EMORY

NELL HODGSON WOODRUFF SCHOOL OF NURSING

Evaluation of Large Language Models in Tailoring Educational Content for Underserved Cancer Survivors and Their Caregivers

Darren Liu[1,2], Xiao Hu[1,2,3], Canhua Xiao[1,4], Jinbing Bai[1,4], Zahra A. Barandouzi[1,4], Stephanie M. Lee[1], Caitlin I. Webster[1], La-Urshalar Brock[1,4], Lindsay J. Lee[5], Delgersuren Bold[1,2], Yufen Lin[1,4]

[1]Nell Hodgson Woodruff School of Nursing, Emory University [2]Center for Data Science, Emory University [3]The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology [4]Winship Cancer Institute, Emory University, [5]Department of Medicine, University of Florida

INTRODUCTION

- Underserved cancer survivors and their caregivers face a disproportionately increased risk of **symptom burden** from cancer and its treatments.
- **Large language models (LLMs)** offer researchers an opportunity to develop educational materials tailored to these populations.
- This study aimed to **evaluate different LLMs in tailoring educational content for underserved cancer survivors and their caregivers**.

METHODS

Data sources:

- Education materials are from the **national guidelines** (e.g., National Cancer Institute, National Comprehensive Cancer Network).
- We selected **30 distinct topics** such as fatigue, depression, anxiety, physical activity, coping skills, and more.

Models:

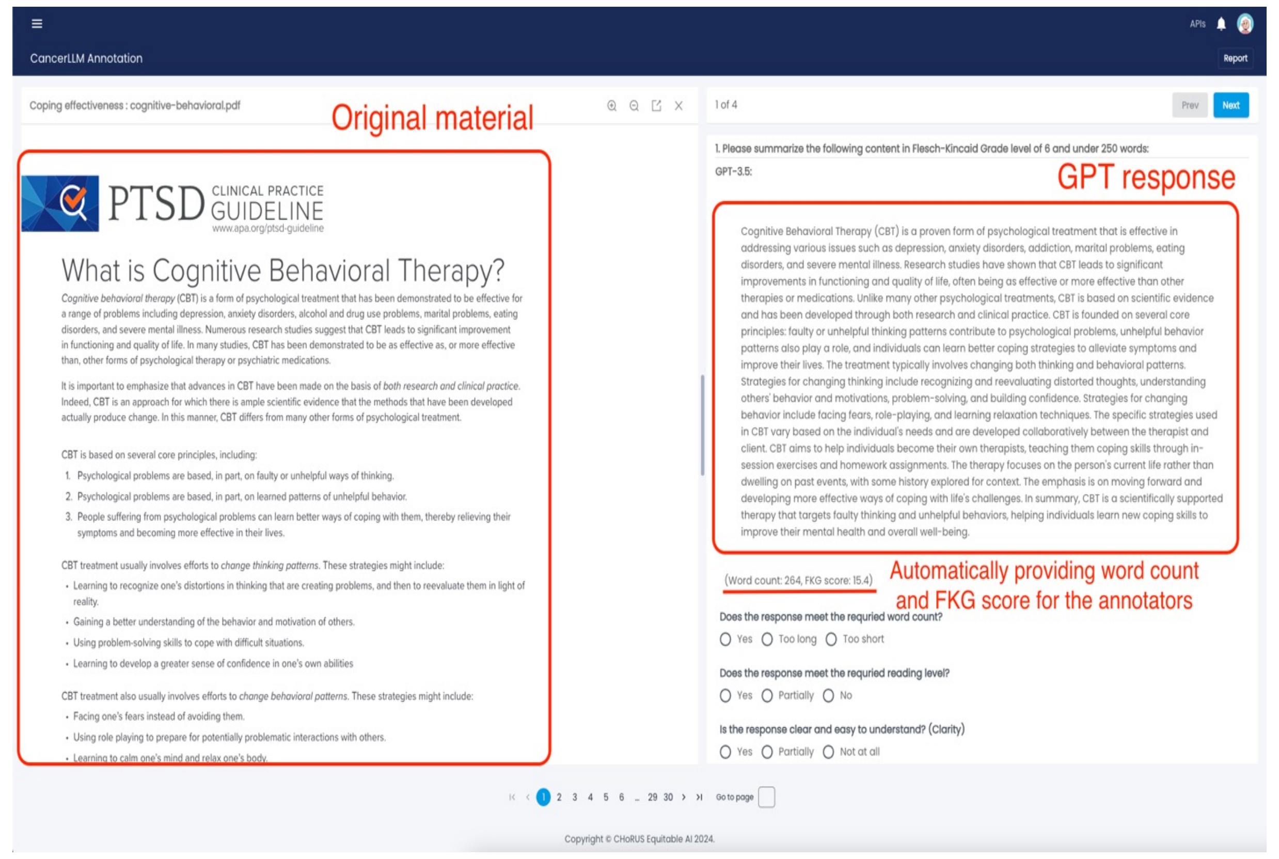
GPT-3.5 Turbo, GPT-4, and GPT-4 Turbo were utilized through Microsoft Azure OpenAI API to generate tailored content.

