Leveraging Large Language Models for Personalized Public Messaging

Amit Kumar Das
Computer Science
Stony Brook University
Stony Brook, New York, USA
amitkumar.das@stonybrook.edu

Cindy Xiong Bearfield
Interactive Computing
Georgia Tech
Atlanta, Georgia, USA
cxiong@gatech.edu

Klaus Mueller
Computer Science
Stony Brook University
Stony Brook, New York, USA
mueller@cs.sunysb.edu

Abstract

We present a novel methodology for crafting effective public messages by combining large language models (LLMs) and conjoint analysis. Our approach personalizes messages for diverse personas – context-specific archetypes representing distinct attitudes and behaviors – while reducing the costs and time associated with traditional surveys. We tested this method in public health contexts (e.g., COVID-19 mandates) and civic engagement initiatives (e.g., voting). A total of 153 distinct messages were generated, each composed of components with varying levels, and evaluated across five personas tailored to each context. Conjoint analysis identified the most effective message components for each persona, validated through a study with 2,040 human participants. This research highlights LLMs' potential to enhance public communication, providing a scalable, cost-effective alternative to surveys, and offers new directions for HCI, particularly for the design of adaptive, user-centered, persona-driven interfaces and systems.

CCS Concepts

• **Human-centered computing** → **User studies**; *Empirical studies in HCI*; • **Computing methodologies** → *Machine learning methodologies*; • **Information systems** → Personalization; • **Applied computing** → Psychology.

Keywords

Large Language Models (LLM), Public Communication, Message Personalization, Conjoint Analysis, Persona-based Messaging, and human-AI collaboration

ACM Reference Format:
Amit Kumar Das, Cindy Xiong Bearfield, and Klaus Mueller. 2025. Leveraging Large Language Models for Personalized Public Messaging. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3706599.3720018

ACM ISBN 979-8-4007-1395-8/25/04
https://doi.org/10.1145/3706599.3720018

1 Introduction

In today's fast-paced digital age, creating effective public messages is vital yet challenging, particularly when addressing diverse audiences [38]. Traditional methods like surveys and focus groups [5] provide insights but are often slow, costly, and may fail to capture real-time shifts in public opinion [43]. Further, individual differences, such as personality traits, significantly influence how messages are received [12]. For example, extroverted individuals may respond differently to messages than introverted ones [30]. These factors complicate efforts to tailor public communications effectively.

Recent advancements in large language models (LLMs) [22] offer new opportunities for personalized communication by simulating human-like responses and adapting to diverse personas [47]. Inspired by this, we combine LLMs with conjoint analysis [11, 28], a statistical technique that models preferences by examining trade-offs between message components. This approach provides a scalable framework for tailoring messages to diverse audiences and contexts.

We focus on two critical scenarios: health (e.g., COVID-19 preventive measures) and civic engagement (e.g., voting behavior). These areas represent distinct public messaging challenges, with health mandates requiring immediate action and civic engagement involving longer-term attitudes. By applying our framework to these scenarios, we aim to generalize insights to other contexts, such as climate change communication and misinformation countering.

Our work also contributes to HCI by exploring persona-oriented communication in adaptive systems [20], enabling tools and interfaces to tailor interactions based on user preferences dynamically. This fosters applications like intelligent tutors that adapt to individual learning styles, personalized health assistants offering context-aware advice, and error messages that adjust tone and detail to user expertise, reducing frustration and improving task completion rates.

To guide our investigation, we address the following research questions:

- **RQ1:** How accurately can LLMs simulate human responses to public messages, and what implications does this have for using LLMs in evaluating public messages compared to traditional methods?
- **RQ2:** How do different message components resonate with various personality types and scenarios? Are there universal preferences across contexts?

Through this research, we contribute to personalized communication and HCI by developing a framework that uses LLMs to simulate responses from diverse personality types, reducing the need for extensive surveys and enabling scalable personalization. We applied conjoint analysis to identify effective message components and validated our LLM-based approach with human participants, highlighting its reliability and limitations. Finally, we provide guidelines for crafting effective, audience-specific public messages.

2 Related Work

Our research draws on several key areas in HCI and related fields, particularly personalization, adaptive interfaces, and the integration of large language models (LLMs) in human-centric applications. Personalization has long been a focus in HCI, with foundational work by Mackay [29] establishing principles for adaptive interfaces that learn from user behavior. Subsequent studies, such as those by Nov et al. [34] and Thomas et al. [41], demonstrated the potential of personality-targeted design to improve engagement on digital platforms. In public health communication, Bhattacharjee et al. [3] showed that personalized mobile interventions led to better adherence, emphasizing the importance of tailored approaches in critical domains.

