

Mental Fitness and the Structure of Adaptation

Evidence from a Systematic Review Toward a Measurable, Trainable Science of Mental Fitness

Derek Cabrera^{1,2,@}, Sriharan Lakshmanaprasath^{3,2}, Dave Silberman^{2,4}, Laura Cabrera^{1,2}

¹ Cornell University

² Cabrera Research Lab

³ George Washington University

⁴ Boston University

@ dac66@cornell.edu

ABSTRACT: Mental fitness is increasingly invoked as a proactive answer to escalating cognitive strain and burnout, yet the term is often treated as a slogan rather than a measurable construct. To test whether “mental fitness” is becoming scientifically coherent—or remaining a loose rhetorical umbrella—we conducted a citation-weighted systematic literature review of 46 highly cited, conceptually relevant studies spanning psychology, neuroscience, education, health, and performance contexts. We use mental fitness rather than brain fitness to emphasize that the target is not the organ in isolation, but embodied, situated, and connected adaptive functioning—how cognitive organization is expressed through coupled emotional, motivational, and physical constraints in real contexts. Instead of asking whether any single intervention “works,” we ask a more foundational question: what is the structure of the mental-fitness literature itself? Three review hypotheses evaluated (i) whether the field exhibits conceptual convergence, (ii) whether measurement has stabilized or remains fragmented, and (iii) what the joint pattern implies for cumulative science. Results show striking convergence in how mental fitness is framed—with near-universal overlap across four recurring elements: systemic metacognition, repetitive practice, multidomain integration, and adaptive capacity—despite the absence of a single formal definition or unified mechanism. In sharp contrast, the field remains measurement-fragmented, distributing operationalizations across six distinct modes with no dominant standard. This combination—shared conceptual grammar without shared measurement infrastructure—marks an intermediate stage of construct maturity and implies a non-trivial need for a convergent, edumetric, performance-sensitive metric capable of tracking change with practice and performance-in-context. We argue that this emerging profile also clarifies a practical cognitive phenomenon underexplained by dual-process models: a third mode of cognition—System III thinking—in which structured metacognition and practice enable fast, accurate, and adaptable performance under pressure. Together, the findings reposition mental fitness as an emergent, empirically tractable science and define the constraints that any candidate metric or training system must satisfy to move the field forward.

KEYWORDS: mental fitness | systems thinking | metacognition | adaptive capacity | measurement fragmentation | edumetrics

Cite this paper: Cabrera, D., Lakshmanaprasath, S., Silberman, D., & Cabrera, L. (2025). Mental fitness and the structure of adaptation: Evidence from a systematic review toward a measurable, trainable science of mental fitness. *Journal of Systems Thinking*, 5(1), 1–40. <https://doi.org/10.54120/jost.v25i1.dc6mfsa6>

1.0 Introduction

Indicators of psychological strain continue to rise despite decades of advances in neuroscience, counseling, psychology, and public health. Anxiety, depression, burnout, and cognitive overload now affect a substantial portion of the population across personal and professional domains. In 2022, 18.2% of U.S. adults reported symptoms of anxiety and 21.4% reported symptoms of depression, reflecting a sharp post-pandemic increase [1]. In the workplace, burnout rates reached historic highs by 2025, with 66% of employees reporting burnout, and more than half describing their levels of burnout, depression, or anxiety as moderate to severe [2]. Notably, reports of high engagement often co-occur with exhaustion, suggesting a paradox of motivation without sustainability [3,4]. Globally, nearly half of workers report taking time off for mental health reasons [5].

Taken together, these patterns suggest that prevailing approaches remain insufficient, fragmented and largely reactive in face of real-world complexities. With an emphasis focused on managing symptoms after they appear rather than building durable, transferable cognitive and psychological capacity. This gap exposes a growing developmental need in mental fitness as a developmental alternative.

1.1 Mental Fitness as a Review Problem

Mental fitness is increasingly framed as a proactive counterpart to mental health—emphasizing capacity, resilience, and adaptability rather than pathology. However, unlike established psychological constructs, mental fitness lacks a stabilized definition, shared theoretical core, or standardized measurement infrastructure. The term is applied to a wide range of practices, including cognitive training programs, mindfulness interventions, persistence- and grit-based models, and multidomain wellness initiatives. Mental fitness vs. brain fitness (one paragraph for the introduction — insert early, right after you first define/position “mental fitness” as the focal construct): Although brain fitness is common in popular and commercial discourse, it is conceptually narrower than what the scientific literature is actually invoking. Brain fitness foregrounds neural substrate and invites a single-organ optimization frame; mental fitness foregrounds whole-person adaptivity—the capacity to stay aligned with situational reality as conditions change. In this review, mental fitness is treated as embodied and coupled: cognitive organization is a primary capacity term, but its real-world expression is gated by interacting emotional, motivational, and physical constraints. This framing avoids reducing performance to “cognition alone,” while keeping cognition structurally central as the locus of explicit organization and metacognitive correction that can be trained and measured (e.g., via organization-sensitive metrics).

This diversity raises a foundational question: is mental fitness a coherent construct with shared structural assumptions, or an umbrella term applied to conceptually distinct and only loosely related approaches? Without conceptual convergence and measurement coherence, cumulative progress is constrained: interventions cannot be meaningfully compared, mechanisms remain opaque, and claims of effectiveness are difficult to evaluate, leaving the field unable to determine what should be trained, how improvement should be measured, or whether adaptation actually transfers to real-world performance.

The purpose of this paper is therefore not to advocate for a particular intervention or to assert causal claims about mental fitness outcomes. Instead, we examine whether the mental-fitness literature demonstrates the structural coherence, conceptual convergence, and measurement conditions required for classification as an emerging scientific construct.

1.2 Review Logic and Hypotheses

To address this question, we conducted a systematic literature review (SLR) [6] designed to test three review-level hypotheses about the structure of the field:

- *H1: Conceptual Convergence.* If mental fitness is functioning as an emerging construct rather than a rhetorical label, the literature should cluster around a limited set of recurring conceptual framings rather than proliferating idiosyncratic definitions.
- *H2: Measurement Fragmentation.* If the construct is immature, measurement approaches should remain fragmented, with no dominant or de facto standard.
- *H3: Directional Implication.* If the literature converges conceptually on trainable, adaptive, performance-in-context capacities while remaining fragmented in measurement, this jointly implies a practical need for a convergent, edumetric mental-fitness metric.

These hypotheses are structural rather than causal. They evaluate observable patterns in how mental fitness is conceptualized and operationalized, not whether any given program or intervention “works.”

1.3 Anticipated Conceptual Core

Although construct formalization is deferred to the Results, preliminary review of the literature suggests repeated emphasis on four recurring principles:

1. *Systemic metacognition* — the capacity to reorganize thought using explicit structural patterns rather than merely observing mental content. This extends beyond awareness-based models of mindfulness, which often lack mechanistic specificity or measurable restructuring processes [7].
2. *Repetitive practice* — structured, feedback-rich engagement that produces durable cognitive change, consistent with findings from learning science and neuroscience [8–12].

3. *Multidomain integration* — coordination across cognitive, emotional, motivational, and physical domains, supported by evidence linking physical activity, emotion regulation, and motivation to cognitive performance [13,14] .
4. *Adaptive capacity* — flexible adjustment to changing real-world conditions, distinguishing adaptability from persistence alone and extending beyond grit-based framings .[15–17]

Crucially, these principles rarely appear as an integrated system within individual studies. Instead, they tend to surface in overlapping but partial configurations, raising the question of whether this overlap reflects superficial similarity or genuine conceptual convergence.

1.4 From Dual-Process Models to System III Thinking

Much of contemporary cognition research is organized around the distinction between fast, automatic processing and slow, deliberative reasoning. While this dual-process framing is descriptively useful, it leaves unresolved a practical problem: how skilled performers achieve speed *and* accuracy under pressure.

Across domains such as surgery, military training, and complex decision-making, structured practice and metacognitive skill development have been shown to support reliable performance in high-stakes environments [18,19]. These findings suggest the possibility of a third mode of cognition—here termed *System III thinking* [20]—in which structured metacognition and practice enable fast, accurate, and adaptable performance.

