Beyond Additive Design

An Empirical Taxonomy of Multimodal STEM Accessibility Systems

Madjid Sadallah **Benoît Encelle**
LIRIS, Université Claude Bernard Lyon 1 · CNRS · INSA Lyon - France

CHI EA '26 · Barcelona · April 2026

doi.org/10.1145/3772363 ANR COOBRA

Accessibility ∞ Multimodality Cognitive Load

The Gap — and the Fallacy

The Persistent Gap

- ✗ **Cognitive Overload** despite technological advances
- ✗ Users struggle to form mental models
- ✗ Poor real-world adoption

Root Cause

Not a **sensor** problem — an **orchestration** problem.

How channels are combined matters more than how many.

The Additive Fallacy

More ≠ Better

The widespread assumption that adding sensory channels inherently improves access.

*In practice: modalities are **stacked**, not integrated — leaving users to perform the fusion mentally.*

This Paper

We systematically distinguish **accumulation** from **true perceptual integration** across 66 systems.

Method — Corpus & Scoring Framework

Corpus

66 Systems (2015–2025)

- Math: 42%
- Diagrams: 33%
- Controlled exp.: 58%
- $ICC_{total} = 0.92$

2 independent raters

Disagreements <12%, resolved by consensus.

5-Dimensional Scoring Framework

Dimension	Construct	Max	Wt
Temporal Sync.	Orchestration timing	/4	25%
Semantic Int.	Cross-modal meaning	/4	25%
Topol. Fidelity	Spatial accuracy	/3	18.75%
Modal Approp.	Affordance matching	/3	18.75%
Performance	Measured outcomes	/2	12.5%
Total Structural Score		**/16**	

Sync + Int weighted at 50% to reflect orchestration primacy — both show the highest inter-rater reliability (ICC = 0.91–0.93).

Three Architectural Regimes

ADDITIVE

37.9% avg 4.1/16

Channel stacking.
Users bridge the gap mentally.

Sync = 0.78/4
Int = 1.06/4

AUGMENTATIVE

36.4% avg 8.3/16

Partial coordination.
A necessary stepping stone.

Sync = 2.27/4
Int = 2.67/4

INTEGRATIVE

25.8% avg 12.3/16

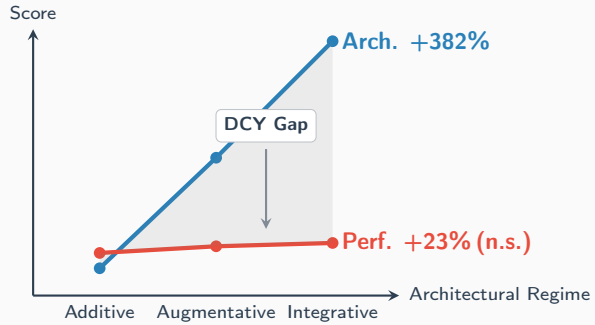
Orchestrated fusion.
Rising post-2019.

Sync = 3.76/4
Int = 3.38/4

Statistical separation: $\eta^2 = 0.83$ Additive $\leftrightarrow$ Integrative: **Cohen's** $d = 6.11$ — qualitatively distinct paradigms, not a continuum.

The Performance–Architecture Paradox

The DCY Phenomenon



Differential Cognitive Yield

Architecturally distinct systems impose **dramatically different cognitive costs** while yielding **similar task performance**.

Why does performance not reflect this?

- 87% avg. success rate — **ceiling effects**
- Only 18% of studies use **NASA-TLX**
- Accuracy $\neq$ cognitive cost

Evidence-Based Guidance

Thresholds to reach Integrative status:

- **Sync $\geq$ 3.2/4** temporal coordination first
- **Int $\geq$ 2.9/4** semantic coherence follows

89% classification accuracy ($\beta_{\text{sync}} = 0.67$, $\beta_{\text{int}} = 0.71$, $R^2 = 0.68$)

3 Design Principles:

1. Sync first — coordination unlocks semantic integration
2. Match affordances: Space $\rightarrow$ Tactile · Time $\rightarrow$ Audio
3. Evaluate with **Perf / Load / Transfer** triad, not accuracy alone

Key Takeaways

Stop Stacking
Modality count is a **vanity metric**.

Measure Load
Accuracy hides exhaustion. Use **NASA-TLX**.

Integrate
Temporal sync is the foundation of **mental models**.

*"The future of accessibility is not more channels, but **better orchestration** of the ones we have."*