Message tailoring has also evolved significantly with advancements in technology. Early principles from Peppers and Rogers [37] have been refined by modern research, revealing how context, platform, and privacy concerns shape user receptiveness [32, 42]. Leveraging AI further enhances personalized messaging, as shown by Costello et al. [8], who demonstrated how AI-mediated communication (AI-MC) can address specific user concerns. Hancock et al. [17] provided a framework for AI-MC, emphasizing user agency, while Dhillon et al. [10] proposed guidelines for designing AI writing assistants that balance autonomy with AI capabilities.

Public health and civic engagement are prominent domains of personalized communication. Kamal et al. [23] developed an adaptive health messaging system that improved engagement by 31% through real-time preference learning. Similarly, Kreuter et al. [25] reported an 18% higher quit rate for smoking cessation programs with tailored messaging. Civic engagement research by Harrington et al. [18] showed how personalized platforms increased voter participation, building on earlier findings by Green and Gerber [15].

Personality modeling has been central to HCI, with the "Big Five" traits [31] providing a foundation for nuanced approaches. Wu et al. [45] demonstrated that personality-aware recommender systems improve user satisfaction. Domain-specific personality scales, such as those developed by Deng et al. [9] and Marta et al. [33], have proven more effective in predicting behavior in contexts like health decisions and civic engagement. Our study builds on these findings by synthesizing insights from Roozenbeek et al. [39] and Lin et al. [26] to develop a customized personality scale tailored to public health and civic engagement contexts.

LLMs have emerged as transformative tools in HCI, enabling natural and adaptive communication through advanced pattern recognition [46]. Their applications span personalization [24], accessibility [1], and market research [4], offering cost-effective and scalable solutions. Studies by Evan et al. [40] highlight their ability to simulate human responses while maintaining ecological validity. These advancements demonstrate the potential of LLMs to open new opportunities for personalized interactions in HCI systems.

3 Methodology

We focused on two public communication scenarios: health mandates (precisely, COVID-19 preventive measures) and civic engagement (voting behavior). We chose these areas because they represent different public messaging challenges and are relevant to a broad audience. We conducted studies to examine the impact of messaging content and audience personality on messaging effectiveness.

First, we generated messages with varying content and identified personality types to experiment with (see Sections 3.1 and 3.2). Next, we instructed an LLM to act as a virtual human participant and presented the LLM with pairs of messages. We prompted the LLM to adopt a specific personality type when reacting to the message and determine which message would be more persuasive for that personality. We analyzed the results via *conjoint analysis* (see Section 3.3).

To assess the influence of information relevance, we evaluated how both relevant and irrelevant risk factors in the message content, as well as relevant and irrelevant personality types, affected LLM's reactions and preferences to messages (included in the supplementary materials). Finally, we compared the LLM-generated results with a human-subject study to evaluate the LLMs' ability to simulate human responses, with details of the human study described in Section 3.4.

3.1 Designing the Messages

We designed 153 unique messages for each scenario by combining components tailored to different personality types. Each message consisted of up to three 'features,' with varying 'levels' for each feature. These messages were tested using both LLM simulations and human participants. The three features are:

- *Feature 1*: Core message content with 17 variations covering risk factors, statistics, and value-based appeals to test persuasiveness across personality types.
- *Feature 2*: Three levels of local context comparing situations between counties to measure geographic influence.
- *Feature 3*: Three variations of outcome framing using success or failure stories to test how message tone affects different personalities.

Table 1 show the features and levels used in the COVID-19 health mandates and voting scenarios, respectively. For the COVID-19-related messages, we identified relevant features based on the study conducted by Coelho et al. [6]. For the voting-related messages, we utilized ChatGPT to generate a range of reasons for poor road conditions, from which we selected a diverse subset to craft the final message components.

3.2 Designing the Personality Types.

To tailor messages to diverse personality types, we identified five profiles spanning from resistance to full support. For the COVID-19 scenario, profiles range from Extreme Opposition (P1), marked by active resistance, to Compliance (P5), characterized by full adherence to mandates. Similarly, for voting, personalities range from Skeptical Critics (P1), who distrust government efficacy, to Enthusiastic Advocates (P5), who strongly support government initiatives.