Table 1. Comparing System I, System II, and System III Thinking

Thinking System	Description	Benefits	Costs	Net Effect
System I Thinking (fast, error-prone, automatic)	Intuitive, effortless, autonomic responses shaped by habit and heuristics.	Very fast; requires no effort or training; efficient for routine survival tasks.	Highly error-prone; biased; oversimplifies complexity.	Speed without accuracy; requires no effort.
System II Thinking (slow, less error, requires effort)	Conscious reasoning and reflection, engaged when intuition is insufficient.	Less error-prone; reflective; provides checks on System I.	Slow and effortful; cognitively costly; lacks specificity for fast or complex situations.	Accuracy without speed; effort required.
System III Thinking (fast, accurate, requires effort)	Structured practice that uses systemic metacognition (DSRP & mental movements) to align thinking with real-world complexity.	Combines speed and accuracy; adaptable; produces reliable judgments under pressure; complexity-friendly.	Requires deliberate training, feedback, and continual practice; not innate or autonomic.	Speed <i>and</i> accuracy; practice effort required.

In this paper, System III is not introduced as a speculative theory but as a *directional implication* to be evaluated in light of the SLR results.

1.5 Structure of the Paper

Section 2 details the systematic review methodology [6] and formal review hypotheses. Section 3 reports results on conceptual convergence, measurement fragmentation, and their joint implications. Section 4 interprets these findings, evaluates their significance for theory and measurement, and discusses the requirements for a convergent, performance-based mental-fitness metric.

Together, the paper reframes mental fitness not as a collection of popular but partial framings—such as narrow cognitive training, awareness-only mindfulness, or persistence-based grit—but as an emergent, structurally coherent domain capable of empirical maturation and practical application.

2.0 Methods

This review used a citation-weighted, relevance-ranked corpus construction approach to assess the current condition of knowledge on the emerging construct “mental fitness” for placement on the KMMM [6,21]. A pilot term-verification step reduced an initial set of 35 candidate terms to a finalized 17-term Search Terms List (v2.0). Each term was queried in Google Scholar using Publish or Perish, and the top 50 most-cited results per term were retained to prioritize the most widely circulated and influential literature. Across all terms, this yielded an initial pool of approximately 850 records (with overlap across queries).

Records were screened at the abstract level using a four-factor relevance logic: studies were retained if they satisfied at least one of the following—(a) direct alignment with the project’s three-part working definition of mental fitness (metacognition, repetitive practice, organizational patterns), (b) empirical or conceptual contributions that expanded or clarified the construct, (c) evidence indicating definitional ambiguity in the field, or (d) explicit contradiction of the working definition (retained for dialectical value). Borderline cases were retained only if they were highly cited. Records were removed if they lacked a meaningful conceptual or empirical connection to mental fitness, were narrowly confined to clinical pathology without broader relevance to fitness/adaptability, or were duplicate/low-signal entries. Applying this process yielded a final analytic corpus of 46 studies.

Each included study was coded on four dimensions: (1) disciplinary origin, (2) methodological approach (observational, experimental, theoretical, mixed), (3) empirical status (yes/no), and (4) thematic alignment with mental fitness subcomponents. Codes were aggregated descriptively (counts and proportions) to characterize the corpus and to provide baseline context for the subsequent H1–H3 analyses.

2.3 Hypotheses

Because this study is a systematic literature review (SLR), the hypotheses tested here are not causal hypotheses about mental fitness itself. They are *review hypotheses* about what *patterns should be observable in the literature* if “mental fitness” is functioning as an emerging, coherent construct rather than a loose or purely rhetorical label. Each hypothesis is framed as a falsifiable claim about (i) conceptual clustering, (ii) measurement infrastructure, and (iii) directional implications visible in the SLR corpus. Corresponding null hypotheses (H_0) are specified to make the interpretive logic explicit and testable.

Table 2. Review hypotheses expressed in symbolic logic and plain-English form. Hypotheses are formulated as falsifiable claims about observable structure in the SLR corpus (not causal claims about mental fitness).

Hypothesis	Alternative (H_A) & Null (H_0) Hypothesis (Symbolic)	Hypothesis (Description)
H1 Conceptual Convergence	H1_A : $\exists k$ such that $ F = k$ and $k \ll L $ H1₀ : $ F \approx L $	The literature clusters into a limited set of recurring conceptual framings (a small typology), rather than being conceptually unique paper-by-paper.
H2 <i>Measurement Fragmentation</i>	H2_A : $ \text{Modes} \geq m \wedge \max_{x \in \text{Modes}} \text{Share}(x) < \sigma$ H2₀ : $\max_{x \in \text{Modes}} \text{Share}(x) \geq \sigma$	The mental-fitness literature uses multiple measurement approaches, and no single approach has become the field's de facto standard.
H3 Directional Implication from Field Structure	H3_A : $\text{Fragmentation}(\text{Modes}) = \text{True} \wedge \text{Orientation}(\text{Trainable} \wedge \text{Adaptive} \wedge \text{Performance}) = \text{High} \rightarrow \text{Need}(\text{ConvergentMetric}) = \text{True}$ H3₀ : $\text{Fragmentation}(\text{Modes}) = \text{False} \vee \text{Orientation}(\text{Trainable} \wedge \text{Adaptive} \wedge \text{Performance}) = \text{Low} \rightarrow \text{Need}(\text{ConvergentMetric}) = \text{False}$	If the mental-fitness literature converges conceptually on adaptive, trainable, performance-in-context capacities while remaining fragmented in measurement, this jointly implies a practical need for a convergent mental-fitness metric with an edumetric, performance-based, and integrative design.

2.2.1 H1 Methods: Conceptual Convergence

We first tested whether the mental-fitness literature groups into a limited set of recurring conceptual framings rather than exhibiting an unlimited number of highly idiosyncratic definitions. Formally, let L be the set of included mental-fitness studies. Let $f(l)$ denote the conceptual framing assigned to study l , and let $F = \{f(l) \mid l \in L\}$ be the set of distinct framings observed in the corpus. The null hypothesis ($H1_0$) is that framings are diffuse, such that $|F| \approx |L|$ (i.e., most studies introduce a unique or effectively unique framing). The alternative hypothesis ($H1_A$) is that framings are

clustered, such that $|F| \ll |L|$ (i.e., many studies share the same framing types). The testable claim is that there exists a small k such that $|F| = k$ and $k \ll |L|$; if most studies map to a small number of recurring framings, H1 is supported.

H1 examined whether the mental-fitness literature clusters into a limited set of recurring conceptual framings rather than exhibiting highly idiosyncratic definitions. The unit of analysis was the individual study. The corpus comprised the full-text PDFs of all 46 papers included in the systematic review.

A dictionary-based content analysis was employed. Four conceptual constructs were specified *a priori*: Multidomain Integration, Systemic Metacognition, Adaptive Capacity, and Repetitive Practice. Each construct was operationalized as a binary variable (present/absent) using explicitly defined lexical triggers and logical decision rules. All keyword lists and rules were frozen prior to full-text analysis and applied uniformly across the corpus.

Full texts were converted to machine-readable form and searched using case-insensitive pattern matching. Within each construct, keywords were evaluated using OR logic unless otherwise specified. Multidomain Integration was coded as present if a paper either explicitly used the term *multidomain* (or equivalent) or referenced at least two of four functional domains (cognitive, emotional, motivational/conative, physical/physiological). The remaining constructs were coded based on explicit reference to their defining conceptual features. No semantic inference, weighting, or negation handling was applied.

For auditability, trigger-level indicators were logged for all constructs, including domain-specific hits and qualification pathways. Descriptive statistics were used to summarize construct prevalence, overlap, and domain distributions. No inferential statistics were applied, as H1 concerns conceptual clustering rather than causal or population inference. Full keyword dictionaries and formal decision rules are provided in Appendix A.

Limitations. The H1 analysis relies on dictionary-based detection of explicit language and therefore captures *conceptual presence* rather than theoretical integration, mechanistic specificity, or causal commitment. The use of full-text PDFs and a permissive multidomain threshold increases sensitivity to cross-domain vocabulary and contributes to high construct prevalence. This saturation reflects the hybrid linguistic character of the mental-fitness discourse rather than post-hoc expansion of coding rules, which were frozen prior to analysis. Importantly, H1 does not test the strength, coherence, or effectiveness of these constructs, only their recurrence as framing elements. Subsequent hypotheses address differentiation, structure, and empirical coherence beyond lexical co-occurrence.

