quotient (TQ) advances this gap by providing the first empirically validated skill-metric for systems thinking, continuing a research program that began with the Systems Thinking and Metacognitive Inventory (STMI). Unlike traditional psychometric instruments that measure traits, TQ functions as an edumetric assessment: it quantifies learnable cognitive performance based on structural organization of thought, not personality or belief. Grounded in O-Theory, DSRP Theory, and the Cognitive Law of Thinking ($\mathbb{M} = \mathbb{I} \cdot \mathbb{O}$), TQ operationalizes the four universal cognitive patterns—Distinctions, Systems, Relationships, and Perspectives (DSRP)—as measurable elements of structural thinking. This study reports further validation of TQ (STMI v4.0), including reliability analyses, ordinal adjustments, and a bifactor confirmatory factor analysis (CFA) using polychoric estimation. The results demonstrate high internal consistency ($\alpha = .87$, $\omega_{\text{total}} = .88$ overall), excellent model fit (CFI = .986, TLI = .985, RMSEA = .016), and strong evidence for a dominant general systems-thinking factor ($\omega_{\text{hierarchical}} \approx .70$) supported by smaller, domain-specific DSRP subskills ($\omega_{\text{specific}} \approx .10-.25$). Together, these findings confirm that systems thinking can be scientifically measured as a trainable, structural skill—one that develops through explicit practice (the Six Moves), can be tracked quantitatively (TQ), and bridges the gap between theory and real-world performance improvement.

KEYWORDS: Thinking Quotient (TQ) | systems thinking | DSRP Theory | O-Theory | structural cognition | edumetric skill-metric | cognitive law of thinking | model–world alignment | organization of information | Love–Reality Loop

Cite this paper: Cabrera, D., Valath Bhuan Das, S. K., & Cabrera, L. (2025). The thinking quotient (TQ): Further validation of a skill-metric for systems thinking. *Journal of Systems Thinking*, 5(1), 1–18.
<https://doi.org/10.54120/jost.v25il.dc6tqsm6>

1.0 Introduction

Background

Systems thinking has long been recognized as a critical 21st-century skill, yet for most of its history it has been conceptually diffuse and practically immeasurable. Despite more than half a century of theorizing—from von Bertalanffy’s *General Systems Theory* to Senge’s *Fifth Discipline*—there has been no scientifically validated means of measuring *how* people think systemically, or whether that ability can be deliberately developed.

This gap left systems thinking as a kind of “folk construct”—important in principle but undefined in structure. Practitioners could teach *about* systems, but not reliably measure *thinking as a system*. Over the past two decades, our research has sought to close that gap by identifying, testing, and validating the underlying cognitive mechanics of systems thinking itself.

Through a series of empirical discoveries summarized in Table 1, we have established a unified theoretical and empirical foundation—collectively known as O-Theory—that defines how both mind and reality organize information through a shared structural grammar. These findings culminate in the development of TQ: a precise, measurable, and trainable skill-metric for systems thinking.

Theoretical Foundation

Our work begins with the discovery of a universal grammar of thought and reality, formalized as DSRP Theory—Distinctions, Systems, Relationships, and Perspectives (Cabrera, 2006–2025). DSRP identifies four structural patterns that govern both cognition ($\mathbb{M}$) and the natural world ($\mathbb{R}$). This theory is grounded in the Cognitive Law of Thinking:

$$\mathbb{M} = I \cdot O$$

where I represents information and O represents the organization applied to it.

This law reveals that understanding depends not on the quantity of information but on the structure applied to it. The implication is profound: learning, culture, and cognition are not merely accumulations of data but transformations of information through universal organizational patterns.

Subsequent research formalized *O-Theory*, demonstrating that these same structural patterns operate simultaneously and interdependently in both mind and matter. Empirical validation across multiple domains—visual, textual, and tabular—confirmed the *model–world alignment* principle ($\mathbb{M}/\mathbb{R} \approx 1$), showing that accurate understanding arises when mental models share the same organization as reality.

Further studies established the *Cognitive Pareto Law*, showing that approximately 20% of recurrent cognitive structures account for 80% of organized thought. This finding led to the identification of the *Six Pareto Moves*—a minimal set of DSRP-based micro-operations that surface and correct the majority of cognitive biases. The *20/80 Thinking Advantage Principle (TAP)* formalized this efficiency: mastering a

small number of structural operations yields disproportionate gains in intelligence, clarity, and adaptability.

This program of research also introduced the *Love–Reality Loop (LRL)*, a structural feedback model describing how mental models ($\mathbb{M}$) are continually tested against reality ($\mathbb{R}$), with misalignment producing error and alignment producing learning and adaptation.

Together, these discoveries transformed systems thinking from a philosophical idea into a measurable science of cognition and organization.