Table 1: COVID-19 Health Mandates and Voting Messages (Feature 1, Feature 2, and Feature 3)

Feature	Level	COVID-19 Health Mandate Messages	Voting Messages
1	0	Please consider your health and our community's safety; complying with mask-wearing and social distancing mandates is essential to reduce COVID-19's spread.	Your vote is your voice. Make sure it is heard by participating in the upcoming election to help improve the roads and transportation in your county.
	1	[F1, L0] + Scientific studies have shown that certain socio-economic risk factors can accelerate the spread of the virus. Your county has several such risk factors.	[F1, L0] + Studies show that certain socio-economic factors can increase the risk of poor road conditions. Your county has several such risk factors.
	2	[F1, L1] + Your county has one such risk factor.	[F1, L1] + Your county has one such risk factor.
	3	[F1, L1] + Your county has one such risk factor: the presence of residents who lack a high school diploma.	[F1, L1] + Your county has one such risk factor: high annual snowfall.
	4	[F1, L1] + In your county, the risk factor is that 24% of residents lack a high school diploma.	[F1, L1] + In your county, the risk factor is an average annual snowfall of 60 inches.
	5	[F1, L1] + In your county, the risk factor is that 24% of residents lack a high school diploma, significantly worse, exceeding the US average.	[F1, L1] + In your county, the risk factor is an average annual snowfall of 60 inches, significantly higher than the national average.
	6	[F1, L1] + In your county, the risk factor is that 24% of residents lack a high school diploma, while the US average is 14%.	[F1, L1] + In your county, the risk factor is an average annual snowfall of 60 inches, while the national average is 28 inches.
	...		
	16	[F1, L1] + In your county, the three risk factors are that 24% of residents lack a high school diploma (the US average is 14%), 15% do not have health insurance (US 12%), and 70% are minorities (US 26%).	[F1, L1] + In your county, the three risk factors are an average annual snowfall of 60 inches (national average is 28), a distance of 100 miles from the nearest urban center (national average is 40 miles), and an unemployment rate of 8% (national average is 3.6%).
2	0	No Description	No Description
	1	There is a county nearby in the US, neighboring yours, that shares similar risk factors with your area.	There is a nearby county in your state that faces similar road condition challenges.
	2	There is a county elsewhere in the US, not neighboring yours, that shares similar risk factors with your area.	There is a county in another state that faces similar road condition challenges.
3	0	No Description	No Description
	1	Residents in another county with risk factor(s) similar to yours, and high COVID-19 death rates, adhered to health mandates such as mask-wearing and social distancing. A few months later, the death rates in that county stabilized.	In the last election, a county with historically low voter turnout saw a 30% increase in participation. The newly elected officials implemented a comprehensive plan to address road infrastructure. As a result, this county now has much better roads, leading to better economic opportunities and safer, more convenient travel.
	2	Residents in another county with risk factor(s) similar to yours, and high COVID-19 death rates, did not adhere to health mandates such as mask-wearing and social distancing. A few months later, the death rates in that county increased further.	In a nearby county, voter turnout remained low in the last election, with only 25% of eligible voters participating. Consequently, the local government failed to prioritize critical issues such as road infrastructure maintenance. As a result, the county's roads further deteriorated, traffic slowed even more, and car owner repair bills skyrocketed.

Table 2: Personality Scenarios: COVID-19 Health Mandates and Voting

Personality	Scenario: COVID-19 Health Mandates	Scenario: Voting
Personality 1	**Extreme Opposition:** Imagine you are a person who vehemently opposes government health mandates and may actively resist or defy them. You may also engage in protests, civil disobedience, or legal challenges to express your opposition.	**Skeptical Critic**: Imagine you are a person who believes that the government is ineffective and often makes things worse. Bureaucracy, corruption, and waste are rampant, and government interventions typically create more problems than they solve. You also believe that minimal government involvement and that private sector solutions are almost always better.
...		
Personality 5	**Compliance:** Imagine you are a person who fully complies with government health mandates and supports them as necessary measures to control the spread of the pandemic. You may also believe in the importance of collective action for public health and safety.	**Enthusiastic Advocate**: Imagine you are a person who believes that the government is highly effective and plays a crucial role in ensuring social welfare, justice, and economic stability. Public programs and regulations are essential for addressing inequalities and providing opportunities for all citizens. You also believe in the power of government to bring about positive change.