2.3.2 H2 Methods: Measurement Fragmentation

H2 examined whether the mental-fitness literature exhibits convergence on a single dominant measurement approach or remains fragmented across multiple measurement modes. The unit of

analysis was the individual study. The corpus comprised the full-text PDFs of all 46 papers included in the systematic review.

Formally, let L be the set of included studies. Let $m(l)$ denote the primary measurement mode used by study l to operationalize or empirically assess mental fitness. Let $Modes = \{m(l) \mid l \in L\}$ be the set of distinct measurement modes observed in the corpus. For each mode $x \in Modes$, let $Share(x)$ be the proportion of studies for which $m(l) = x$. The null hypothesis ($H2_0$) is that the field converges on a dominant measurement mode, such that there exists a mode x for which $Share(x) \geq \sigma$, where σ represents a de facto standard threshold. The alternative hypothesis ($H2A$) is that measurement remains fragmented, such that multiple modes exist and no single mode meets or exceeds this threshold.

Measurement modes were defined *a priori* as mutually exclusive, high-level categories reflecting the primary empirical operationalization of mental fitness. The mode taxonomy included: (i) psychometric self-report scales, (ii) performance-based or cognitive task measures, (iii) physiological or biomarker proxies, (iv) behavioral or observational coding schemes, (v) composite or index-based measures combining multiple indicators, and (vi) qualitative-only operationalizations (e.g., interviews, thematic analysis). Each paper was assigned exactly one primary mode based on the measurement approach that directly generated the study's principal results. When multiple measures were used with equal analytic weight, the study was coded as composite.

Full-text PDFs were converted to machine-readable form and searched for explicit measurement descriptions, including instrument names, task descriptions, physiological indicators, and analytic procedures. Mode assignment was rule-based and relied solely on explicit textual evidence; no inferential interpretation of author intent or construct validity was applied. For auditability, representative text excerpts supporting each mode classification were logged at the paper level.

Descriptive statistics were used to summarize the number of distinct measurement modes and the proportional share of each mode across the corpus. The primary test of $H2$ compared the observed maximum mode share to the predefined dominance threshold σ . No inferential statistics were applied, as the hypothesis concerns observable structural properties of the reviewed literature rather than population-level estimation.

Limitations. The $H2$ analysis classifies measurement at the level of broad mode rather than specific instruments or psychometric properties. Consequently, it captures structural convergence or fragmentation in measurement approach, not quality, validity, or effectiveness of the measures themselves. Studies employing multiple instruments within a single mode are treated as convergent at the mode level, even if substantial heterogeneity exists at the instrument level. This conservative framing biases the test *against* finding fragmentation; any observed nonconvergence therefore reflects robust structural dispersion rather than fine-grained methodological disagreement.

2.3.3 H3 Methods: Directional Implication from Field Structure

H3 does not introduce additional descriptive or empirical coding beyond that performed for H1 and H2. Instead, it evaluates a *directional implication* that follows from the joint structure of the literature as established by those prior analyses. Specifically, H3 tests whether the combination of (i) conceptual convergence around mental fitness as an adaptive, trainable, performance-in-context capacity (H1), and (ii) persistent fragmentation of measurement approaches with no dominant standard (H2), implies a practical need for a convergent mental-fitness metric with a constrained measurement profile.

Formally, H3 is evaluated as a *conditional inference* rather than as a statistical or causal hypothesis. Let Fragmentation(Modes) denote the absence of a dominant measurement mode in the literature, as operationalized and tested in H2. Let Orientation(Trainable $\wedge$ Adaptive $\wedge$ Performance) denote high conceptual convergence on mental fitness as a capacity characterized by learning, adaptation, and context-sensitive performance, as established by H1. H3 tests whether the joint satisfaction of these conditions implies Need(ConvergentMetric)=True.

No new variables were coded and no additional frequency or inferential analyses were conducted for H3. Instead, a predefined inference rule was applied:

If a field conceptualizes its target construct as trainable and performance-based, yet lacks a convergent or standardized measurement infrastructure, then cumulative progress, cross-study comparability, and skill-development tracking are structurally constrained unless a convergent metric exists.

The test of H3 therefore consists of verifying whether both antecedent conditions are satisfied (Sections 3.1 and 3.2) and assessing whether the above inference rule applies. H3 is supported if the literature simultaneously exhibits high orientation toward trainable, adaptive, performance-in-context mental fitness and persistent measurement fragmentation, thereby implying a non-trivial need for a convergent mental-fitness metric.

This approach preserves falsifiability: H3 would not be supported if either conceptual convergence were low (i.e., mental fitness were treated primarily as a static trait) or if measurement convergence were high (i.e., a dominant standard already existed). Importantly, H3 does not evaluate or validate any specific instrument. The identification of candidate metrics that may satisfy the derived measurement requirements is deferred to the Discussion.

3.0 Results

The final analytic sample comprised *46 studies* drawn from an initial pool of approximately *850 records*, reflecting a substantial narrowing from broad discourse to a relatively small core of highly cited and conceptually relevant work. The resulting corpus indicates that mental fitness remains an emergent and weakly consolidated construct rather than a mature, standardized field.

By disciplinary origin, the literature is heavily concentrated in *psychology* (34/46; 73.9%). *Education* accounts for a secondary cluster (5/46; 10.9%), followed by *medicine/health sciences* (3/46; 6.5%). *Biology* and *computer science* are sparsely represented (2/46 each; 4.3%). This distribution indicates that mental fitness is predominantly theorized and operationalized within psychological paradigms, with limited cross-disciplinary integration from biological, educational systems, or computational perspectives.

Methodologically, 31 of 46 studies (67.4%) were coded as empirical, while 15 (32.6%) were conceptual or theoretical. Among methodological approaches, *quantitative designs* were most common (16/46; 34.8%), followed by *theoretical reviews or frameworks* (13/46; 28.3%) and *mixed-methods studies* (9/46; 19.6%). *Qualitative-only* designs were comparatively rare (6/46; 13.0%), and *randomized or tightly controlled experimental studies* were minimal (2/46; 4.3%). This pattern suggests that the literature is better developed for describing associations, constructs, and proposed mechanisms than for establishing strong causal claims.

Thematic coding revealed the greatest concentration of studies in *measurement tools and frameworks* (10 studies) and *training, learning, and transfer* (8 studies). Additional clusters included *conceptualization of mental fitness* (6 studies), *intervention-based research* (6 studies), and *neurobiological or physiological correlates* (6 studies). Fewer studies directly addressed long-term adaptability, transfer across domains, or structural mechanisms underlying mental fitness, indicating fragmentation across subcomponents rather than a unified theoretical core.

Overall, the general results show a field characterized by high conceptual activity, moderate empirical presence, and limited experimental rigor. Mental fitness is studied primarily through psychological lenses, with substantial effort devoted to defining and measuring the construct, but relatively few studies capable of isolating causal mechanisms or demonstrating generalizable effects of specific mental fitness interventions. These baseline results contextualize the hypothesis-specific analyses summarized in Table 3.

Table 3. Summary of hypothesis results (common analytic corpus). All hypothesis evaluations use the same full-text analytic corpus (n = 46). Evidence bases are derived from hypothesis-specific coding of this shared sample; H3 is a logical synthesis of H1 and H2.

Hypothesis statement (short form)	Primary result	Level of support	Notes on interpretation
H1: Conceptual framing converges despite definitional variability	Near-universal presence and high overlap of four framing elements: systemic metacognition (44/46), repetitive practice (45/46), adaptive capacity (44/46), and multidomain integration (46/46)	Strongly supported	H1 evidence base = all full-text coded records in <i>H1_dictionary_coding_results_fulltext_46_v2.csv</i> (n = 46). Results indicate convergence at the level of framing vocabulary and conceptual emphasis, not a single formal definition or unified mechanism.

H2: Measurement fragmentation (no de facto standard mode)	Measurement modes are distributed across multiple categories with no dominant approach (psychometric 11; qualitative-only 11; physiological 10; performance/task 6; behavioral/observational 4; composite/index 4); largest mode share = 11/46 = 0.239 (23.9%)	Strongly supported	H2 evidence base = all records in <i>H2_measurement_mode_audit.xlsx</i> (n = 46). Measurement diversity persists even where conceptual framings converge.
H3: Fragmented measurement + strong trainability/adaptation orientation implies need for a convergent metric	Given strong conceptual convergence on adaptive, trainable, performance-oriented mental fitness (H1) alongside persistent measurement fragmentation (H2), the structure of the literature implies a practical need for a convergent mental-fitness metric	Strongly Supported Inference	H3 is evaluated via logical synthesis of H1 and H2 rather than new empirical coding. The “need” is a directional implication from jointly observed field structure, not a claim about the validity of any specific metric.

