First Evidence of Reliability and Validity of SARD
An Exploratory Pilot Study in UAE Primary-Grade Native Arabic Readers

Hanada Taha Thomure
Zai Arabic Research center – Zayed University
Dubai- UAE

Why SARD?

Arabic reading difficulties require early identification.

Few validated Arabic screening tools exist.

SARD = web-based, automated Arabic reading diagnostic with 16 tasks.

Goal: establish initial reliability and validity evidence.

Study Design

- N = 354 native Arabic-speaking students.
- Grades 2–6, UAE schools.
- 16 assessment tasks covering decoding, naming, comprehension, memory and phonological skills.
- Focus of this pilot: reliability and construct validity.

What Does SARD Measure?

Letter reading (isolated and connected forms)

Syllable reading

Real-word and diacritized-word reading

Pseudoword reading

Text reading, comprehension, naming, memory, phonological awareness.

Key Finding 1: Developmental Progression

1. Performance increased consistently from Grade 2 to Grade 6.
2. Letter-level skills stabilized early.
3. Word, pseudoword and text reading continued improving across grades.
4. Supports developmental sensitivity of the battery.

Key Finding 2: Excellent Reliability

1. Cronbach's α = 0.914 for the seven decoding tasks.
2. Item-total correlations ranged from 0.62 to 0.84.
3. Every decoding task contributed meaningfully to the overall construct.
4. Strong evidence of internal consistency.

Key Finding 3: Convergent Validity

1. Decoding tasks correlated with reading comprehension.
2. Grade 2: strongest links for real words, diacritized words, pseudowords and text reading.
3. Grade 6: passage reading ($\rho = .46$) and pseudoword decoding ($\rho = .29$) remained significant.
4. Demonstrates connection between decoding and comprehension.

Key Finding 4: Discriminant Validity

1. After controlling for letter naming accuracy, letter-level effects largely disappeared.
2. Word-level tasks remained significant.
3. Suggests SARD captures higher-order reading processes beyond simple naming ability.

Key Finding 5: PCA Results

1. First principal component explained 66.7% of variance.
2. All decoding tasks loaded positively.
3. Evidence for a strong general reading proficiency factor.
4. Supports a coherent latent construct.

Key Finding 6: Hierarchical Clustering

1

Two clear clusters emerged:

1. Letter-level skills
2. Word/text-level skills

2

Shows SARD distinguishes foundational and advanced reading processes.

Interpretation

- **SARD measures a common reading ability while preserving meaningful subskills.**
- **Captures both foundational orthographic knowledge and lexical processing.**
- **Patterns align with established models of reading development.**

Limitations

- Cross-sectional design.
- No test–retest reliability yet.
- No external diagnostic benchmark.
- Need criterion validity and predictive validity studies.

Implications & Future Directions

Potential large-scale Arabic reading screening tool.

Automated administration and scoring.

Expand to other Arabic-speaking countries.

Evaluate diagnostic accuracy and longitudinal prediction.

Take-Home Message

SARD demonstrated strong reliability and promising validity.

Developmental patterns were theoretically meaningful.

Evidence supports SARD as a screening tool for Arabic reading difficulties.

Next step: large-scale validation.