Our framework draws on the Theory of Planned Behavior [2], which links attitudes to behavioral intentions, and public health response models [13], which identify behavioral spectrums from resistance to acceptance. Instead of the Big Five [14], we used profiles tailored to specific scenarios, informed by studies on public health [26, 39] and civic engagement. We used GPT-3.5 and GPT-4 to simulate responses for each personality profile in the COVID-19 and voting scenarios. Scripts were created for each of the five profiles per scenario, resulting in 20 unique simulations (5 profiles × 2 scenarios × 2 LLMs). Each script described the personality type and guided the model to respond as a human with that persona. Each script was run twice for consistency, using OpenAI's API with settings optimized for precision and focus (e.g., temperature set to 0.2). This ensured detailed, consistent outputs for each personality-scenario combination. For instance, responses for Personality 1 in the COVID-19 scenario were cross-validated between runs to confirm consistency. LLMs were prompted to assume the role of the described personality and evaluate which message from a pair was more persuasive.

3.3 Evaluating Message Preferences with Conjoint Analysis

We used choice-based conjoint analysis to identify the most effective message components for each personality type. This method presents two messages at a time for comparison, helping assess feature importance. Our study focused on three message features: Feature 1 (17 levels), Feature 2 (3 levels), and Feature 3 (3 levels), resulting in 153 possible combinations. To ensure reliable results, we sampled 4,250 pairwise comparisons from the 11,628 possible pairs [21, 35], ensuring each feature level appeared equally often. This balanced design minimized bias and allowed fair evaluation of all components. Each personality type was tested independently in separate sessions to avoid cross-contamination of LLM memory. This setup ensured that preferences and patterns remained isolated for each profile. By analyzing these responses, we identified key features influencing message persuasiveness overall, variations in feature preferences across personality types, and universal preferences effective across all profiles.

3.4 Human Participant Study

To validate our LLM-based approach, we conducted a parallel study with human participants, focusing on the COVID-19 scenario and the related personality types (Table 2: Scenario COVID-19 health mandates). We recruited 2,040 participants to cover the five different personality types identified in our COVID-19 scenario. This sample size was determined based on recommendations from conjointly.com [7], which suggested that 400 participants per scenario would yield reliable and statistically significant results given our study setup.

Participants were recruited through Prolific.com [36] and various social media platforms, including Reddit groups, ensuring a diverse sample in terms of age, gender, and educational background. This approach to participant recruitment aimed to mirror the diversity of perspectives that our LLM simulations attempted to capture.

We chose to focus on the COVID-19 scenario for the human participant study because the simulations for both the COVID-19 and voting scenarios revealed similar patterns of results, which suggests that results likely can generalize across contexts.

During the study, participants first responded to demographic questions via Qualtrics. Next, participants were presented with descriptions of five distinct personalities and asked to select the one that best represented them. Based on their choice, they were redirected to the Conjointly.com platform [7], where we replicate the same conjoint study previously conducted with the LLM. We used the same conjoint analysis method to analyze the data from the human participants, so we can directly compare the results from our LLM simulations with real human preferences.

4 Results

Our analysis highlights one representative personality types: P1 (Extreme Opposition). Results for P2 (Opposition), P3 (Neutral), P4 (Conditional Support), and P5 (Compliant) followed similar trends and are included in the supplementary materials.

4.1 LLM Simulation Results

The analysis of relative feature importance across personality types (shown in Fig. 1a revealed consistent patterns in how different groups evaluate messages. Feature 1 (Risk Factor Information) dominated message effectiveness across all personality types.

Feature 3 (Success/Failure Stories) emerged as the second most crucial factor but with notable variations. All groups clearly preferred success stories (Level 1) over failure stories, indicating that positive framing works across personality types. Finally, Feature 2 (Neighboring Counties) showed consistently low importance across all groups. This suggests that comparing situations with other locations has low message effectiveness, regardless of personality type.

4.2 Key Findings from the LLM Study.

The COVID-19 scenario revealed that Feature 1 (Risk Factors Information) dominated the importance rankings (see fig. 2a). Participants favored detailed, statistically backed information, shown by their preference for Level 16 in Feature 1. Feature 3 (Success/Failure stories) emerged as the second most influential element. Feature 2 (Neighboring Counties) showed minimal impact, suggesting that local comparisons play a limited role in effective messaging. These patterns, which also appeared in the voting scenario, suggest that despite some personality-based variations, the most effective COVID-19 health messages focus on detailed, data-driven content and positive outcome stories.