3.1 H1 Results: Conceptual Convergence

H1 tested whether the mental-fitness literature clusters into a limited set of recurring conceptual framings rather than exhibiting highly idiosyncratic or diffuse definitions. Using frozen, dictionary-based coding rules (Appendix A), four conceptual constructs were assessed across the full-text PDFs of all 46 included studies: Multidomain Integration, Systemic Metacognition, Adaptive Capacity, and Repetitive Practice. Table 4 summarizes the prevalence of each construct across the corpus.

Table 4: H1 Dictionary Coding Results on Zip PDFs (n=46): Binary presence of each H1 construct using Appendix A dictionaries on available full-text PDFs.

Construct	Count	Percent
Multidomain Integration (MD)	46	100.0%
Systemic Metacognition (SM)	44	95.7%
Adaptive Capacity (AC)	44	95.7%
Repetitive Practice (RP)	45	97.8%

All four constructs were highly prevalent. Multidomain Integration was detected in 100% of studies, Systemic Metacognition and Adaptive Capacity were each present in 95.7% of studies, and

Repetitive Practice was present in 97.8% of studies. Only a small number of papers failed to meet the criteria for any given construct, and no construct was confined to a narrow subset of the literature.

To clarify how Multidomain Integration was satisfied, Table 5 reports the qualification pathways under the rule $MD = I \vee (C + E + M + P \geq 2)$.

Table 5. Multidomain Qualification Pathways: How papers qualified for MD under $MD = I \vee (C + E + M + P \geq 2)$ (n=46).

MD trigger pathway	Count	Percent
Explicit multidomain term present (I)	1	2.2%
≥ 2 domains present (C/E/M/P threshold)	46	100.0%
Both I and domain-threshold	1	2.2%
Only I	0	0.0%
Only domain-threshold	45	97.8%

Only one study explicitly used the term *multidomain* or equivalent integration language. All studies qualified via the domain-threshold pathway, meaning they contained references to at least two of the four functional domains (cognitive, emotional, motivational/conative, physical/physiological). The vast majority of papers qualified exclusively through this content-based criterion rather than through explicit self-labeling. Table 6 reports the distribution of domain counts for Multidomain Integration.

Table 6. Domain Count Distribution for MD, Number of detected domains among {Cognitive, Emotional, Motivational/Conative, Physical/Physiological} per paper (n=46).

Domains detected (0–4)	Papers	Percent
2	1	2.2%
3	7	15.2%
4	38	82.6%

Most studies referenced three or four domains, with over 80% containing language from all four domains. Only one study referenced exactly two domains, meeting the minimum threshold for inclusion. Finally, Table 7 shows the degree of overlap across the four H1 constructs.

Table 7. Construct Overlap Across the Four H1 Areas. Number of H1 constructs present per paper (n=46).

# constructs present (0–4)	Papers	Percent
2	1	2.2%
3	3	6.5%
4	42	91.3%

Over 90% of studies exhibited all four constructs simultaneously. Only one study contained fewer than three constructs, and no study fell outside the overall conceptual typology.

Taken together, these results support H1. The mental-fitness literature does not fragment into a large number of unique or incompatible conceptual framings. Instead, it consistently clusters around a small set of recurring constructs that co-occur across most studies. While the specific definitions, methods, and outcomes vary substantially, the underlying conceptual vocabulary shows a high degree of convergence, indicating partial coherence at the level of framing even in the absence of a unified theoretical model.

3.2 H2 Results: Measurement Fragmentation

H2 tested whether the mental-fitness literature converges on a dominant measurement approach or remains fragmented across multiple measurement modes. Using rule-based classification, each of the 46 included studies was assigned a single primary measurement mode based on the approach used to operationalize mental fitness in its principal analyses. Measurement modes were defined *a priori* and applied uniformly across the full-text PDFs, with supporting textual evidence logged for auditability.

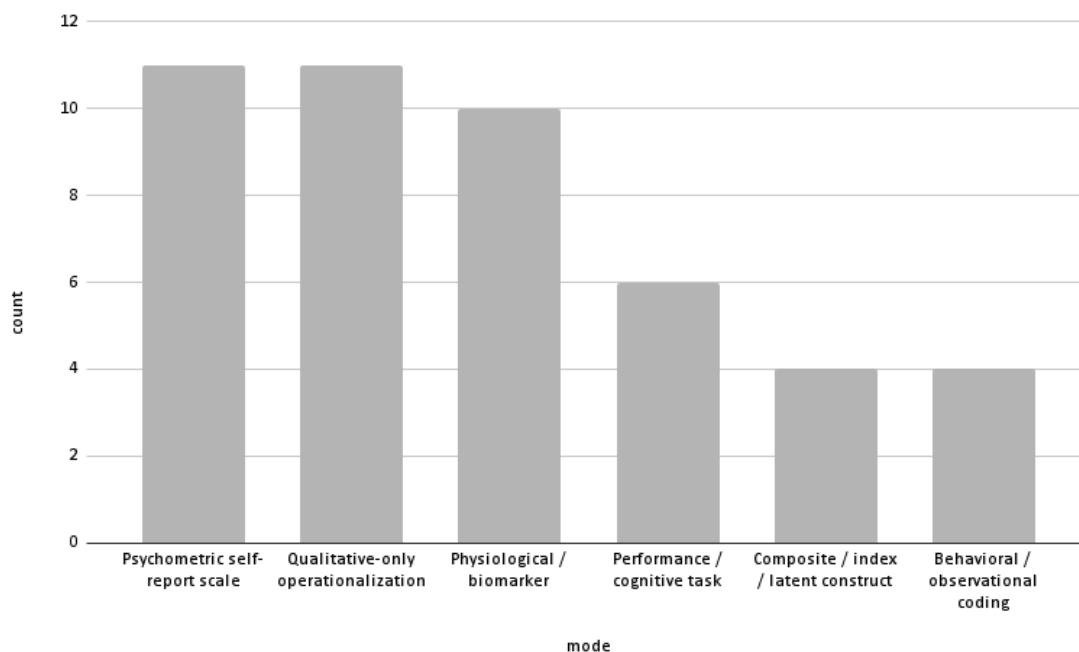
Table 8. H2 Measurement Mode Distribution (n = 46). Primary measurement mode used to operationalize mental fitness in each study.

Measurement Mode	Count	Percent
Psychometric self-report scale	11	23.9%
Qualitative-only operationalization	11	23.9%
Physiological / biomarker proxy	10	21.7%
Performance / cognitive task	6	13.0%
Composite / index-based measure	4	8.7%
Behavioral / observational coding	4	8.7%

Six distinct measurement modes were observed in the corpus. No single mode accounted for more than 23.9% of studies. The two most common modes—psychometric self-report scales and qualitative-only operationalizations—were each used in fewer than one-quarter of the literature. All remaining modes appeared at substantially lower frequencies.

Figure 1 visualizes the distribution of primary measurement modes, illustrating the absence of a dominant approach and the dispersion of measurement strategies across the field.

Figure 1. Distribution of Primary Measurement Modes in the Mental-Fitness Literature (n = 46). Bar chart showing counts of studies by primary measurement mode.



The maximum observed mode share ($11/46 \approx 0.239$) falls well below commonly used thresholds for de facto standardization. Consistent with the coding rules, studies employing multiple measures with equal analytic weight were classified as composite rather than inflating the prevalence of any single mode. This conservative assignment biases the analysis toward detecting convergence; nevertheless, fragmentation persists.

Robustness note (σ sensitivity). The conclusion that measurement remains fragmented is robust to reasonable variation in the dominance threshold σ . Under a strict criterion ($\sigma = 0.50$), no mode approaches majority use. Under a more permissive criterion ($\sigma = 0.33$), intended to capture a strong plurality standard, the largest observed share (0.239) still falls well below threshold. Thus, H2A is supported across the full plausible range of σ values, and H2_o is rejected irrespective of whether de facto standardization is defined conservatively or liberally.

Taken together, these results support H2. The mental-fitness literature does not converge on a single dominant measurement approach. Instead, it exhibits persistent measurement fragmentation across multiple modes. When considered alongside the strong conceptual convergence observed in H1, this pattern indicates a field that shares common framing language but lacks a unified measurement infrastructure.

