

FutureFLO in Nursing Education Grant Evaluation

Initial experiments in generative chatbot development

Regina Russell, PhD, MA, MEd, VUSN

Md. Mashiur Rahman Khan, Vanderbilt University

Mike Wang, MEd, UNC at Wilmington

Marianna LaNoue, PhD, MS, VUSN

Hannah McGrew, MSN, RN, APRN, CNM, FNP-C, VUSN

Samantha Swift, VUSN

Patty Sengstack, DNP, RN-BC, FAAN, FACMI, VUSN

Julia Phillippi, PhD, RN, CNM, FACNM, FAAN, VUSN

Intro

- Advancements in artificial intelligence (AI), especially large language models (LLMs), offer promising enhancements to program evaluation in nursing education.
- Effective evaluation is crucial for improving educational outcomes and securing grant funding in nursing education.
- We investigate the use of a generative chatbot, Future-FLO, designed to support evaluation processes.
- This study employed a mixed-methods approach to evaluate the Future-FLO system and its components: the FutureFLO persona definition, the Improver-Bot, and task-specific prompts.

Methods

- Utilizing workplan data from the Maternity Care Nursing Workforce Expansion program, we conducted a pilot test to generate and assess Gantt charts from data in static .pdf files.
- ChatGPT 4.0 outputs were assessed by both human evaluators and an automated Improver-Bot for accuracy, completeness, and usefulness

Acknowledgements

- VUSN MEADOW Team (HRSA MATCARE Grant 23-120)
- VU Chancellor's Generative AI Seed Grant
- VU Amplify System Team

VUSN Research Day

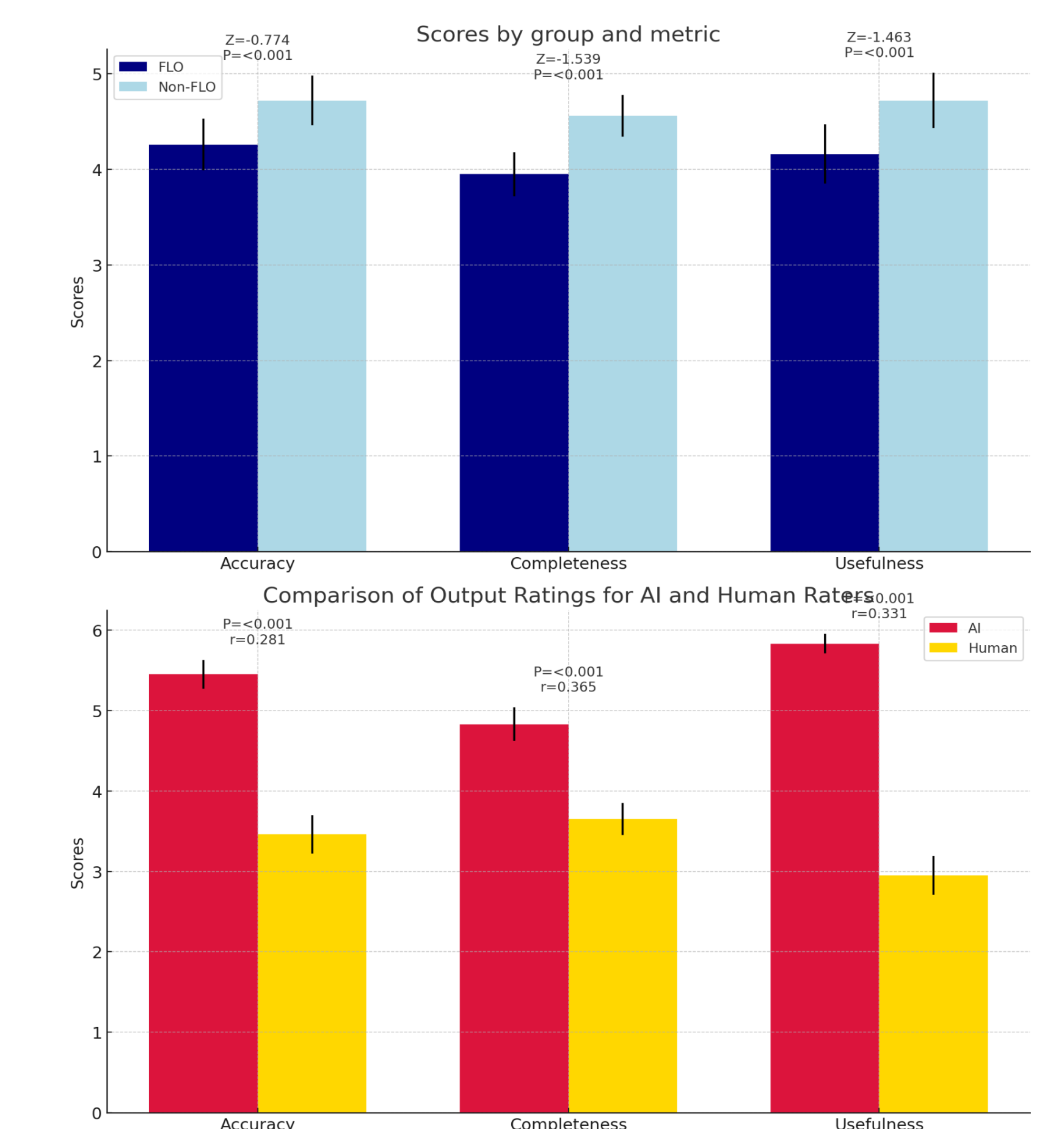
- August 21, 2024
- Center for Research Development and Scholarship
- Vanderbilt Center for Research on Inequality and Health

The evaluation revealed that the FutureFLO system did not improve the outputs as expected. The integration of AI into evaluation must be approached cautiously, ensuring alignment with practical needs and ethical standards.