Table 1. Top 14 Research Findings Supporting O-Theory

#	Finding	Summary Technical Insight	Plain English Summary & Implications	Citations
1	Empirical validation of the cognitive law $\mathbb{M} = \mathbb{I} \cdot \mathbb{O}$ and $\mathbb{M}/\mathbb{R}$ (Love Reality Loop)	Controlled experiments across tabular, visual, and textual domains demonstrated that model performance depends multiplicatively on both information ($\mathbb{I}$) and organization ($\mathbb{O}$). The product-only model ($\mathbb{M} = \mathbb{I} \cdot \mathbb{O}$) matched the fit of more complex models, confirming the law's existence and effect.	How information is organized—not how much there is—determines understanding. Establishes $\mathbb{M} = \mathbb{I} \cdot \mathbb{O}$ as a measurable law of cognition and defines learning ($\Delta\mathbb{M}$), organizational learning ($\Delta\mathbb{M} \times a$), and culture (shared $\mathbb{M}$).	[1]
2	Theoretical establishment of DSRP-483 Theory	Defines four universal cognitive and ontological patterns—Distinctions, Systems, Relationships, Perspectives—that organize all information in both mind and nature.	Thinking and reality follow the same simple rules; a unified model of structure.	[2–8]
3	DSRP patterns exist universally in nature (ontologically)	Evidence shows D, S, R, P—and their elements ($i:o$, $p:w$, $a:r$, $v:p$)—structure physical, biological, and social systems.	Nature organizes information using the same structural logic as cognition.	[9] includes a growing database of over 200 studies
4	DSRP patterns exist universally in mind (cognitively)	Cognitive and neuroscience research confirm that these patterns organize mental models and perception.	The human mind runs on the same code that organizes the universe.	[10–14]
5	Dynamics of DSRP (equality, coimplication, simultaneity)	Each pattern's elements act together and define one another in real time; none operate alone.	Thinking is dynamic and co-creative; patterns emerge simultaneously.	[10–14]
6	Interdependence of DSRP patterns (“enclosure”)	Empirical and formal results show D, S, R, P are recursively nested; activating one activates all.	You can't make a distinction without also creating a system, relationship, and perspective.	[10–14]
7	DSRP patterns measurably improve performance	Controlled experiments show large effect sizes in systems thinking, metacognition, and emotional intelligence.	When people learn how thinking actually works, they become smarter and more collaborative.	[14] numerous case studies
8	Most ST frameworks lack scientific validity (Echo Chamber Index)	ECI analysis shows 7 of 8 major frameworks cite themselves and lack empirical grounding.	Much of the field is based on belief, not evidence. DSRP meets validity standards.	[15]
9	Critical Thinking is a form of bias	Corpus and discourse analyses show CT definitions are mostly negative and evaluative.	“Critical thinking” often teaches judgment instead of understanding. DSRP replaces critique with clarity.	[16]

10	The 80/20 Cognitive Pareto Law (CPL)	Data show ~20 % of cognitive structures generate ~80 % of organized thought.	A few mental moves do most of the work—intelligence can be trained efficiently.	[17]
11	Six Moves embody the 20/80 Thinking Advantage	Analysis of Pareto patterns led to six universally efficient cognitive Moves and the Thinking Advantage Principle (TAP).	Six simple mental moves capture the 20 % of thinking that drives 80 % of understanding.	[17,18]
12	System 3 Thinking (“fast and right”)	Training in DSRP patterns automates accurate, high-speed reasoning that combines intuition with rigor.	Deep thinking can become second nature—people can think fast and right.	[17]
13	Six Moves improve performance by ≈550 %	Studies found average 550 % gains in systems-thinking skill when learners practiced the Six Moves on complex problems. (N = 988, ages 18–75; 1,976 observations; $\eta^2 \approx .45$.)	The Six Moves make people up to five times better at seeing connections and solving problems.	[19] + numerous case studies
14	Empirical validation of TQ (Thinking Quotient)	The TQ is the first validated skill-metric for systems thinking ($\alpha \approx .7$) confirming DSRP skills are measurable and interdependent.	TQ proves that systems thinking can be measured and improved universally.	[20] + new paper

Thinking vs. Types of Thinking

Thinking isn’t what we think it is. In modern culture, “thinking” has been splintered into dozens of branded subtypes—each treated as though it were a separate mental process:

Common Sense, Critical Thinking, Emotional Intelligence, Creative Thinking, Growth Mindset, Design Thinking, Leadership Thinking, Problem-Solving Thinking, Strategic Thinking, Systems Thinking, Analytical Thinking, General Intelligence, Interpersonal Intelligence, Intrapersonal Intelligence, Grit, Resilience, Anti-Fragility, Scientific Thinking, Fast/Slow Thinking, Convergent/Divergent Thinking, Lateral Thinking, Reflective Thinking, Abstract Thinking, Inductive Thinking, Deductive Thinking, Situational Thinking, Clear Thinking, Complexity Thinking, Ecological Thinking, Adaptive Thinking, Enterprise Thinking, Thinking Like a Scientist, Synthetic Thinking, Interdisciplinary Thinking, Prosocial Thinking, and Bias.

All of these are *thinking*. They differ in content, purpose, and cultural emphasis—but not in *structure*. Every one of them, when examined closely, depends on the same universal cognitive patterns identified by DSRP Theory: making **Distinctions**, organizing **Systems**, recognizing **Relationships**, and taking **Perspectives**. These patterns are not optional “styles” of thinking—they are the universal *grammar of cognition* itself. Every act of thought, regardless of its context, involves these four interdependent operations.

Within this broader landscape, *Systems Thinking* occupies a special place. As it is popularly defined today, it appears as just one “type” among many—often focused on complexity, feedback, and interdependence. But when understood more precisely—as *Systems (ontological realism) + Thinking (epistemological*

constructivism)—it represents the *ultimate goal* of thinking itself: to align our internal mental models with the external reality they seek to represent. In other words, Systems Thinking is not only one type of thinking—it is *the success condition of all thinking*. To think systemically is to think in a way that mirrors the structure of reality.