3.3 H3 Results: Directional Implication from Field Structure

The results of H1 and H2 jointly satisfy the antecedent conditions specified in H3. H1 demonstrated strong conceptual convergence across the mental-fitness literature around adaptive, trainable, and performance-in-context capacities, with recurring emphasis on learning, regulation, metacognition, and transfer rather than static traits or content mastery. H2 demonstrated persistent fragmentation of measurement approaches, with multiple distinct modes in use and no dominant or de facto standard.

Given this joint pattern, the predefined inference rule for H3 applies. When a field frames its focal construct as trainable and performance-based while lacking a convergent measurement infrastructure, cumulative progress, cross-study comparability, and systematic skill tracking are structurally constrained in the absence of a convergent metric.

Accordingly, H3A is supported. The structure of the mental-fitness literature implies a practical need for a convergent mental-fitness metric that is edumetric (sensitive to change with practice), performance- and context-sensitive, and integrative across domains. This conclusion follows directly from the observed field structure and does not depend on the evaluation of any specific instrument. The assessment of candidate metrics that may satisfy these requirements is addressed in the Discussion.

Table 9. Stepwise Logic Underlying H3: From Observed Field Structure to Measurement Implication. This table summarizes the inferential steps linking the empirical results of H1 and H2 to the directional implication tested in H3. Rows represent premises and constraints, not additional hypotheses.)

Step	Source	Observed pattern in the SLR (n=46)	Inference or constraint imposed on measurement
1	H1	Mental fitness is consistently framed as adaptive, contextual, and performance-oriented rather than as static knowledge or abstract trait	A viable metric must measure capability-in-context and performance change, not content mastery or fixed traits
2	H1	Mental fitness is framed as trainable and improvable through practice and feedback	A viable metric must be edumetric and sensitive to learning trajectories and practice effects
3	H1	Mental fitness is treated as multi-domain and integrative, spanning cognition, metacognition, emotion, and regulation	A viable metric must support an integrative architecture with a unified core and interpretable subcomponents

4	H2	Measurement approaches are fragmented; no dominant or standardized mode exists	A convergent metric is required to enable comparability and cumulative measurement
5	H3 inference	Conditions in Steps 1–4 are jointly satisfied	The field structurally implies a need for a convergent mental-fitness metric with the above properties

This inference meets a strong-support criterion because it follows from the joint satisfaction of two independently supported conditions: high conceptual convergence (H1) and persistent measurement non-convergence (H2), a configuration that, under standard measurement and construct-maturity theory, necessitates the development of a convergent metric.

4.0 Discussion

4.1 Interpreting Conceptual Convergence in the Mental-Fitness Literature (H1)

The results of this systematic review reveal a field with a distinctive and nontrivial structural profile. With respect to H1, the mental-fitness literature exhibits *strong conceptual convergence* rather than conceptual dispersion. Across all 46 reviewed studies, a small set of recurring conceptual framing elements—systemic metacognition, repetitive practice, multidomain integration, and adaptive capacity—appears with near-universal frequency and extensive overlap. This pattern directly supports H1: mental fitness is not framed idiosyncratically on a paper-by-paper basis, but instead converges around a limited and recurring conceptual core.

Importantly, this convergence is observed at the level of *conceptual framing and vocabulary*, not at the level of unified theory or shared mechanisms. The H1 analysis was designed to detect the presence and co-occurrence of framing elements, not to adjudicate their depth, coherence, or causal structure. As such, the results indicate that authors across domains are largely talking about the *same kind of capacity*, even when they operationalize or explain it differently.

A notable feature of the H1 findings is the saturation of multidomain language. Although explicit use of the term *multidomain* was rare, all studies referenced multiple functional domains (cognitive, emotional, motivational/conative, and/or physical/physiological). This suggests that multidomain framing has become implicit in the discourse of mental fitness, even when not explicitly labeled as such. In other words, the field is linguistically multidomain by default, reflecting a shared assumption that mental fitness involves coordinated functioning across domains rather than isolated traits or skills.

Taken together, these patterns indicate that mental fitness has achieved a level of *conceptual stabilization* uncommon in purely rhetorical or fad-driven constructs. While definitions vary and

measurement practices remain heterogeneous, the underlying framing elements recur with sufficient regularity to constitute a shared conceptual grammar. This finding distinguishes mental fitness from loosely defined umbrella terms that lack consistent internal structure.

4.1.1 Conceptual Convergence Without Theoretical or Measurement Unification

The strong support for H1 must be interpreted alongside the results of H2. Although the literature converges conceptually, it remains *structurally fragmented* with respect to measurement and operationalization. No single measurement mode dominates the field; psychometric self-report instruments, task-based performance measures, physiological proxies, qualitative assessments, and composite indices all coexist, each representing a minority share of the literature.

This combination—high conceptual convergence paired with low measurement convergence—is characteristic of an *emergent or transitional domain*. Researchers broadly agree on what kind of capacity mental fitness refers to, but lack shared tools for assessing it in a cumulative, comparable way. Conceptual convergence has therefore outpaced methodological consolidation.

From the perspective of H1, this pattern is informative rather than problematic. It suggests that disagreement in the field is no longer primarily about *what mental fitness is*, but about *how it should be measured, trained, and evaluated*. The field appears to have moved beyond definitional proliferation into a phase where conceptual agreement exists without corresponding infrastructural alignment.

4.1.2 Interpreting Conceptual Convergence and Public–Academic Misalignment

The convergence detected in H1 occurs at the level of *conceptual framing* rather than at the level of shared definitions or measurement standards. To further interpret the nature and implications of this convergence, we examined two adjacent questions: (1) whether constructs most commonly associated with “mental fitness” in public discourse are actually present in the scholarly literature, and (2) whether definitions offered in the academic sphere converge on precise, operationally measurable formulations. Together, these analyses reveal a pronounced misalignment between public and academic conceptions of mental fitness, as well as a shared limitation within the literature itself: convergence without formal specification.

As shown in Table 10, widely circulated framings—mindfulness, brain games, and grit—map intuitively onto the four convergent constructs identified in H1 (systemic metacognition, repetitive practice, multidomain integration, and adaptive capacity), but each captures only a partial aspect of a more integrated capacity. These framings are therefore better understood as *entry-point metaphors* rather than as coherent or sufficient models of mental fitness.

Table 10. *Interpretive mapping of common public framings onto the convergent conceptual core identified in H1.*

Mental Fitness Construct (H1)	Common Framing in the Public Sphere	Structural Limitation Observed	Proposed Structural Contribution
Systemic Metacognition	Mindfulness	Emphasizes awareness without specifying mechanisms for restructuring thought or measuring change	Makes metacognitive awareness actionable, improvable, and measurable through explicit structure
Repetitive Practice	Brain games	Produces narrow, task-specific gains with limited far transfer	Shifts emphasis from task proficiency to transferable capability via deliberate practice
Multidomain Integration	Grit	Collapses resilience into persistence, neglecting interactions among domains	Reframes fitness as coordinated interaction across cognitive, emotional, motivational, and physical domains
Adaptive Capacity	All three	Tends toward triviality (games), vagueness (mindfulness), or rigidity (grit)	Clarifies the functional target: flexible adjustment to changing real-world conditions

To assess whether the prominence of these framings in the public sphere is reflected in the scholarly literature, we conducted a literal string audit of the 46-paper corpus. Results are summarized in Table 11.

Table 11. *Literal presence of popular mental-fitness framings in the 46-paper corpus (case-insensitive). Presence indicates literal term use, not conceptual equivalence.*

Term	Papers with term (n = 46)	% of corpus	How the term is used when present
Mindfulness	10	21.7%	Typically appears as a component, module, or intervention ingredient within broader programs or indices; rarely presented as a unifying definition
Grit	1	2.2%	Mentioned in passing as a protective or resilience-related factor; not defined or operationalized as mental fitness
Brain games	1	2.2%	Used as a categorical label for cognitive training approaches; not framed as a defining or explanatory model

This pattern mirrors well-established critiques of these constructs. Brain games demonstrate, at best, short-term and near-transfer effects: training improves performance on the trained task but

rarely generalizes to broader cognitive adaptability or real-world performance [22,23]. Mindfulness programs are widely promoted as panaceas, yet systematic reviews reveal heterogeneous interventions, vague operational definitions, and overstated claims, with only moderate and inconsistent effects on stress and well-being [7,24]. Grit metaphors frame resilience as sheer perseverance, but the construct has been criticized as conceptually vague, heavily overlapping with conscientiousness, difficult to measure reliably, and offering limited incremental predictive power once related traits are controlled [15–17].