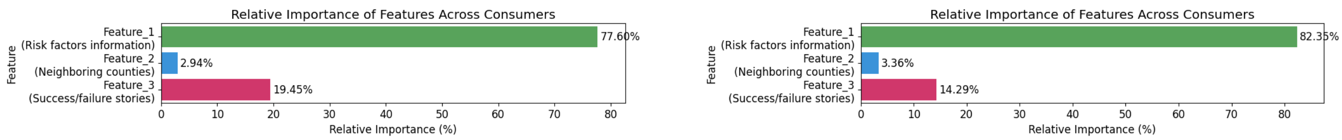
4.3 Human Participant Study Results

We conducted a parallel study with human participants for the COVID-19 scenario to validate our LLM approach. Our study included 2,040 participants (930 women, 120 non-binary, 990 men) from the United States, with a relatively balanced distribution across personality types. Figures 1b and 2b display the importance values and level preferences for the Extreme Opposition personality type. The results from our human study strongly corroborate the findings from our LLM simulations, demonstrating a remarkable alignment between LLM predictions of human preferences and actual human preferences. We calculated Spearman's rank correlations to examine the relationship between LLM and human responses based on their 23 relative level importance values. The analysis revealed very strong correlations (see Fig. 3) across all five personality types ($\rho <$ 0.001 across all types).

Feature 1 (Risk Factor Information) emerged as the dominant factor across all personality types. Across all personality types, Feature 1 was the most important. All personalities strongly preferred detailed, statistically supported information (Level 16 in Feature 1). Feature 3 (Success/Failure Stories) was consistently the second most important feature, with a preference for success stories across all personality types. Feature 2 (Neighboring Counties) had low importance across all personalities, but its importance increased for the neighboring county.

4.4 Cost-Effectiveness Analysis

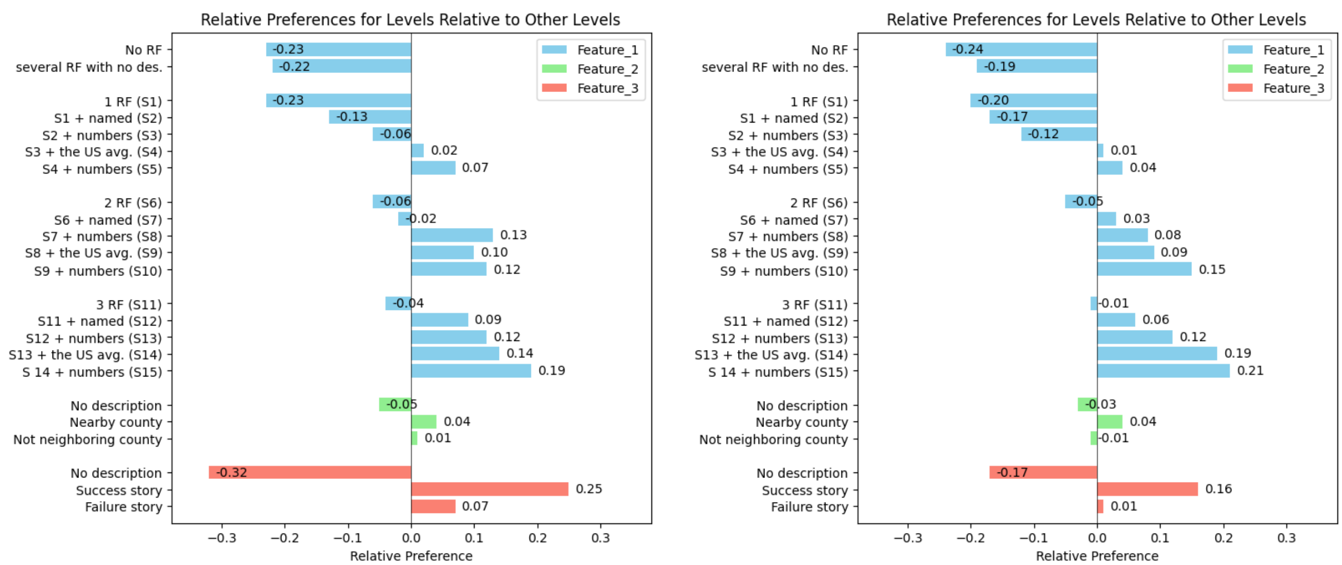
We compared three approaches: Conjointly.com, Prolific.com, and GPT-4 (see Table 3). For studying one personality type with 400 participants: Conjointly.com costs $1,895 but doesn't allow personality targeting. Prolific.com costs $1,200 ($3 per participant) and supports personality targeting. GPT-4 Turbo costs $15.48 total ($15.02 for input tokens, $0.46 for output tokens). GPT-4 cuts costs by over 98% compared to both platforms and reduces data collection time to one hour. While GPT-4 provides simulated responses instead of real participant data, our earlier results showed strong correlations between the two methods.