This perspective also explains why so much of what passes for “critical” or “systems” thinking in popular and academic discourse becomes circular or ideological. Detached from the structural mechanics of cognition, these frameworks drift into bias and echo chamber effects—focusing on opinion rather than organization. DSRP and O-Theory correct this by showing that thinking *is* systemic by nature, because cognition and reality share the same structural grammar.

The Thinking Quotient (TQ) measures that grammar in action. It quantifies how well individuals *organize information in structurally systemic ways*—how closely their thinking mirrors the systemic organization of reality itself. In this sense, TQ doesn’t measure a “type” of thinking; it measures the *quality* of thought’s alignment with reality. It is both inclusive of Systems Thinking and larger than it: a measure of the general (*g*) factor of systemic thought, the structural skill that underlies and unites all other forms of thinking.

From STMI to TQ (v4.0)

Building on these foundations, we developed the Systems Thinking and Metacognitive Inventory (STMI) as the first empirical instrument designed to measure growth in systems thinking. Across four major versions ($n = 1,580$), the STMI achieved progressively higher reliability ($\alpha \approx .69$ overall in early versions). The STMI’s purpose was twofold: to measure *Competence* (actual systems-thinking skill) and *Confidence* (metacognitive awareness of that skill).

The *Thinking Quotient (TQ)* represents the next stage—STMI v4.0—refined through updated theory, improved item architecture, and upgraded analytics. TQ retains the dual-scale design but redefines it within the framework of O-Theory:

- **Competence** measures the *organization of information* (O = DSRP)—how people actually structure their thinking.
- **Confidence** measures *meta-awareness of organization* (I_about_O)—how people perceive or believe they structure their thinking.

Competence and Confidence as Dual Scales

Where traditional psychometric instruments measure *traits*, TQ measures *skills*. Competence represents the measurable performance dimension; Confidence captures metacognitive self-assessment. Together, they form a two-layer model of cognitive performance—observable skill and self-awareness of that skill.

In earlier STMI studies, Competence and Confidence showed consistent internal reliability ($\alpha \approx .70-.80$ each) and moderate positive correlations ($r \approx .45-.55$), indicating that most people tend to overestimate their systems-thinking skill—a metacognitive blind spot that training can correct.

In the current study, we focused on the Competence scale for structural validation, as it is the true skill-metric component suitable for bifactor and latent modeling. Confidence remains part of the overall TQ instrument and will be examined relationally in the Discussion.

Purpose of the Present Study

The goal of this study was to provide further validation of TQ as a reliable and structurally coherent skill-metric for systems thinking. Specifically, we sought to:

1. Evaluate internal reliability of TQ across both scales (Competence and Confidence) using Cronbach's α and McDonald's ω_{total} .
2. Assess subscale reliabilities across the five Competence domains (D, S, R, P, M).
3. Estimate ordinal (Spearman-adjusted) reliability to test robustness across data types.
4. Conduct a bifactor confirmatory factor analysis (CFA) of the Competence items using polychoric correlations and WLSMV estimation to test for a dominant general systems-thinking factor (g) alongside DSRP-specific subskills.
5. Examine the conceptual and empirical relationship between Competence and Confidence as distinct but related dimensions of cognitive performance.

Collectively, these analyses test whether TQ satisfies the criteria of a valid skill-metric:

- **Reliable** (internally consistent)
- **Structural** (organized according to DSRP theory)
- **Trainable** (sensitive to improvement)
- **Measurable** (quantifies performance, not personality)

2.0 Methods

Testing the Validity and Reliability of Skill-Based Measurement

Prior research in DSRP Theory and O-Theory has already established that thinking is the *organization of information* ($T = I \rightarrow O$) [cite], and that this organization follows four universal structural patterns [cite]: making Distinctions, seeing Systems, recognizing Relationships, and taking Perspectives.

These patterns are not hypothetical—they have been empirically verified as the underlying mechanics of both cognition and reality.

The next question, and the one addressed here, is whether these structural patterns can be measured in people—not through attitudes or self-reports, but through actual *performance of thinking*. Can we observe, quantify, and track how well individuals organize information according to DSRP? Can we measure improvement over time as they practice and develop these skills?

The Thinking Quotient (TQ) was designed to answer precisely these questions. It evolved directly from this theoretical foundation to serve as a skill-metric—an instrument that measures *how well* people can

demonstrate the universal patterns of structural thinking in practice. Unlike psychometric tools that capture traits or preferences, TQ quantifies *can-do* ability: the cognitive organization a person can actually perform and improve.

Instrument Design

TQ consists of two parallel scales:

Competence Scale – 47 scored performance items measuring applied use of DSRP patterns in problem-solving, reasoning, and categorization tasks. Items are grouped into five subdomains: D (Distinctions), S (Systems), R (Relationships), P (Perspectives), and M (Mix). Each item has a single objectively correct target structure but is scored on a 0–4 scale to capture degrees of accuracy and completeness in reasoning—allowing partial credit for responses that demonstrate partial structural correctness.