Despite these limitations, such partial framings have received disproportionate public attention, producing a landscape heavy on accessibility and appeal but light on causality, reproducibility, and structural coherence. In contrast, their sparse and peripheral appearance in the scholarly mental-fitness literature suggests that the field itself has not converged on these constructs as defining or sufficient explanations. Instead, they are used instrumentally—as components, examples, or adjacent influences—rather than as coherent models of mental fitness.

To examine convergence within the academic sphere itself, we reviewed all explicit definition statements and framework-like claims labeled “mental fitness” (or close substitutes) in the corpus. Table 12 summarizes the results.

Table 12. Definitions and framework claims related to mental fitness in the 46-paper corpus.

Paper ID	Title	Explicit definition?	Definition or description	Model/framework type
MF2	Conceptualising and measuring mental fitness: A Delphi study	Yes	Mental fitness defined as “the modifiable capacity to utilise resources and skills to flexibly adapt to challenges or advantages, enabling thriving”; accompanied by guiding principles	Consensus definition + principle set
CFECA2	Cognitive Enrichment, Self-Regulation, Life Satisfaction and Aging...	Yes	“Mental fitness is a state of mind...” in a goal-setting/empowerment context	Conceptual/aspirational framing
MF3	Internet-Based Intervention to Promote Mental Fitness...	No	Mental fitness instantiated through the Psyfit multi-module intervention and outcome measures	Intervention-as-framework
MF4	Assessing a Structured Mental Fitness Program...	No	Mental fitness treated as an applied wellness/performance capacity via a structured training program	Occupational program model

MF1	Humor as a Mental Fitness Indicator	No	Humor proposed as an indicator of underlying mental fitness	Indicator model
CF4	Crowdsensing: Assessment of Cognitive Fitness Using ML	No	Cognitive fitness inferred from data streams via ML	Measurement framework (adjacent)
CF5	Resiliency... cognitive fitness and confidence...	No	Cognitive fitness modeled as a predictor/moderator of resilience	Domain-specific explanatory model

Taken together, these findings indicate that academic convergence exists primarily at the level of *abstract capacity language* (e.g., adaptability, thriving, state of mind), while explicit definitions remain rare and weakly operationalized. Where definitions are absent, mental fitness is instead defined implicitly through programs, indicators, or domain-specific predictors—effectively by example rather than by shared mechanism. As a result, the academic literature converges quietly but imprecisely, while the public sphere converges loudly around concrete but methodologically fragile substitutes.

4.1.4 Conclusion: Implications of H1

H1 demonstrates that mental fitness is not a conceptually diffuse or arbitrary label. Rather, the literature converges around a small set of recurring framing elements that collectively describe a trainable, adaptive, multidomain capacity expressed through practice and performance. However, this convergence remains largely abstract and weakly specified, lacking shared definitions and measurement infrastructure. At the same time, public discourse has converged on more concrete but structurally incomplete constructs—such as mindfulness, brain games, and grit—that are poorly supported as standalone explanations.

This dual misalignment helps explain why mental fitness remains simultaneously popular and scientifically underdeveloped. Conceptual agreement is now visible, but it has not yet translated into cumulative measurement, theory testing, or reproducible intervention science. The following sections examine how this gap manifests in measurement fragmentation and assess its implications for the future development of the field.

4.2 Interpreting Measurement Fragmentation in the Mental-Fitness Literature (H2)

The results of H2 indicate that, despite strong conceptual convergence, the mental-fitness literature remains fragmented with respect to measurement. Across the 46 reviewed studies, six distinct measurement modes were observed, and no single mode accounted for more than approximately one-quarter of the corpus. This pattern supports H2 and rejects the null hypothesis of measurement

convergence: mental fitness does not yet possess a de facto standard for empirical operationalization.

This fragmentation is not merely a matter of different instruments within a shared paradigm, but of fundamentally different approaches to what it means to “measure” mental fitness. Some studies rely on self-reported traits or states, others on task-based performance, physiological proxies, behavioral observation, qualitative assessment, or composite indices. These modes reflect different epistemic commitments—subjective versus objective, proximal versus distal, trait-oriented versus performance-oriented—rather than incremental methodological variation around a common core.

Importantly, the absence of a dominant measurement mode is robust to reasonable definitions of standardization. Even under permissive thresholds for de facto convergence, no measurement approach approaches plurality dominance. This indicates that fragmentation is structural rather than marginal, and cannot be attributed to outliers or a small subset of atypical studies.

4.2.1 Conceptual–Measurement Mismatch as a Structural Feature of an Emergent Field

When interpreted alongside H1, the H2 findings reveal a pronounced conceptual–measurement mismatch. The field largely agrees on what mental fitness *is like*—a trainable, adaptive, multidomain capacity expressed through practice and performance—yet lacks shared tools for assessing that capacity in a consistent and cumulative manner. Conceptual stabilization has occurred without corresponding measurement stabilization.

This pattern is characteristic of an emergent or transitional scientific domain. In such fields, shared language and framing often precede the development of standardized measurement infrastructure. Researchers recognize the same underlying phenomenon but approach it through the methodological traditions of their home disciplines, resulting in parallel but non-interoperable measurement practices.

Crucially, the H2 results should not be interpreted as evidence of conceptual confusion or theoretical incoherence. On the contrary, measurement fragmentation appears to coexist with—and perhaps follow from—conceptual convergence. Once a construct is broadly agreed upon, disagreement and experimentation tend to shift toward how it should be operationalized, quantified, and evaluated.

4.2.2 Implications of Measurement Fragmentation for Cumulative Science

The absence of measurement convergence has direct implications for synthesis and theory development. Without shared measurement modes, results cannot be easily compared, aggregated, or meta-analyzed across studies. Claims about effectiveness, trainability, or transfer remain tied to specific operationalizations rather than to mental fitness as a unified construct.

As a result, the field risks accumulating parallel literatures that speak about the same underlying capacity but cannot be directly integrated. Conceptual agreement alone is insufficient for cumulative science; it must be paired with some degree of measurement interoperability to enable comparison, replication, and theory testing.

At the same time, the diversity of measurement modes reflects the multidimensional and context-sensitive nature of mental fitness as currently understood. Fragmentation may therefore represent exploratory breadth rather than disorder, particularly in a field that spans psychology, neuroscience, education, health, sport, and performance science. The H2 findings do not imply that any given measurement mode is incorrect, only that none has yet achieved field-wide adoption.

4.2.3 Summary of H2 Implications

In sum, H2 demonstrates that the mental-fitness literature remains measurement-fragmented despite strong conceptual convergence. The field has converged on a shared framing of mental fitness as a trainable, adaptive, multidomain capacity, but has not yet converged on how that capacity should be measured. This divergence constrains cumulative empirical progress and complicates synthesis, even as it reflects active exploration and disciplinary diversity.

Taken together, H1 and H2 describe a field that has stabilized conceptually but not infrastructurally. The next section examines the directional implications of this pattern and assesses whether the structure of the literature itself implies the need for greater measurement convergence, as articulated in H3.

4.3 Directional Implications of Conceptual–Measurement Structure (H3)

H3 addressed whether the joint pattern established by H1 and H2 carries a directional implication for the future development of the mental-fitness field. Specifically, H3 asked whether strong conceptual convergence around mental fitness as a trainable, adaptive, multidomain, performance-in-context capacity—combined with persistent measurement fragmentation—implies a practical need for a convergent measurement framework.

The results of H1 and H2 jointly satisfy the antecedent conditions specified in H3. As shown in Sections 4.1–4.4, the literature converges on a shared conceptual grammar that frames mental fitness as developable through practice, expressed in performance, and distributed across interacting cognitive, metacognitive, emotional, and regulatory processes. As shown in Sections 4.5–4.8, this conceptual convergence is not accompanied by corresponding measurement convergence; instead, the field relies on a heterogeneous set of largely non-interoperable measurement modes, with no dominant standard.

