Developmental Appropriateness of AI-generated Wimmelbooks

Zuzana Fedorková

Comenius University Bratislava
fedorkova24@uniba.sk

1 Introduction

Generative AI is increasingly integrated into children's literature. It raises questions about whether AI-generated content can match the developmental quality of human-authored material, particularly in complex genres, such as wimmelbooks.

Wimmelbooks (from German Wimmelbuch, "wimmeln" = "to teem") are wordless picturebooks characterised by rich, large-format, complex illustrations "teeming" with characters and details Rémi (2017). Across each double page, the characters "narrate" many different mini-stories that the child can construct and follow. Wimmelbooks primarily depict everyday situations and are thus a rich source for seeking causal explanations for things happening daily, which children usually search for between 4 and 7 years Linka et al. (2023).

This work presents a research concept with preliminary results from exploratory testing. Using wimmelbook genre, it seeks to understand whether AI-generated content meets the developmental needs of preschool children and how prompting methods affect genre accuracy and final result.

2 Methods

Drawing from the research in developmental visual attention Linka et al. (2023) and scene grammar Henderson and Hollingworth (1999), this work proposes an analytical framework for wimmelbooks and applies it to AI-generated examples. For exploratory testing, one distant scene type was generated using the GPT Image 2 model and three prompting strategies (minimal, detailed, and developmentally explicit), which were then qualitatively analysed. Going forward, more scene types (closer, distant) and a mixed quantitative-qualitative approach is planned to examine how AI-generated illustrations conflict with the developmental needs of preschool children and whether their quality differs across scene types and prompting strategies.

3 Results

The preliminary exploratory results show partial success: the GPT-image 2 model reproduced general genre features, such as many characters engaged in varied simultaneous actions, but struggled with planar distribution and scene grammar. It also failed at detail generation, particularly hands - the primary attentional target for preschool children Linka et al. (2023). While different prompting strategies suggest a tension between developmental appropriateness and genre adherence, the effects of scene type and prompting strategy on developmental appropriateness still require further analysis.

4 Discussion

The present work looks at the appropriateness of AI use in children's literature. Preliminary exploratory results indicate that GPT Image 2 model reproduces surface genre features but struggles to fully deliver the developmental qualities of more complex genres. In order to draw stronger conclusions, conducting a systematic study with a larger image set is needed. The findings might be of interest to illustrators, AI designers and people with an interest in the use of AI for children's content.

References

Henderson, J. M. and Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*.

Linka, M., Sensoy, O., Karimpur, H., Schwarzer, G. and de Haas, B. (2023). Free viewing biases for complex scenes in preschoolers and adults. *Scientific Reports*, 13(1):11803.

Rémi, C. (2017). Wimmelbooks. In *The Routledge Companion to Picturebooks*. Routledge.