Confidence Scale – parallel self-report items measuring metacognitive awareness and self-assessed ability across the same DSRP domains. Confidence items are rated on a 4-point scale (“Not confident” → “Highly confident”). Both scales are administered digitally via a digital platform. All items were developed through iterative expert review and pilot testing to align explicitly with DSRP theory, avoiding content overlap or verbal bias. Additional information about the original instrument design can be found in Cabrera, et. al [20].

Data and Participants

We analyzed the most recent version of the TQ dataset ($n = 289$) consisting of adult learners across diverse professional and educational settings. Participants completed both Competence and Confidence scales in a single session. Participants completed both scales in approximately 20 minutes on average ($M = 19.4$, $SD = 4.8$), with the Competence portion taking about 14 minutes and the Confidence portion about 6 minutes.

Data cleaning and validation steps included: removal of incomplete records (>30% missing data); conversion of all item scores to numeric format; standardization of response scaling across items; identification of valid subscale groupings based on “Question Block” (QB) mapping.

This version (STMI v4.0 → TQ) represents the culmination of four major iterations of the instrument (total $n \approx 1,580$ across versions), each improving item clarity, balance, and psychometric precision.

Analytical Procedures

Internal Reliability: Reliability quantifies the degree to which TQ consistently measures the same underlying construct. High α and ω demonstrate that TQ is not only conceptually coherent but also statistically stable. We computed Cronbach’s α and McDonald’s ω_{total} for both the full-scale and subscales using the *Competence* and *Confidence* datasets. Cronbach’s α was calculated as a measure of

internal consistency (equal-weight assumption). McDonald's ω_{total} was estimated using a principal-components-based approach to capture true-score variance without assuming item equivalence. Subscale α and ω values were computed for each of the five Competence domains (D, S, R, P, M).

Ordinal-Adjusted Reliability: If reliability remains strong under ordinal treatment, the instrument is robust to scaling assumptions—meaning its precision holds regardless of how responses are coded numerically. Because TQ item responses are ordinal (4-point), we re-estimated α and ω using Spearman correlations as a nonparametric proxy for polychoric correlations. This adjustment ensures that reliability estimates are not inflated or deflated by treating ordinal data as interval-level.

Bifactor Confirmatory Factor Analysis (CFA): We conducted a bifactor CFA on the Competence items using the *lavaan* package in R (estimator = WLSMV; ordered = TRUE). The model tested a dominant general factor (g) representing overall systems-thinking ability, alongside five specific orthogonal factors representing D, S, R, P, and M subskills. Model constraints included: general factor g loads on all items; specific factors load only on items within their domain, and; that all factors are orthogonal (covariances = 0). Model fit was assessed using CFI, TLI, RMSEA, and SRMR, with thresholds of $\geq .90$ for acceptable fit and $\geq .95$ for excellent fit. The bifactor CFA tests whether systems thinking is best represented by one general structural factor with domain-specific refinements—precisely what DSRP theory predicts.

Omega Hierarchical and Specific (ω_g , ω_{specific}): From the bifactor CFA, we extracted hierarchical reliability indices using `semTools::bifactorIndices()`. ω_g estimates the proportion of reliable variance explained by the general factor (TQ g). ω_{specific} estimates variance explained by each specific factor (D, S, R, P, M) beyond g. High ω_g confirms that systems thinking operates as a unified skill, while moderate ω_{specific} values confirm that each DSRP pattern adds measurable specificity without functioning independently.

Competence–Confidence Relationship: We computed correlations between average Competence and Confidence scores across domains to quantify metacognitive alignment ($r \approx .45\text{--}.55$). This relationship reflects the metacognitive gap: most individuals are more confident than competent. Tracking that alignment allows TQ to diagnose overconfidence and guide more effective learning interventions.

Data Transparency and Reproducibility

All analyses were conducted using established, open statistical tools (Python for reliability analyses, and R with *lavaan* and *semTools* for the bifactor CFA). The analytical workflow—including scoring procedures, code used for reliability estimation, and the full model specifications—is documented to ensure methodological clarity and reproducibility. Because the TQ instrument and its item-level content are proprietary, raw items are not publicly released; however, the structure of the analyses can be independently replicated using parallel ordinal and bifactor modeling approaches. All results were cross-validated using both parametric and nonparametric methods to confirm their robustness.

Table 2. Summary of Methodological Rigor and Validation Layers

Validation Layer	Analytical Test	Outcome	Implication

Internal Consistency	Cronbach's α , ω_{total}	High (.81–.88 overall)	TQ measures a coherent, unified skill.
Ordinal Robustness	Spearman α , ω_{total}	High (.81–.84)	Reliable regardless of response scaling.
Structural Fit	Bifactor CFA (CFI=.986, TLI=.985, RMSEA=.016)	Excellent model fit	Systems thinking operates as a unified structural skill.
Factor Reliability	$\omega_{\text{c}} \approx .70$, $\omega_{\text{m}} \approx .10\text{--}.25$	Strong general factor with measurable subskills	Confirms DSRP simultaneity.
Dual-Scale Validity	Competence–Confidence correlation $r \approx .50$	Positive but imperfect alignment	Supports dual cognitive–metacognitive architecture.

Together, these methods provide a rigorous evaluation of TQ as both a theoretical construct and a practical measurement tool. Whereas earlier versions (STMI) established the measurability of systems-thinking skill, the present study tests the structural validity and precision of that measurement—whether the instrument reliably captures the unified cognitive organization predicted by DSRP Theory and O-Theory, and whether its metrics are robust across scaling, structure, and context. The next section presents the empirical results, interpreted through JoST's dual lens of technical evidence and practical meaning.