Given this combination, the directional implication specified in H3 follows. When a field defines its focal construct as trainable and performance-based, yet lacks a convergent measurement infrastructure, cumulative progress is structurally constrained. Comparability across studies is limited, claims about improvement remain instrument-bound, and learning trajectories cannot be

tracked or synthesized at the construct level. Under these conditions, the absence of a convergent metric is not merely a methodological inconvenience but a limiting factor for theory testing, replication, and cumulative science.

Accordingly, H3 is supported. The structure of the mental-fitness literature implies a non-trivial need for a convergent measurement approach capable of capturing change with practice, performance in context, and integration across domains. This implication does not entail that existing measures are invalid or that diversity of methods is undesirable. Rather, it suggests that without some form of shared measurement reference, the field risks continued conceptual agreement without proportional empirical consolidation.

Importantly, H3 does not identify or validate any specific measurement instrument. It establishes a set of measurement requirements implied by the field's own commitments: sensitivity to learning and practice, emphasis on performance rather than static traits, and integration across multiple functional domains. Whether and how these requirements might be satisfied is a subsequent question. In the following section, we consider how existing and proposed measurement approaches—including candidate convergent metrics—align with these derived constraints.

5.0 Conclusion

This review assessed the current condition of knowledge surrounding the emerging construct of *mental fitness* by synthesizing the most widely circulated and conceptually relevant literature. Across a corpus of 46 studies, the field shows a consistent structural pattern. No single formal definition, mechanism, or measurement standard dominates the literature. However, there is strong convergence at the level of framing: mental fitness is repeatedly treated as an *adaptive, trainable, performance-oriented capacity* grounded in *metacognitive regulation* and expressed across domains. In contrast, operationalization remains highly fragmented, with no de facto standard mode or convergent metric evident.

Taken together, these findings place the field at an intermediate stage of construct maturity. Conceptual convergence appears sufficient to support synthesis and comparison, yet measurement infrastructure lags behind—constraining benchmarking, cumulative progress, and translation into practice. The core limitation identified here is therefore not an absence of relevant constructs or empirical activity, but an absence of convergent instrumentation capable of capturing the shared structure implied by the literature.

To make this configuration interpretable at a glance, *Figure 2* situates mental fitness within “intelligence in the wild” as a coupled model–world system in which real-time information (*I*) must be organized (*O*) under both external constraints and internal state constraints (C/EMP). In this framing, mental fitness is the trainable capacity to (i) reorganize information as reality changes (with measurable organization, e.g., DSRP) and (ii) regulate state variables so that performance is expressed reliably under pressure rather than only under controlled conditions [25]. This depiction provides a direct bridge between the present review's central pattern—conceptual convergence

alongside measurement fragmentation—and the need for candidate metrics that can quantify *change in adaptive organization over time*.

Figure 2. Intelligence in the wild as adaptive DSRP organization of reality—including internal C/EMP state constraints (adapted from [25]).

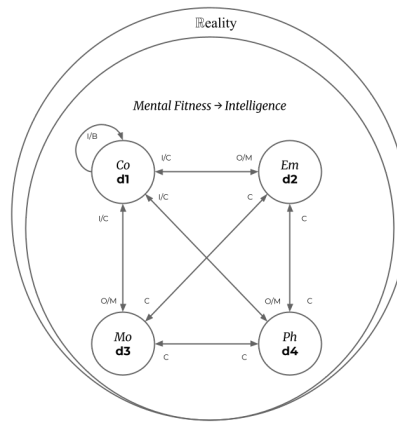
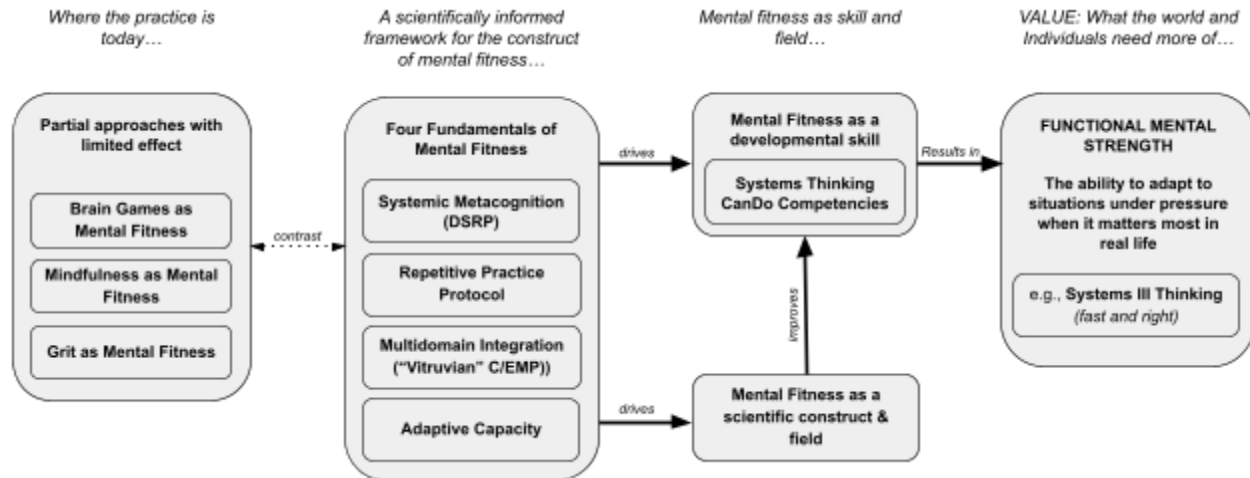


Figure 3 extends this bridge by positioning the review’s results within the broader public–academic landscape. It distinguishes (a) highly salient but structurally partial public framings (mindfulness, brain games, grit), (b) the four convergent fundamentals identified in H1 (systemic metacognition, repetitive practice, multidomain integration, adaptive capacity), and (c) the developmental and scientific implications of treating mental fitness as both a trainable skill and a cumulative field of study. In other words, the figure summarizes the central interpretive claim of the review: the public sphere often collapses mental fitness into accessible but limited substitutes, while the academic sphere converges on an integrated capacity but typically leaves it underspecified and difficult to measure. The functional target implied by the convergent core is therefore not simply “well-being” in general, but *functional mental strength*: the ability to adapt under pressure when it matters most in real life.

Figure 3. From partial public framings to an integrated construct: mapping “mental fitness” from common substitutes (mindfulness, brain games, grit) to the convergent four-fundamental core and the functional target of adaptive performance under pressure.



Recent theoretical work has begun to address this gap at the level of construct clarification. In *From Zoo Smart to Jungle Smart* [25], mental fitness explicitly re-anchors “intelligence in the wild” in adaptivity under changing real-world constraints, formalizing model quality as a multiplicative function of information and organization and treating state variables—cognitive, emotional, motivational, and physical conditions (C/EMP)—as a gating function on performance expression rather than a parallel “intelligence.” In this account, mental fitness is not reducible to IQ/trait performance in standardized conditions (“zoo smart”), but is defined by the trainable capacity to maintain alignment with situational reality in vivo (“jungle smart”) through measurable organization (DSRP), plus state-dependent expression and debt dynamics.

Building on this foundation, *The Thinking Quotient (TQ)* [26] has been proposed as a candidate metric designed to operationalize mental fitness as an organizational and trainable capacity rather than a static trait. The TQ paper positions TQ as an edumetric skill-metric (not a psychometric trait test), grounded in O-Theory/DSRP and the Cognitive Law ($M = I \cdot O$) [27], and reports further validation work focused on reliability and structural coherence (including internal consistency analyses and latent-structure modeling aimed at a dominant general factor alongside DSRP subskills). In practical terms, this represents one concrete attempt to convert the review’s “convergent-but-unspecified” core into measurable change over time.

The present review does not test or validate any specific metric. Rather, it establishes the empirical and conceptual conditions under which candidate metrics become both necessary and evaluable. By clarifying where the field converges, where it fragments, and what that configuration implies—especially in relation to public-facing substitutes—this review shifts the agenda from whether mental fitness is coherent to the more tractable problem of disciplined metric evaluation. Future work can now compare candidate instruments against the constraints derived here, advancing the field toward cumulative measurement and real-world application.

