KU LEUVEN

The Impact of AI-Supported Learning on Financial Literacy: Evidence from a Randomized Controlled Trial

Jaime Polanco-Jimenez & Kristof De Witte

KU Leuven

April 3, 2025

This Paper

Question: Can AI, particularly tailored AI chatbots, effectively improve financial literacy outcomes and potentially bridge learning gaps in the Belgian context?
Motivation: Growing importance of financial literacy, coupled with persistent learning gaps. Exploring the potential of scalable AI chatbot solutions (comparing generic vs. tailored approaches) within the specific Belgian educational system.
How:

- RCT identifying the causal effect of AI chatbots vs. traditional learning path.
  - **Compared:** Control (Traditional Path) vs. T1 (Reduced TP + AI) vs. T2 (Tailored AI).
  - **Main Outcome:** ↑ Increase in Financial Literacy Scores 10pp in Std. for both AI groups relative to Control.
- Data from N=2236 students in Belgium that have started the learning path.

Heterogeneities:

- Heterogeneity by prior knowledge and student characteristics explored.

Justification for Financial Education in Belgium

- **Low Basic Financial Literacy:** Significant portion of Flemish students struggle with basic financial decisions. [1]
- **Socio-Economic Gap:** Strong correlation between socio-economic status and financial literacy scores highlights inequality. [2]
- **Teacher Training Imperative:** Many teachers lack sufficient financial literacy skills to effectively teach the subject. [3]
- **Curriculum Development:** Continuous curriculum development is needed to address evolving financial challenges. [4]

[1] OECD (2017). PISA 2015 results: Students' financial literacy. Paris: OECD Publishing.

[2] De Beckker, K., De Witte, K., & Van Campenhout, G. (2019a). Identifying financially illiterate groups: An international comparison. International Journal of Consumer Studies, 43(5), 490–501.

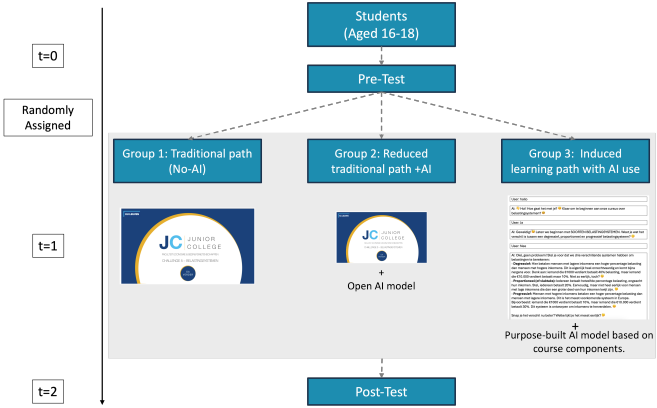
[3] De Beckker, K., Compen, B., De Bock, D., & Schelfhout, W. (2019). The capabilities of secondary school teachers to provide financial education. Citizenship, Social and Economics Education, 18(2), 66-81.

[4] European Communities. (2007). Key competences for lifelong learning. European reference framework. Retrieved from http://hdl.voced.edu.au/10707/285153.

Contribution

- Delivers causal evidence from a RCT in Belgian secondary education on the effectiveness of AI chatbots in improving financial literacy scores compared to traditional instruction.
- Uniquely assesses both cognitive outcomes (financial literacy) and affective/experiential factors (e.g., motivation, user experience) related to AI tool adoption in education.
  - Different to Gregory Kestin (WP. 2024) and Owen Henkel et al. (WP. 2025) that focus on learning in mathematics.
- Provides novel insights into the application of AI (comparing generic vs. tailored chatbots) within financial literacy, informing AI tool design for educational contexts.

Study Design: Randomized Controlled Trial (RCT)



The sample size of 732 was determined by a power analysis ($\alpha = 0.05$, power $= 0.80$), assuming a medium effect size (Cohen's d $= 0.2$) for the primary outcome: the post-test score difference between the combined AI groups and the control group.

Randomization Check: Baseline Score Balance

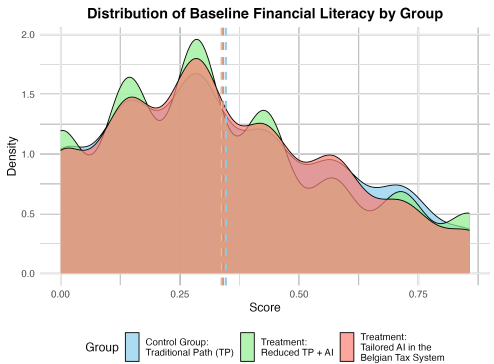


Figure: Kernel density plot of baseline financial literacy scores for Control (likely Blue), Treatment 1 (Reduced TP+AI, likely Green), and Treatment 2 (Tailored AI, likely Red). Distributions appear very similar.