3.0 Results

Evaluating Reliability and Structure

We analyzed reliability, dimensionality, and model fit to determine whether TQ measures a coherent, trainable skill. This section summarizes results for both *Competence* (skill-performance) and *Confidence* (metacognitive awareness), integrating classical reliability coefficients, ordinal adjustments, and the bifactor CFA outcomes.

Reliability Analyses

We calculated internal reliability using Cronbach's α and McDonald's ω_{total} for the full instrument and for each DSRP domain. Ordinal (Spearman-based) versions of these statistics were also computed to test robustness when treating items as ordered categories (Table 2).

Table 2. Overall and Subscale Reliability

Metric / Scale	Result	Plain-English Interpretation
Cronbach's α (overall, 57 items)	0.866	The total TQ score is highly consistent; items measure one coherent skill.
McDonald's ω_{total} (overall)	0.877	About 88 % of variance reflects real differences in systems-thinking skill.
Cronbach's α (mapped Competence items)	0.814	Even using only mapped DSRP items, reliability remains strong.
McDonald's ω_{total} (mapped Competence items)	0.843	Over 84 % true-score variance $\rightarrow$ precise, balanced design.
Subscale α (Competence)	D =.48 S =.49 P =.53 R =.59 M =.59	Moderate internal consistency; D and S could be refined.
Subscale ω_{total} (Competence)	D =.70 S =.65 P =.67 R =.74 M =.73	Stable subskills; R and M most reliable.
Spearman α (ordinal-adjusted)	≈ 0.81 overall	Reliability remains very strong with ordinal treatment.
Spearman ω_{total} (ordinal-adjusted)	≈ 0.84 overall	Consistency holds regardless of response scaling.
Confidence scale α / ω_{total}	$\alpha \approx 0.79$ $\omega \approx 0.82$	Metacognitive self-ratings are themselves reliable.
Competence–Confidence correlation	$r \approx 0.50$ ($p < .001$)	Moderate positive link—people are somewhat accurate yet often overconfident.

TQ shows reliability levels comparable to or exceeding established cognitive-skill measures. The close agreement between α and ω confirms that results are not an artifact of item weighting. The strong ordinal reliabilities demonstrate that the instrument's precision holds across scoring assumptions, an important property for digital, adaptive delivery.

Structural Modeling

To test the theoretical claim that systems thinking is a unified but multifaceted skill, we conducted a bifactor confirmatory factor analysis (CFA) on the Competence items. We specified a model with one general factor (g) loading on all items and five specific orthogonal factors representing D, S, R, P, and M. The analysis used polychoric correlations with the WLSMV estimator (Table 3).

Table 3. Model Fit and Factor Reliabilities

Index / Estimate	Result	Interpretation
CFI	0.986	Excellent comparative fit.
TLI	0.985	Excellent incremental fit.
RMSEA	0.016	Minimal residual error.
SRMR	0.111	Slightly above ideal (.08) due to within-domain item wording overlap.
ω_{total}	0.84	Overall latent reliability remains high.
$\omega_{\text{hierarchical}} (\omega_{\text{g}})$	≈ 0.70	$\approx 70\%$ of reliable variance explained by the general systems-thinking factor.
$\omega_{\text{specific}} (\omega_{\text{r}}, \omega_{\text{p}})$	$\approx 0.10\text{--}0.25$ each	Each DSRP domain contributes unique but smaller reliable variance.
Strongest specific factors	R (Relationships), P (Perspectives)	These subskills show the most distinct structure beyond g.

The CFA confirms the DSRP simultaneity principle: systems thinking functions as one organized skill composed of interdependent patterns. The dominant general factor (g) accounts for most true variance, while R and P add meaningful specialization. Excellent fit indices verify that the DSRP structure accurately models observed cognition.

Competence–Confidence Alignment

Competence measures structural performance; Confidence measures self-awareness of that performance. We examined how these relate across participants by correlating total and domain-level Competence scores with corresponding Confidence ratings. Across all domains, Confidence exceeded Competence on average by ≈ 0.6 points ($SD = 0.3$) on a 4-point scale. The mean correlation ($r \approx 0.50$) indicates partial but imperfect alignment between perceived and actual skill (Table 4). This metacognitive gap is consistent with expected cognitive bias patterns and provides the foundation for subsequent interpretation in the Discussion.

Table 4. Overall Summary of Findings

Finding	Empirical Evidence	Interpretation
TQ shows high reliability and precision.	$\alpha = .87$; $\omega = .88$	Measures a single coherent cognitive skill.

Competence subscales are measurable.	$\omega = .65-.74$ per domain	DSRP patterns function as structured subskills.
Model fit is excellent.	CFI = .986; RMSEA = .016	Confirms DSRP-based structural model.
General factor dominates.	$\omega^2 \approx .70$	Unified systems-thinking ability underlies all items.
Specific factors add nuance.	$\omega^2 \approx .10-.25$	R and P most distinct subskills.
Confidence scale reliable.	$\alpha = .79$; $\omega = .82$	Self-assessment dimension stable.
Competence–Confidence gap.	$r \approx .50$; Confidence > Competence	Reveals metacognitive bias; target for training.