Prompts:

- At a low reading level (FKG ≤ 6);
- At a word limit of 250;
- In both textual and bulleted formats;
- Providing Spanish and Chinese translations for each topic.

Annotation:

- A panel of **nine oncology experts**, comprising four oncology professors, four doctoral students, and one medical resident.
- Each expert was assigned 10 topics to evaluate based on **seven criteria** and was required to provide feedback on the **errors**.


≡
CancerLLM Annotation
APIs
Report
Original material
GPT response
Automatically providing word count and FKG score for the annotators
Prev
Next
PTSD CLINICAL PRACTICE GUIDELINE
www.apa.org/ptsd-guideline
Go to page

Coping effectiveness : cognitive-behavioral.pdf

What is Cognitive Behavioral Therapy?

Cognitive behavioral therapy (CBT) is a form of psychological treatment that has been demonstrated to be effective for a range of problems including depression, anxiety disorders, alcohol and drug use problems, marital problems, eating disorders, and severe mental illness. Numerous research studies suggest that CBT leads to significant improvement in functioning and quality of life. In many studies, CBT has been demonstrated to be as effective as, or more effective than, other forms of psychological therapy or psychiatric medications.

It is important to emphasize that advances in CBT have been made on the basis of *both research and clinical practice*. Indeed, CBT is an approach for which there is ample scientific evidence that the methods that have been developed actually produce change. In this manner, CBT differs from many other forms of psychological treatment.

CBT is based on several core principles, including:

1. Psychological problems are based, in part, on faulty or unhelpful ways of thinking.
2. Psychological problems are based, in part, on learned patterns of unhelpful behavior.
3. People suffering from psychological problems can learn better ways of coping with them, thereby relieving their symptoms and becoming more effective in their lives.

CBT treatment usually involves efforts to *change thinking patterns*. These strategies might include:

- Learning to recognize one's distortions in thinking that are creating problems, and then to reevaluate them in light of reality.
- Gaining a better understanding of the behavior and motivation of others.
- Using problem-solving skills to cope with difficult situations.
- Learning to develop a greater sense of confidence in one's own abilities

CBT treatment also usually involves efforts to *change behavioral patterns*. These strategies might include:

- Facing one's fears instead of avoiding them.
- Using role playing to prepare for potentially problematic interactions with others.
- Learning to calm one's mind and relax one's body.

1 of 4

1. Please summarize the following content in Flesch-Kincaid Grade level of 6 and under 250 words:

GPT-3.5:

Cognitive Behavioral Therapy (CBT) is a proven form of psychological treatment that is effective in addressing various issues such as depression, anxiety disorders, addiction, marital problems, eating disorders, and severe mental illness. Research studies have shown that CBT leads to significant improvements in functioning and quality of life, often being as effective or more effective than other therapies or medications. Unlike many other psychological treatments, CBT is based on scientific evidence and has been developed through both research and clinical practice. CBT is founded on several core principles: faulty or unhelpful thinking patterns contribute to psychological problems, unhelpful behavior patterns also play a role, and individuals can learn better coping strategies to alleviate symptoms and improve their lives. The treatment typically involves changing both thinking and behavioral patterns. Strategies for changing thinking include recognizing and reevaluating distorted thoughts, understanding others' behavior and motivations, problem-solving, and building confidence. Strategies for changing behavior include facing fears, role-playing, and learning relaxation techniques. The specific strategies used in CBT vary based on the individual's needs and are developed collaboratively between the therapist and client. CBT aims to help individuals become their own therapists, teaching them coping skills through in-session exercises and homework assignments. The therapy focuses on the person's current life rather than dwelling on past events, with some history explored for context. The emphasis is on moving forward and developing more effective ways of coping with life's challenges. In summary, CBT is a scientifically supported therapy that targets faulty thinking and unhelpful behaviors, helping individuals learn new coping skills to improve their mental health and overall well-being.

(Word count: 264, FKG score: 15.4)

Does the response meet the requried word count?

○ Yes ○ Too long ○ Too short

Does the response meet the requried reading level?

○ Yes ○ Partially ○ No

Is the response clear and easy to understand? (Clarity)

○ Yes ○ Partially ○ Not at all

‹ 1 2 3 4 5 6 … 29 30 ›

Copyright © CHoRUS Equitable AI 2024.