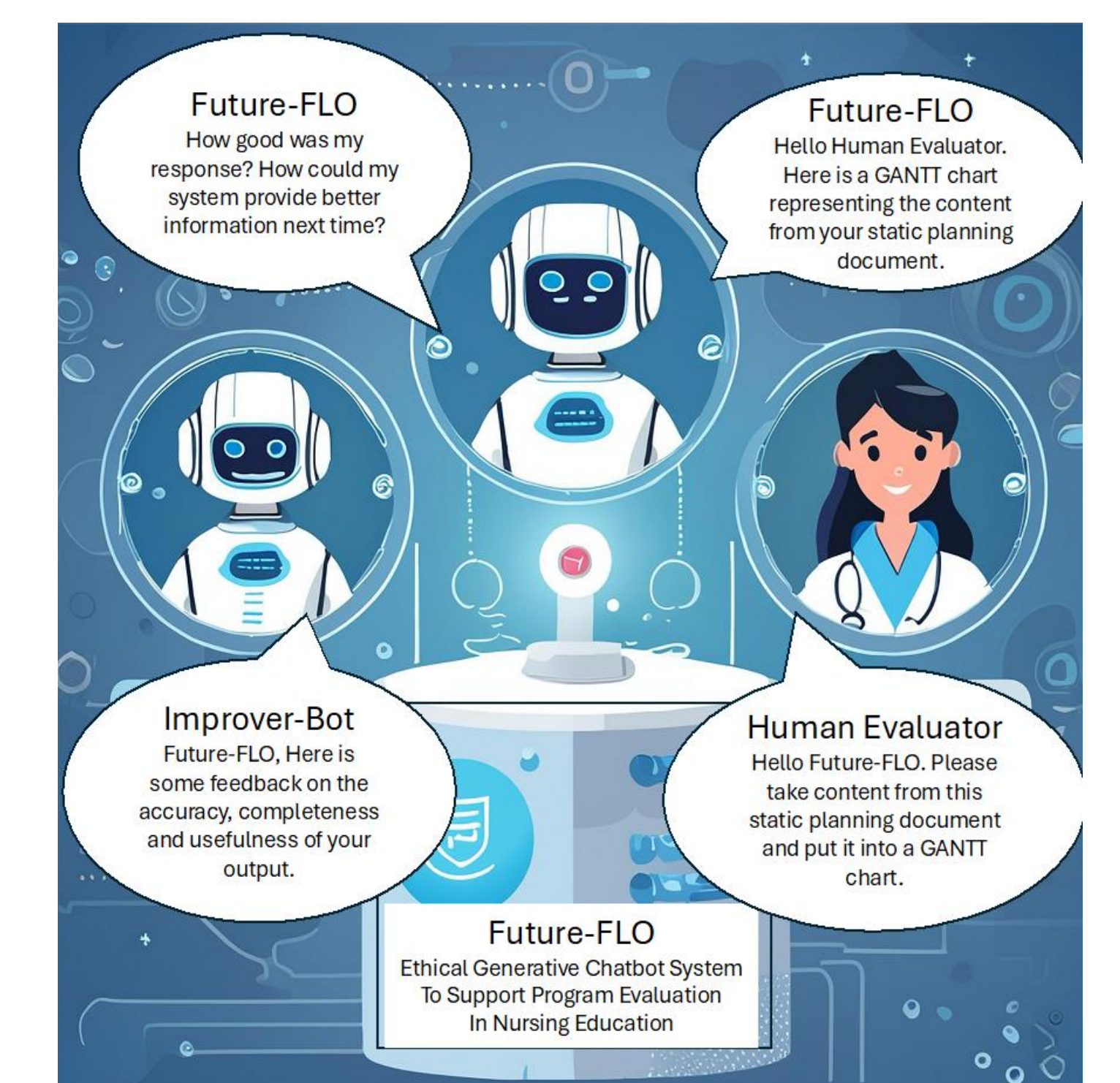
Results

- 8 outputs rated from ChatGPT
 - 4 w/FLO persona, 4 without
- 6 Human Raters and 37 ratings
- 1 Improver-Bot and 40 ratings
- Contrary to expectations, Future-FLO did not significantly improve the evaluation outputs.
- Outputs rated by the Improver-Bot showed potential biases in assessment.



Conclusions

- The results highlight the challenges of implementing generative AI in evaluation tasks.
- Indicating the need for further development and careful integration to align with practical and ethical standards.



Building FutureFLO: A generative chatbot system to support program evaluation in nursing education