Integrating Evidence and Meaning

- Scientific Significance:** This is the first empirical confirmation that the structural mechanics of DSRP — and by extension O-Theory — manifest as a single measurable latent skill.
- Practical Significance:** Because the TQ captures that structure reliably, practitioners can now measure *how* people think, track improvement over time, and design interventions that directly build systems-thinking capacity.
- Developmental Significance:** The correlation between Confidence and Competence closes the loop between cognition and metacognition, enabling individualized feedback for learning and leadership development.

The results confirm that TQ performs precisely as a skill-metric, not a psychometric trait test: it quantifies changeable cognitive structure. The next section discusses how these findings extend DSRP and O-Theory, their implications for practice, and how TQ enables a measurable science of systems-thinking development.

Table 5. Technical Results and Plain-English Interpretations for the TQ

Technical Result / Metric	Plain English Interpretation
Cronbach's α (overall, 57 usable items) = 0.866	The total TQ score is highly consistent. Responses across items align well, showing that the test measures a single, coherent skill rather than random noise.
McDonald's ω_{total} (overall) = 0.877	About 88% of score differences reflect real variation in systems-thinking skill, not measurement error — excellent reliability for a skill-based measure.

Cronbach's α (mapped Competence items) = 0.814	Even when using only the explicitly mapped D, S, R, P, and M items, the test remains highly reliable. The core Competence structure stands alone as a complete skill measure.
McDonald's ω_{total} (mapped Competence items) = 0.843	The mapped subset still captures over 84% true-score variance, confirming the internal precision and balanced design of the core test.
Subscale α (Competence) — D = .48, S = .49, P = .53, R = .59, M = .59	Each DSRP domain shows moderate internal consistency. They function best as parts of the total score, though Distinctions (D) and Systems (S) could be strengthened with more refined items.
Subscale ω_{total} (Competence) — D = .70, S = .65, P = .67, R = .74, M = .73	Each DSRP domain measures a stable, distinct skill pattern. Relationships (R) and Meta/Integration (M) are the most reliable subskills; all fall within acceptable early-stage ranges for skill metrics.
Spearman α (ordinal-adjusted) $\approx$ 0.81 overall	Treating items as ordinal (ranked) instead of numeric yields similar reliability. This supports the use of polychoric correlations and WLSMV estimation in confirmatory models.
Spearman ω_{total} (ordinal-adjusted) $\approx$ 0.84 overall	Confirms that the test remains consistent when modeled as ordinal data — TQ's structure is robust regardless of scale type.
Bifactor CFA (lavaan, WLSMV)	The model tested one dominant <i>general systems-thinking</i> factor (g) plus five domain-specific factors (D, S, R, P, M). Fit was excellent: CFI = .986, TLI = .985, RMSEA = .016, SRMR = .111.
$\omega_{\text{hierarchical}} \approx .65\text{--}.70$	Roughly two-thirds of the reliable variance in all items is explained by a single general TQ factor — meaning most of what TQ measures is unified systems-thinking ability.
ω_{specific} (D, S, R, P, M) $\approx$.10–.25 each	Each DSRP domain adds unique variance beyond the general skill, confirming that the four cognitive patterns are measurable but interwoven. Relationships (R) and Perspectives (P) show the strongest unique effects.
Fit indices summary — CFI = .986, TLI = .985, RMSEA = .016, SRMR = .111	The bifactor model fits the data extremely well. Slightly elevated SRMR indicates minor wording overlap among items, not structural misspecification.
Test–retest expectation (edumetric interpretation)	Because TQ measures a <i>learnable</i> skill, scores should rise with deliberate practice. Growth reflects learning and improved cognitive structure — not instability or error.

Overall interpretation of results	TQ is a reliable, valid, and learnable measure of systems-thinking skill. The total TQ score best represents overall cognitive organization, while D, S, R, P, and M provide meaningful, trainable subskills within that unified structure.
--	---

The bifactor CFA confirms that:

- *A single general systems-thinking factor* accounts for most reliable variance ($\omega_{\square} \approx .7$).
- *DSRP subskills (i:o, p:w, a:r, v:p) are real but secondary*, supporting the theory that they operate simultaneously within one structural field.
- *Model fit is excellent*, validating the internal architecture of the TQ as a modern edumetric instrument.

4.0 Discussion

From Conceptual Ideal to Measurable Skill

For decades, systems thinking has been described as vital yet remained largely unmeasurable—a “black box” between knowing and doing. Theoretical models proliferated, but few demonstrated empirical precision or developmental measurability.

Our program of research, summarized in the *Top 10 Findings in Systems Thinking*, revealed that cognition itself is organized through universal structural patterns. Those patterns—Distinctions, Systems, Relationships, and Perspectives—constitute the deep architecture of both mind and nature.

The Thinking Quotient (TQ) operationalizes this discovery. It measures *organization* rather than *opinion*, making thinking visible and improvable. This study’s results confirm that systems thinking can now be treated as a *skill*—structured, learnable, and quantifiable.

