

Rapid Online Assessment of Reading: Validation of an Italian Single-Word Reading Tool

Elisa Bassoli¹, Giuditta Smith¹, Yagmur Ozturk², Emily Arteaga Garcia³, Wanjing Anya Ma³, ROAR Developer Consortium³, I-ROAR Data Collector Consortium⁴, Jason D. Yeatman³, Marilina Mastrogiuseppe^{2,4}, Sindy Caffarra^{1,3}

¹ University of Modena and Reggio Emilia, Italy

² University of Trieste, Italy

³ Stanford University, USA

⁴ Consiglio Nazionale delle Ricerche, Istituto per la Ricerca e l'Innovazione Biomedica, Italy

Background: In both research and large-scale screening, cognitive abilities such as reading are typically assessed using in-person paper tests, which are expensive and require trained administrators. This results in significant costs both in terms of resources and time. In the context of the literacy crisis reported in Italy and other low and middle income countries [1, 2], there is an urgent need for fast and efficient tools to identify reading difficulties. As far as we know, current Italian options for online reading assessment are often subscription based and time consuming. We propose the I-ROAR-Word as a potential alternative.

The Rapid Online Assessment of Reading is a free web-based platform designed for rapid screening of reading and pre-reading skills in children through short, self-administered tasks. The lexical decision task (LDT) from ROAR has been shown to approximate in-person standardized measures with high accuracy in English and Spanish [3–5]. Here, we present its Italian adaptation, I-ROAR-Word, and examine its psychometric properties, specifically answering the following questions: 1. Is I-ROAR-Word a valid and reliable approximation of a traditional standardized test for Italian single-word reading abilities (Concurrent validity)?; 2. Is it possible to replicate lexical effects already reported in the literature, such as lexicality effect, word frequency and word length effect (Replicability)?; 3. Can this online tool distinguish struggling readers from typical readers (Construct validity)?

Method: A total of 876 children (ages 6–11), recruited from primary schools in northern Italy, completed the I-ROAR-Word, which records accuracy and response times (RTs). Participants also completed standardized measures of Italian reading ability, vocabulary, and non-verbal intelligence [6–8]. The I-ROAR-Word comprises 184 real words, selected in order to capture a broad range of lexical features, and 184 matched pseudowords. Pseudowords were created using Wuggy 0.3.2 [9] and they were matched with real words on length, bigram frequency, orthographic depth and orthographic neighborhood (Figure 1). In order to select the final subset of the stimuli with the highest predictive power ($n=286$), we excluded the items that did not reflect overall reading performance, and applied a Rasch Model to select items with the best combination of difficulty and discriminability. The remaining items were analyzed with mixed linear models including by-subject and by-item random intercepts.

Results: 1. I-ROAR-Word well approximates standardized reading scores and the effect holds even after correcting for age, cognitive measures or vocabulary measures (Accuracy: $\rho = -0.69$, $p < .001$; RT: $\rho = -0.68$, $p < .001$; Figure 2). 2. Word frequency and word length have an effect on I-ROAR-Word performance ($ps < .001$), and real words are read significantly faster than pseudowords ($ps < .001$; Figure 3). These effects are present in all age groups. 3. I-ROAR-Word accuracy differentiates struggling reader children from typical reader children, matched by age and cognitive skills ($t = -6.54$, $p < .001$; Figure 4).

Discussion: The Italian adaptation of the I-ROAR-Word shows a good approximation of a reading standardized test, typically used for screening. The validity of the tool is further confirmed by the replication of lexical effects already reported in the literature (i.e. lexicality, word frequency and word length effect). The I-ROAR-Word can distinguish struggling readers from typical readers. Overall, I-ROAR-Word represents a valid and reliable online tool for the rapid screening of single-word reading difficulties in Italian.

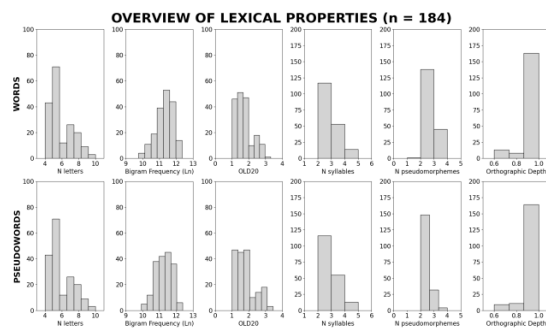


Figure 1: Overview of the lexical properties of the 184 real words and the 184 matching pseudowords.