Variable	Baseline Score
Control Group Mean (Constant)	0.347 (0.009)***
Treatment: Reduced TP + AI	-0.009 (0.013)
Treatment: Tailored AI	-0.006 (0.013)
Observations	2,236

Notes: Baseline balance check. Standard errors in parentheses. N=2,236. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Ref: Control Group.

Interpretation:

- No significant difference in baseline score between treatment arms and control (coefficients are small, p>0.4).
- Supports successful randomization. Groups are comparable at baseline.

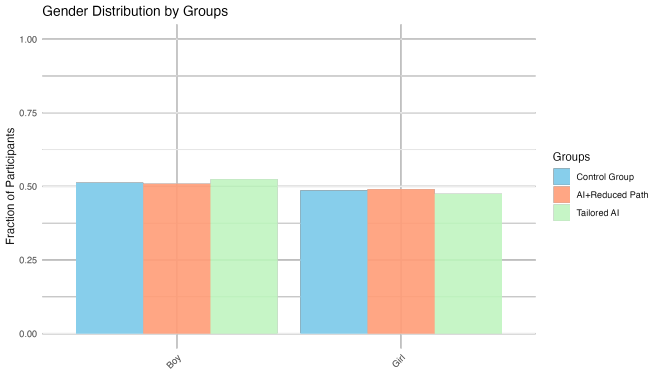
Baseline Outcomes Balance

Table: Descriptive Statistics by Group (Baseline)

Variable	Traditional Path(TP)	Reduced TP+ AI	Tailored AI(Belgian Tax)
Overall Score	0.347 (0.245)	0.338 (0.246)	0.341 (0.239)
Attitude and Motivation	2.861 (0.680)	2.882 (0.679)	2.839 (0.651)
Learning & User Experience	2.783 (0.888)	2.783 (0.884)	2.743 (0.826)
Self-Regulation & Metacognition	2.723 (0.813)	2.649 (0.787)	2.661 (0.780)
Engagement & Commitment	2.544 (0.767)	2.490 (0.704)	2.500 (0.770)
Self-Confidence & Self-Efficacy	2.685 (0.880)	2.689 (0.836)	2.698 (0.863)
Emotional Factors	2.466 (0.803)	2.456 (0.768)	2.488 (0.812)

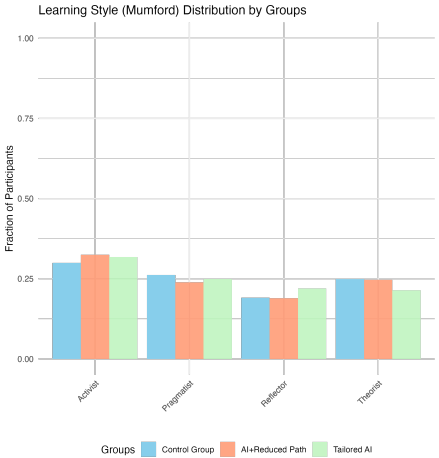
Note: Values are Mean (Standard Deviation). Scales are 1-5 Likert unless noted for Overall Score. *Variable descriptions: Attitude and Motivation:* Interest in learning taxes; Perceived usefulness; Motivation from lesson. *Learning & User Experience:* Usefulness for understanding; Ease of use/intuitiveness; Satisfaction with experience. *Self-Regulation & Metacognition:* Monitoring understanding; Strategies for difficulties; Connecting learning to life. *Engagement & Commitment:* Enthusiasm and energy; Absorption/concentration; Perceived amount learned. *Self-Confidence & Self-Efficacy:* Confidence solving problems; Belief in practical application. *Emotional Factors:* Anticipated test anxiety; Performance expectations.

Baseline Characteristics: Gender Balance



- Gender distribution compared across groups at baseline.
- *Test:* Pearson's $\chi^2(2) = 0.74$, $p = 0.946$.
- **Conclusion:** Groups well-balanced on gender.

Baseline Characteristics: Learning Style (Murford) Balance



- Learning Style distribution compared across groups at baseline.
- *Test:* Pearson's $\chi^2(2) = 5.36$, p $= 0.49$.

Honey and Mumford Learning Style:

- Activist: Learns best by doing and experiencing.
- Reflector: Learns best by observing and thinking things through.
- Theorist: Learns best by understanding concepts, and the logic behind things.
- Pragmatist: Learns best when they see a practical application

Baseline Balance: Categorical Variables

Table: Baseline Balance Check for Categorical Characteristics

Variable	Test Statistic (χ^2(df))	p-value
Gender	$\chi^2(4) = 0.74$	0.946
Type of School	$\chi^2(8) = 10.37$	0.240
Last Dutch Grade (Previous School Year)	$\chi^2(8) = 5.00$	0.757
Last Math Grade (Previous School Year)	$\chi^2(8) = 7.63$	0.471
Predominant Language Used at Home	$\chi^2(4) = 0.73$	0.948
Learning Style (Mumford)	$\chi^2(6) = 5.36$	0.499

Note: All p-values are substantially greater than conventional significance levels (e.g., $\alpha = 0.05$). No statistically significant imbalances were detected. The groups appear well-balanced for these characteristics at baseline.