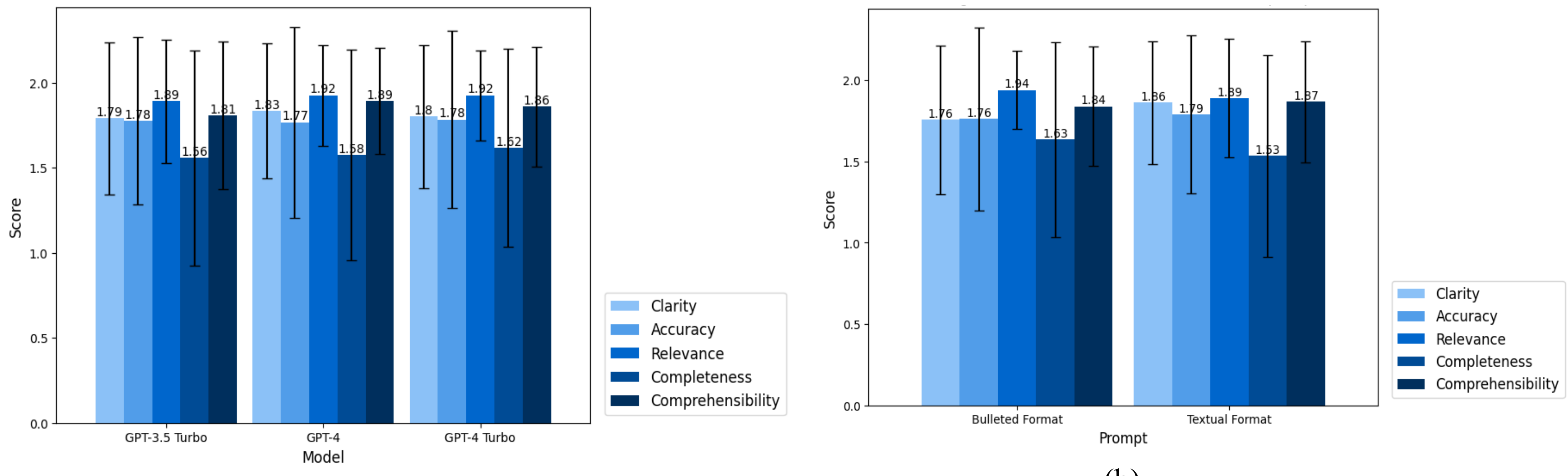
Figure 1. Screenshot of the Cohort Adjudication and Data Annotation application

- ANOVA or Chi-square analyses were employed to compare differences among the various GPT models and prompts.

RESULTS

Table 1. Performance of All Models, Prompts on the Summarization Tasks

	GPT-3.5 Turbo		GPT-4		GPT-4 Turbo	
Prompt	Textual Format	Bullet Points	Textual Format	Bullet Points	Textual Format	Bullet Points
Word Limit (%)	0.467	0.967	0.917	0.767	0.517	0.817
Reading Level (%)	0.183	0.283	0.217	0.217	0.533	0.317
Accuracy	1.767±0.500	1.783±0.49	1.800±0.480	1.733±0.634	1.800±0.48	1.767±0.563
Clarity*	1.833±0.418	1.750±0.474	1.867±0.389	1.800±0.403	1.883±0.324	1.717±0.49
Relevance	1.883±0.415	1.900±0.303	1.883±0.372	1.967±0.181	1.900±0.303	1.950±0.22
Completeness	1.533±0.623	1.583±0.645	1.483±0.624	1.667±0.601	1.583±0.619	1.650±0.547
Comprehensibility	1.817±0.469	1.800±0.403	1.883±0.324	1.900±0.303	1.900±0.303	1.817±0.39
Total Score	8.833±1.748	8.817±1.546	8.917±1.239	9.067±1.26	9.067±1.087	8.900±1.298
Spanish Translation (%)	0.933		0.967		1	
Chinese Translation (%)	0.767		0.867		0.800	



(a) (b)

Figure 2. Average Scores on Each Criterion Between: (a) Different Models; (b) Different Prompts.

- 74.2% (n=360) adhering to the specified word limit and achieving an average quality assessment score of 8.933 out of 10
- Achieving an accuracy of 88.9% for Spanish and 81.1% for Chinese translations
- Errors: inaccurate scope, expression, definition, meaningless points

DISCUSSION

- Overall, it is proven that LLMs are highly effective in tailoring, condensing, and translating educational content for underserved cancer patients and their caregivers.
- The findings from this study can inform the development and implementation of interventions in cancer symptom management and health equity.

CONCLUSION

- This study highlights the application of LLMs in cancer care and education while acknowledging their potential limitations.

ACKNOWLEDGMENT

- The study was supported by the Oncology Nursing Foundation Research Grant (2022 RE03).

yufen.lin@emory.edu