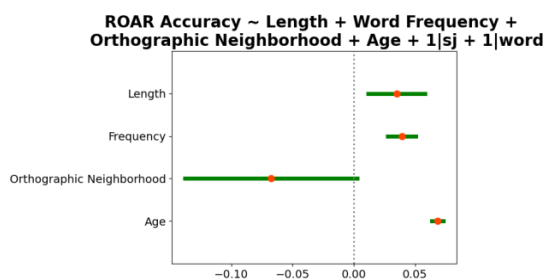


Figure 3: Replicability. Top panel: Confidence intervals (green lines) and beta coefficients (red dot) of the linear mixed model including word frequency and word length effect. Bottom panel: Lexicality effect across all school grades.

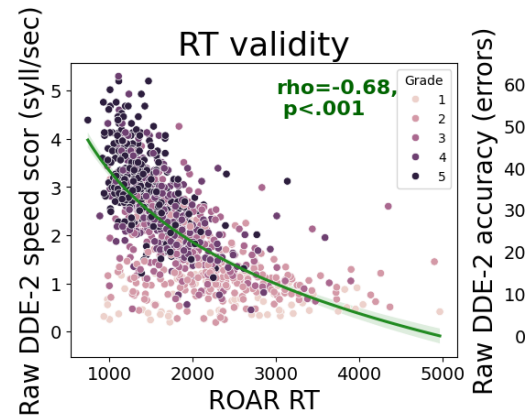


Figure 2: Concurrent validity. Spearman correlation between I- ROAR-Word (response times of correct trials) and standardized reading test (DDE-2).

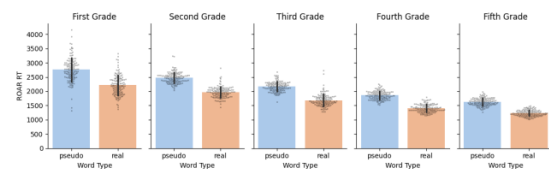


Figure 4: Construct validity. Z-scores of standardized reading tests and I-ROAR-Word from three groups: Struggling readers, Age matched typical readers and Reading matched typical readers.

References

- [1] Contini, D., Tommaso, M. D., Muratori, C., Piazzalunga, D., & Schiavon, L. (2022). Who Lost the Most? Mathematics Achievement during the COVID-19 Pandemic. *The B.E. Journal of Economic Analysis & Policy*, 22(2), 399–408. <https://doi.org/10.1515/bejeap-2021-0447>
- [2] Mundial, B. (2022). The state of global learning poverty: 2022 Update.
- [3] Yeatman, J., Tang, K., Donnelly, P., Yablonski, M., Ramamurthy, M., Karipidis, I., Caffarra, S., Takada, M., Kanopka, K., Ben-Shachar, M., & Domingue, B. (2021). Rapid online assessment of reading ability. *Scientific Reports*, 11(1), 6396. <https://doi.org/10.1038/s41598-021-85907-x>
- [4] Ma, W., Richie-Halford, A., Burkhardt, A., Kanopka, K., Chou, C., Domingue, B., & Yeatman, J. (2023). ROAR-CAT: Rapid Online Assessment of Reading ability with Computerized Adaptive Testing.
- [5] Siebert, J., & Fuentes-Jimenez, M. (2024). A Fair Lexical Decision Task for Monolingual and Multilingual Spanish-speakers. *PsyArXiv*. <https://doi.org/10.31234/osf.io/qfddpb>
- [6] Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test*. American Psychological Association (APA). <https://doi.org/10.1037/t27208-000>
- [7] Raven, J., Raven, J., Court, & Jh. (1998). Section 2: Coloured Progressive Matrices (1998 edition). Introducing the parallel version of the test. Manual for the Raven's Progressive Matrices and Vocabulary Scales.
- [8] Sartori, J., & Tressoldi. (2007). Batteria per la Valutazione della Dislessia e della Disortografia.
- [9] Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior research methods*, 42(3), 627–633. <https://doi.org/10.3758/brm.42.3.627>