Scientific Integration

1. **Existence and Structure Confirmed:** High reliability ($\alpha = .87$; $\omega = .88$) and excellent model fit (CFI = .986; RMSEA = .016) empirically confirm the existence of a unified systems-thinking skill. The bifactor CFA demonstrated that most reliable variance ($\omega_{\square} \approx .70$) resides in a single general factor—exactly as DSRP Theory predicts—while each DSRP domain contributes smaller, meaningful specialization ($\omega_{\square} \approx .10-.25$).
2. **DSRP Simultaneity:** The pattern of loadings validates that Distinctions, Systems, Relationships, and Perspectives co-occur rather than operate independently. Activating one pattern engages the others; cognition is structurally entangled. This is empirical proof of *DSRP simultaneity*, a cornerstone of O-Theory.

3. **Dual-Scale Coherence:** The Confidence scale showed comparable reliability ($\alpha = .79$; $\omega = .82$) and moderate correlation with Competence ($r \approx .50$). This confirms the predicted dynamic: cognition (Competence) and metacognition (Confidence) form a feedback system. Most people overestimate their skill, but that gap closes with explicit structural feedback.
4. **Alignment with the Cognitive Law ($M = I \cdot O$):** The results substantiate that understanding depends multiplicatively on both information (I) and organization (O). TQ directly measures O, providing the missing half of the cognitive equation. Improvements in TQ scores represent measurable increases in structural organization capacity.
5. **Consistency with the Cognitive Pareto Law and TAP:** The discovery that a small set of structural operations (the Six Moves) governs most cognitive performance aligns with the high ω observed here: a small number of structural skills account for most reliable variance in thinking performance. This confirms the 20/80 Thinking Advantage Principle in a measurement context.

From Measurement to Development

1. Establishing a Science of Thinking: TQ provides the empirical substrate for a *science of systems thinking*. For the first time, we can quantify how people organize information, not merely what they know. This transforms systems thinking from philosophy to applied cognitive science.

2. From Psychometrics to Skill-Metrics: Traditional psychometrics measure stable traits; TQ measures *changeable structure*. Because TQ is edumetric, improvement over time is expected and desirable. Rising scores indicate learning and structural reorganization—not unreliability.

3. Bridging Theory and Practice: The TQ unites the theoretical architecture of O-Theory with the practical mechanics of the Six Moves. Practitioners can now use the same metric to diagnose, train, and verify cognitive improvement. It closes the loop between theory (how thinking works) and practice (how to build it).

4. Implications for Education, Leadership, and Systems Work

- **Education:** Enables standards-aligned measurement of systems-thinking skill across K-16 using the same DSRP-based architecture.
 - **Leadership & Organizations:** Provides a quantitative index of organizational cognition—how teams structure problems, decisions, and strategies.
- Research & Evaluation:** Offers a reproducible benchmark for interventions claiming to “improve systems thinking.” Claims can now be empirically tested.

5. Metacognitive Feedback: The dual Competence–Confidence structure allows learners to see where perception diverges from performance. That visibility transforms training from abstract reflection into structural feedback—making thinking and self-awareness mutually reinforcing.

6. Structural Dunning–Kruger Dynamics and the Pareto Bias of Elements: The

Competence–Confidence gap observed in TQ reflects a structural form of the Dunning–Kruger effect, consistent with the Cognitive Pareto Law and the bias of elements discovered in prior O-Theory research. Because thinking is structured by DSRP patterns, overestimation of one’s own systems-thinking ability emerges when those patterns are unevenly developed—when certain elements dominate or others are underrepresented. This structural imbalance limits the capacity for accurate self-assessment. Within the Love–Reality Loop, the resulting overconfidence is not a personality flaw but a predictable model–world misalignment ($\mathbb{M} \neq \mathbb{R}$): the mind’s organization of information does not yet match reality’s organization.

The Six Pareto Moves, identified in earlier studies as the 20% of structural operations that surface 80% of recurring cognitive bias, directly target this misalignment. Practicing the Moves strengthens the underused elements of thought, thereby increasing Competence, calibrating Confidence, and closing the metacognitive gap. This finding demonstrates that TQ not only diagnoses structural bias but also provides a developmental measure of bias reduction over time.

Advancing a Measurable [Systems] Thinking Science

Bifactor modeling is not just a statistical formality—it is a structural test of theory. The excellent fit indices confirm that the DSRP architecture accurately predicts the covariance structure of thinking behavior. In cognitive-science terms, it verifies that the same underlying “grammar” organizes both thought and reality ($\mathbb{M}/\mathbb{R} \approx 1$).

The dominant general factor corresponds to the *organization of thought*, while the specific DSRP factors represent recursive expressions of that organization. This empirically grounds the theoretical claim that thinking is fractal, co-implicated, and dynamically self-similar across scales.

1. **Longitudinal Learning Curves** Because TQ is edumetric, future work will track changes in TQ scores across training programs to model the rate and shape of cognitive-skill acquisition ($\Delta\mathbb{M} = \mathbb{I} \rightarrow \mathbb{O}$ over time).
2. **Cross-Domain Validation** Replicating TQ across disciplines (STEM, management, ecology, policy) will test its universality as a measure of structural cognition.
3. **AI and Cognitive Modeling** Integrating TQ data with machine-learning models can simulate structural cognition, advancing adaptive training systems grounded in real cognitive architecture.
4. **Expanding Confidence Analytics** Future analyses will explore confidence calibration curves—mapping where learners are over- or under-confident—to optimize metacognitive training.
5. **Structural Benchmarks for Systems Thinking Standards** TQ provides the measurement base for O1–O7 Standards and CanDo Levels, enabling cumulative certification of thinking capability.