Main Result - Gained Learning

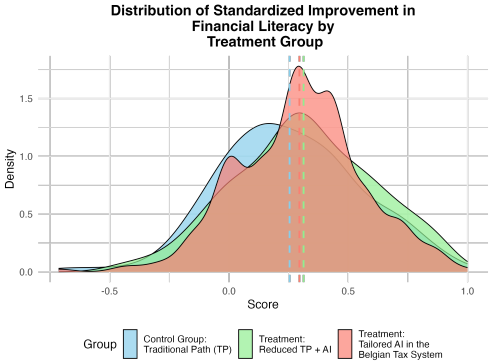


Figure: Kernel density plot of gained learning observed in financial literacy scores

Variable	Baseline Score
Control Group Mean (Constant)	0.254 (0.021)***
Treatment: Reduced TP + AI	0.067 (0.030)*
Treatment: Tailored AI	0.043 (0.025)·
Observations	640

Notes: Baseline balance check. Standard errors in parentheses. N=640 (87% of expected participants by power analysis). $^{***}p = 0$, $^{**}p < 0.001$, $^{*}p < 0.05$, $^{*}p < 0.1$. Ref: Control Group.

Main Result - Financial Literacy Learning

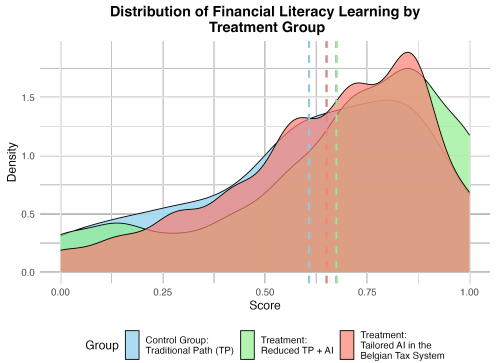


Figure: Kernel density plot of learning observed in financial literacy scores

Variable	Baseline Score
Control Group Mean (Constant)	0.608 (0.020)***
Treatment: Reduced TP + AI	0.067 (0.030)*
Treatment: Tailored AI	0.043 (0.025)·
Observations	640

Notes: Baseline balance check. Standard errors in parentheses. N=640 (87% of expected participants by power analysis). $^{***}p = 0$, $^{**}p < 0.001$, $^{*}p < 0.05$, $^{*}p < 0.1$. Ref: Control Group.

Power Significance Sensitivity:

Effect Size	0.10	0.15	0.20	0.25	0.30	0.35
Power						
0.80	2901.0	1296.0	732.0	474.0	333.0	246.0
0.85	3288.0	1467.0	831.0	534.0	375.0	279.0
0.90	3807.0	1698.0	960.0	618.0	432.0	321.0
0.95	4644.0	2070.0	1170.0	753.0	525.0	390.0

Motivation, Self-Regulation, and Engagement Results

Table: Gained Motivation, Self-Regulation, and Engagement Results

Variable	Constant	Reduced TP+ AI	Tailored AI(Belgian Tax)
Attitude and Motivation	-0.1724 (0.680)***	0.04422 (0.09774)	-0.13610 (0.08398)
Learning & User Experience	-0.2903 (0.07391)***	0.05812 (0.11140)	0.0500 (0.09557)
Self-Regulation & Metacognition	-0.0113 (0.0600)	-0.0596 (0.0902)	0.1271 (0.0776)
Engagement & Commitment	0.13706 (0.06835)	0.00242 (0.10297)	0.06976 (0.08854)
Self-Confidence & Self-Efficacy	-0.24200 (0.07650)**	-0.22372 (0.11488)·	0.20803 (0.09879)*

Notes: OLS Regression. Dependent Variable: Change from pre-test to post-test (Gained Points, Scaled 1-5). Standard errors in parentheses. N=640.
Statistical Significance:
$^{***}p = 0$, $^{**}p < 0.001$, $^{*}p < 0.05$, $^{*}p < 0.1$. Ref: Control Group.

Conclusions

Main Finding: Participation in AI-supported financial literacy interventions led to statistically significant gains in financial literacy scores compared to traditional methods.

- Standardized scores increased by approximately 0.07 to 0.10 standard deviations on average, which is relevant despite only a moderate effect

The Impact of AI-Supported Learning on Financial Literacy: Evidence from a Randomized Controlled Trial[5]

PhD Student
KU Leuven
LEER
Pontificia Universidad Javeriana
Thank you!

https:/polanco-jaime.github.io/
LEER Conference 2025

[5]Corresponding author: Jaime Polanco-Jimenez. Leuven Economics of Education Research, KU Leuven, Belgium; Department of Economics at Pontificia Universidad Javeriana, Colombia
Kristof De Witte: Full Leuven Economics of Education Research, KU Leuven, Belgium; UNU-Merit, Maastricht University, the Netherlands.

KU LEUVEN