References

1. Terlizzi EP, Zablotsky B. Symptoms of Anxiety and Depression Among Adults: United States, 2019 and 2022. National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC); 2024 Nov. doi:10.15620/cdc/64018
2. Mind Share Partners. 2025 Mental Health at Work Report. 2025. Available: <https://www.mindsharepartners.org/2025-mental-health-at-work-report>
3. Daily Telegraph. U.S. Workforce Engagement and Burnout Report. 2025.
4. DHR Global. New DHR Global Report Reveals Burnout Having Little Impact On Employee Engagement. Nov 2024. Available: <https://www.dhrglobal.com/news/new-dhr-global-report-reveals-burnout-having-little-impact-on-employee-engagement/>
5. Koronka P. Why overwhelmed young workers are taking time off for stress. The Times; 17 Jan 2025. Available: <https://www.thetimes.com/uk/healthcare/article/young-workers-taking-more-time-off-for-stress-gkbjwlh6x>
6. Cabrera D, Cabrera L, Cabrera E. The steps to doing a systems literature review (SLR). *Journal of Systems Thinking*. 2023. doi:10.54120/jost.pr000019.v1
7. Goyal M, Singh S, Sibinga EMS, Others. Meditation programs for psychological stress and well-being: A systematic review and meta-analysis. *JAMA Internal Medicine*. 2014;174: 357–368. doi:10.1001/jamainternmed.2013.13018
8. Fitts PM, Posner MI. *Human Performance*. Belmont, CA: Brooks/Cole; 1967.
9. Bjork RA, Bjork EL. A new theory of disuse and an old theory of stimulus fluctuation. In: Healy A, Kosslyn S, Shiffrin R, editors. *From Learning Processes to Cognitive Processes: Essays in Honor of William K Estes*. Hillsdale, NJ: Erlbaum; 1992. pp. 35–67.
10. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychological Review*. 1993;100: 363–406. doi:10.1037/0033-295X.100.3.363
11. Adkins DL, Boychuk J, Remple MS, Kleim JA. Motor training induces experience-specific patterns of plasticity across motor cortex and spinal cord. *Journal of Applied Physiology*. 2006;101: 1776–1782. doi:10.1152/jappphysiol.00515.2006
12. Bassett DS, Yang M, Wymbs NF, Grafton ST. Learning-induced autonomy of sensorimotor systems. *Nature Neuroscience*. 2015;18: 744–751. doi:10.1038/nn.3993
13. Gross JJ. Emotion Regulation: Current Status and Future Prospects. *Psychological Inquiry*. 2015;26: 1–26. doi:10.1080/1047840X.2014.940781
14. Deci EL, Ryan RM. The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*. 2000;11: 227–268. doi:10.1207/S15327965PLI1104_01

15. Credé M, Tynan MC, Harms PD. Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*. 2017;113: 492–511. doi:10.1037/pspp0000102
16. Ivcevic Z, Brackett MA. Predicting School Success: Comparing Conscientiousness, Grit, and Emotion Regulation Ability. *Journal of Research in Personality*. 2014;52: 29–36. doi:10.1016/j.jrp.2014.06.005
17. Rimfeld K, Kovas Y, Dale PS, Plomin R. True Grit and Genetics: Predicting Academic Achievement From Personality. *Journal of Personality and Social Psychology*. 2016;111: 780–789. doi:10.1037/pspp0000089
18. Steinhall N, McPettit R, Bond J, Parks M, Khan M, Sharfarz D, et al. Wicked Solutions for Wicked Problems: Misalignment in Public Policy. *Journal of Systems Thinking*. 2023; 20.
19. Aljrarri 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–XX. doi:10.54120/jost.00000101
20. Cabrera D, Cabrera L. 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
21. Cabrera D, Cabrera L. The Knowledge-Method Matching Matrix (KMMM): Considering the Condition of Knowledge When Choosing Research Methods. *Journal of Systems Thinking*. 2021;20.
22. Melby-Lervåg M, Hulme C. Is Working Memory Training Effective? A Meta-Analytic Review. *Developmental Psychology*. 2013;49: 270–291. doi:10.1037/a0028228
23. Simons DJ, Boot WR, Charness N, Gathercole SE, Chabris CF, Hambrick DZ, et al. Do “Brain-Training” Programs Work? *Psychol Sci Public Interest*. 2016;17: 103–186. doi:10.1177/1529100616661983
24. Van Dam NT, van Vugt MK, Vago DR, Schmalzl L, Saron CD, Olendzki A, et al. Mind the Hype: A Critical Evaluation and Prescriptive Agenda for Research on Mindfulness and Meditation. *Perspectives on Psychological Science*. 2018;13: 36–61. doi:10.1177/1745691617709589
25. Cabrera D, Cabrera L. From Zoo Smart to Jungle-Smart: Training Mental Fitness for Adapting to Real-World Complexity. *Journal of Systems Thinking*. 2026;6: 1–38. doi:10.54120/jost.v26i1.dc6iqtg6
26. Cabrera D, Das SKVB, Cabrera L. The thinking quotient (TQ): Further validation of a skill-metric for systems thinking. *Journal of Systems Thinking*. 2025;5: 1–18. doi:10.54120/jost.v25i1.dc6tqsm6
27. 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

Appendices

Appendix A

Full Frozen Keyword Dictionaries for H1 Conceptual Typology Coding. Complete, versioned regex dictionaries used for binary coding of H1 constructs. Matching is case-insensitive. All terms are evaluated using OR logic within each trigger set, unless otherwise specified in the construct rule.

Construct	Trigger / Domain	Frozen Keyword / Regex List
Multidomain Integration	Integration cue (I)	\bmultidomain\b ; \bmulti[-\s]domain\b
	Cognitive (C)	\bcognit\w*\b ; \battention\w*\b ; \bworking\s+memory\b ; \bshort[-\s]?term\s+memory\b ; \bexecutive\s+function\w*\b ; \binhibit\w*\b ; \bcognitive\s+control\b ; \bself[-\s]?control\b ; \breason\w*\b ; \bproblem[-\s]?sol\w*\b ; \bdecision[-\s]?mak\w*\b ; \bplanning\b ; \blearning\b ; \bskill\s+acquisition\b ; \bperformance\b ; \bprocessing\s+speed\b ; \breaction\s+time\b ; \bfluid\s+intelligen\w*\b
	Emotional (E)	\bemotion\w*\b ; \baffect\w*\b ; \bmood\b ; \banxiety\b ; \bdepress\w*\b ; \bstress\b ; \bwell[-\s]?being\b ; \bemotional\s+regulat\w*\b ; \bemotion\s+regulat\w*\b ; \bself[-\s]?regulat\w*\b ; \bemotional\s+intelligen\w*\b ; \bE\b
	Motivational / Conative (M)	\bmotiv\w*\b ; \bconat\w*\b ; \bvolition\w*\b ; \bgoal[-\s]?direct\w*
	Physical / Physiological (P)	\bphysical\b ; \bphysiolog\w*\b ; \bbody\b ; \bsomat\w*\b ; \bfitness\b ; \bendurance\b ; \bsleep\b ; \bheart\s*rate\b ; \bHRV\b ; \bcortisol\b ; \bautonomic\b ; \bcardio\w*\b ; \bimmune\b ; \binflamm\w*\b ; \bEEG\b ; \bMRI\b ; \bneuro\w*\b ; \bbrain\b ; \bconnectivit\w*\b ; \bnetwork\s+neuroscien\w*\b
Systemic Metacognition	Systems trigger (S)	\bsystem\w*\b ; \bsystemic\b ; \bsystems?\s+thinking\b

	Metacognition trigger (M)	\bmetacogn\w*\b ; \bmeta[-\s]?cogn\w*\b
Adaptive Capacity	Adaptation (A)	\badapt\w*\b ; \badaptabil\w*\b
	Flexibility (F)	\bflexib\w*\b ; \bagil\w*\b
	Resilience (R)	\bresilien\w*\b ; \bcoping\b ; \bstress\s+toler\w*\b
	Plasticity (P)	\bplastic\w*\b ; \bneuroplastic\w*\b
	Transfer / Generalization (T)	\btransfer\b ; \bgeneraliz\w*\b ; \bnear[-\s]?transfer\b ; \bfar[-\s]?transfer\b ; \bretention\b
Repetitive Practice	Practice (P)	\bpractice\b ; \bpractic\w*\b
	Training (T)	\btrain\w*\b ; \btraining\b
	Repetition (R)	\brepetit\w*\b ; \brepeat\w*\b ; \brehears\w*\b ; \bdrill\w*\b
	Intervention / Protocol (I)	\bintervention\b ; \bprogram(me)?\b ; \bprotocol\b ; \bsessions?\b ; \bregimen\b ; \bcurriculum\b