Conclusion

The results close a long-standing gap in the systems-thinking field: we now have both the theory (O-Theory and DSRP) and the tool (TQ) to measure and build systems thinkers. Empirically, TQ

demonstrates that systems thinking is not mystical, intuitive, or purely philosophical—it is a structured, learnable operation of mind. Conceptually, this study integrates decades of theoretical insight into a single measurable law of cognition ($\mathbb{M} = \mathbb{I} \cdot \mathbb{O}$). Practically, it equips educators, leaders, and scientists with a quantitative metric for human thinking performance. Culturally, it reframes systems thinking from an aspirational value to an evidence-based discipline that can be trained, tested, and improved.

The Thinking Quotient (TQ) represents a milestone in the maturation of systems-thinking science. It validates, through empirical evidence, that DSRP and O-Theory describe not only how reality is organized but also how the mind organizes reality. By measuring that organization directly, TQ transforms systems thinking from an art of insight into a measurable, improvable skill—one that connects the structure of thought to the structure of the world.

References

1. Cabrera D, Cabrera L. The structure of thought and reality: empirical evidence for the three universal laws linking mind, matter, and meaning. *Journal of Systems Thinking*. 2025;5: 1–20. doi:10.54120/jost.v25i1.dc6smio6
2. Cabrera D. The Unified Calculus of organization, structure, emergence, and meaning: The patterns that connect mind and nature. *Journal of Systems Thinking*. 2025;5: 1–41. doi:10.54120/jost.2025-000002
3. Cabrera D, Cabrera L. The grammar of reality: The organizing principles of structure and emergence. *Journal of Systems Thinking*. 2025;5: 1–18. doi:10.54120/jost.2025-000003
4. Cabrera D, Cabrera LY. Any Person, Any Study: A Different Kind of Theory of Everything (ToE). *Journal of Systems Thinking*. 2023; 1–14. doi:10.54120/jost.0000030
5. Cabrera D. A Mathematical Theory of Organization: DSRP as Universal Code of Mind and Nature for Organizing, Evolving, and Understanding Information. *Journal of Systems Thinking*. 2024.
6. Cabrera D. *Systems Thinking*. Cornell University; 2006.
7. Cabrera D. Method of teaching thinking skills and knowledge acquisition. Google Patents; 2008. Available: <http://www.google.com/patents/US20080187899>
8. Cabrera D. Systems and methods of mapping, transforming, and sharing data and ideas. US Patent. 11200262, 2021. Available: <https://patentimages.storage.googleapis.com/1c/78/51/b10275012a82bd/US11200262.pdf>
9. Cabrera D, Cabrera LL, Cabrera E. A Literature Review of the Universal and Atomic Elements of Complex Cognition. *Journal of Systems Thinking*. 2023; 1–85. doi:10.54120/jost.0000032
10. Cabrera D, Cabrera L, Cabrera E. Distinctions Organize Information in Mind and Nature: Empirical Findings of Identity–Other Distinctions (D) in Cognitive and Material Complexity. *Systems*. 2022;10. doi:10.3390/systems10020041
11. Cabrera D, Cabrera L, Cabrera E. Systems Organize Information in Mind and Nature: Empirical Findings of Part-Whole Systems (S) in Cognitive and Material Complexity. *Systems*. 2022.

12. Cabrera D, Cabrera L, Cabrera E. Relationships Organize Information in Mind and Nature: Empirical Findings of Action-Reaction Relationships (R) in Cognitive and Material Complexity. Systems. 2022.
13. Cabrera D, Cabrera L, Cabrera E. Perspectives Organize Information in Mind and Nature: Empirical Findings of Point-View Perspective (P) in Cognitive and Material Complexity. Systems. 2022.
14. Cabrera D, Cabrera L, Cabrera E. The “Fish Tank” Experiments: Metacognitive Awareness of Distinctions, Systems, Relationships, and Perspectives (DSRP) Significantly Increases Cognitive Complexity. Systems. 2022;10. doi:10.3390/systems10020029
15. Cabrera D, Silberman D, Cabrera L. Popularity Without Proof: A Systematic Empirical Review of Systems Thinking Frameworks. Systems Thinking. Journal of Systems Thinking. 2025;5: 1–30. doi:10.54120/jost.2025-000001
16. Cabrera D. Critical Thinking: Broken Design Diagnosing Flaws Our Most Trusted Cognitive Tool. Journal of Systems Thinking. 2025;2025: 1–18.
17. Cabrera D CL. The Pareto structure of thought: empirical discovery of the six foundational mental moves. J Syst Think. 2025;2025: 1–41. doi:10.54120/jost.v25i1.dc6par6
18. Di Marco G, Cabrera D. Common biases in systems thinking. Journal of Systems Thinking Preprints. 2024. doi:10.54120/jost.pr000037.v1
19. Aljarri W 'ad, Bustamante J, Ingwell A, Labrada T, Nadig B, Shin J, et al. New Hope for policy schools: Examining systems thinking solutions to improve policy education. Journal of Systems Thinking. 2024; 1–61. doi:10.54120/jost.00000101
20. Cabrera L, Sokolow J, Cabrera D. Developing and Validating a Measurement of Systems Thinking : The Systems Thinking and Metacognitive Inventory (STMI). Journal of Systems Thinking. 2023; 1–43. doi:10.54120/jost.0000042