

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/361842183>

Development and Validation of a Serious–Game Based Literacy Assessment

Poster · July 2022

CITATIONS
0

3 authors:



Laurène Houtin

AD-HOC Lab

44 PUBLICATIONS

64 CITATIONS

SEE PROFILE



Clémence Brun

AD-HOC Lab

42 PUBLICATIONS

64 CITATIONS

SEE PROFILE

READS
12



Alexis Akinyemi

AD-HOC Lab

47 PUBLICATIONS

67 CITATIONS

SEE PROFILE

Development and Validation of a Serious-Game Based Literacy Assessment

Laurène Houtin^{1,2}, Alexis Akingemi^{1,2}, & Clémence Brun^{1,3}, contact: laurene.houtin@adhoclab.fr

¹AD-HOC Lab, ²Univ. Paris Nanterre - Laboratoire Parisien de Psychologie Sociale, ³Univ. Grenoble Alpes, TIMC-IMAG UMR CNRS 5525, ThEMAS team



Illiteracy is one of the main concerns of European policymakers on the topic of education equal opportunities

promotion (OECD, 2021). The Programme for the International Assessment of Adult Competencies (PIAAC) is specially designed to assess literacy worldwide (*i.e.*, 24 countries involved). It provides - every 10 years - insights regarding the level of concern that participants' country policymakers should have regarding the evolution of literacy level in their territories, considering the evolution of other countries levels.

Moreover, the use of serious games to assess the acquisition of these basic skills seems to be a promising assessment method (Serret et al., 2017).

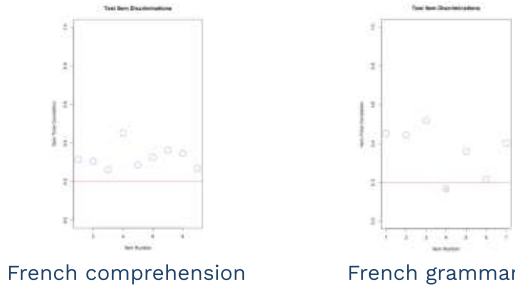
The project **EVA** aimed at providing French career guidance advisors and social workers with a free public service based on a **playful digital tool to assess literacy**, while taking into account their professional everyday constraints, to act quickly enough to prevent the social and professional consequences of illiteracy.

It was designed for young unemployed, and the skills we aimed to assess were: reading, vocabulary, syntax, grammar, punctuation, spelling, written and oral comprehension in French.

LDT : after conventional outlier deletion and score normalization (Niedenthal & Setterlund, 1994), we conducted a factor analysis (Barlett $\chi^2 = 41983$, $dl = 1596$, $KMO = .99$) which confirmed a bi-factorial structure separating words from non-word.

MCQs : factor analyses ran on MCQs binary responses showed that items were distributed according to theory. Although, item discrimination and item difficulty analyses showed that some items' response modalities should be revised to improve the discrimination of participants' levels.

Item discrimination analyses :



Global scores calibration : we computed global scores calibration according to factor analyses outputs. While scores using LDT were close to a normal distribution, MCQs were strongly heavy-tailed to the lower scores. Considering the potential of response time, we computed new global scores using composite scores of response time and response correctness for MCQs which led to a global score distribution suitable for calibration. We used standardization, and n-SD distances from the mean to determine global score clusters that were used for participants' debriefing.

INTRODUCTION

RESULTS

METHOD

DISCUSSION

A digital serious game of 3 modules was developed. The whole playing experience was designed within a common storyline (*i.e.*, an amusement park tour) to preserve users' commitment during the average 30 minutes of the test. The first task was a **lexical decision task** (LDT) designed according to the storyline while carefully respecting the experimental settings of literature (e.g., Harrington, 2006; Niedenthal & Setterlund, 1994). The second and third modules are structurally similar to the PIAAC's ones, commonly divided into several **multiple-choice questions** (MCQs) dedicated to the evaluation of the literacy skills.

Participants were 1077 French young unemployed (*Minage* = 16, *Maxage* = 25). They were given a computer running the tests and headphones and they didn't receive any help or guidance except for the modules introductions. They were then thoroughly debriefed by their career counselors, with a complete description of the test output (*i.e.*, level identified and guidance advice).



LDT was more stable than MCQs even with a graphical and contextual adaptation. This result led us to consider using response time together with response correctness to improve

MCQs accuracy. These results advocate for the interest of response time-based questions in language assessment, particularly for computer-based assessment.

They also showed that the adaptation of the tests in a playful context was not associated with a drop in test accuracy in LDT even though an experimental study dedicated to this question should be conducted.

Finally, providing the improvement of item difficulty indices - with the adjustment of some response modalities - results showed that **this digital serious game was appropriate to assess the literacy level of young unemployed within 30min**, which has positive applied consequences considering the difficulties encountered by professionals in the field.



<https://eva.beta.gouv.fr>

References :

- Harrington, M. (2006). The lexical decision task as a measure of L2 lexical proficiency. *EUROSLA yearbook*, 6(1), 147-168.
- Niedenthal, P. M., & Setterlund, M. B. (1994). Emotion congruence in perception. *Personality and Social Psychology Bulletin*, 20(4), 401-411.
- Rouet, J. F., Britt, M. A., Gabrielsen, E., Kaakinen, J., Richter, T., & Lennon, M. (2021). *PIAAC Cycle 2 assessment framework: Literacy*.
- Serret, S., Hun, S., Thümmel, S., Pierron, P., Santos, A., Bourgeois, J., & Askenazy, F. (2017). Teaching literacy skills to French minimally verbal school-aged children with autism spectrum disorders with the serious game SEMA-TIC: An exploratory study. *Frontiers in psychology*, 8, 1523.