(a) GPT 4o (Personality Type 1)

(b) Human Response (Personality Type 1)

Figure 1: Relative Importance of Features Across Personality Types (COVID-19)



(a) GPT 4o

(b) Human response

Figure 2: Comparison analysis for conjoint analysis for Personality 1

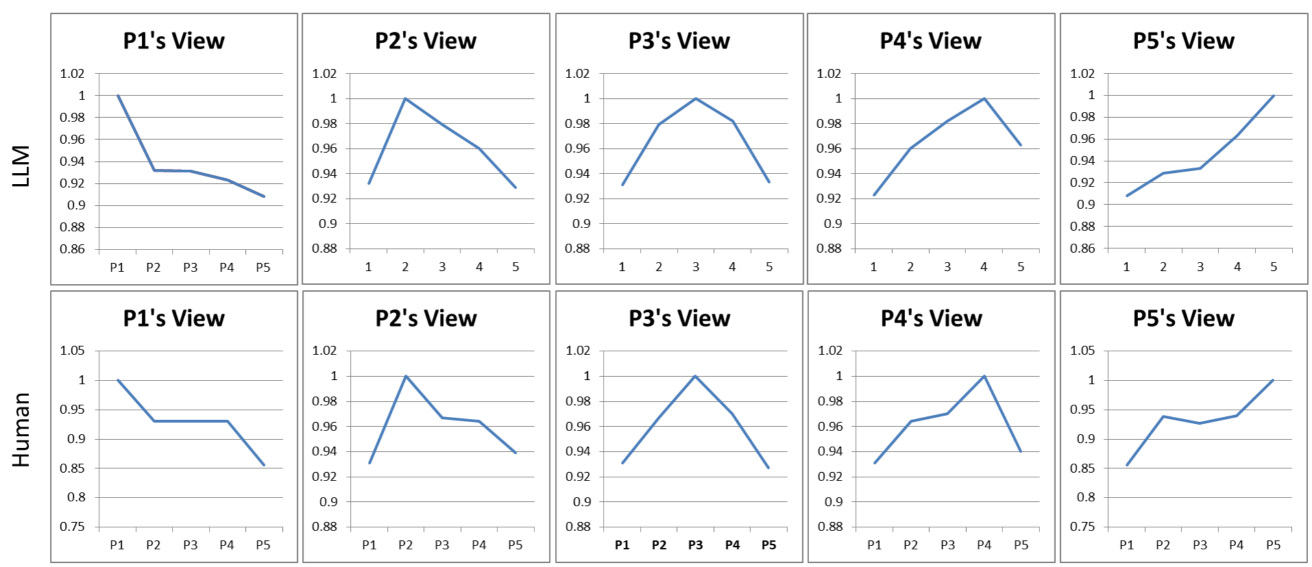


Figure 3: Correlation between pairs of personas in the COVID-19 messaging study, showing both LLM results (top) and human results (bottom). Each section shows how one persona relates to all others.

Table 3: Cost Comparison: LLM vs Human Studies

Category	Conjointly	Prolific	GPT-4
Platform Cost	$1,895	$1,200	-
Participant Pay	Incl.	Incl.	$0
Token Cost (In/Out)	-	-	$15.02 / $0.46
Total (1 personality)	$1,895	$1,200	$15.48
Cost Reduction	-	36.68% (vs Conjointly)	99.18% (vs Conjointly) / 98.71% (vs Prolific)
Est. Data Coll. Time	Not spec.	Variable	~1 hour
Personality Def.	Not possible	Challenging	Precise

5 Summary of findings

In the following we review our two research questions in light of our findings.

RQ1: How accurately can LLMs simulate human responses to public messages, and what are the implications compared to traditional studies? LLMs generate responses closely aligned with human behavior, capturing diverse perspectives and automating public message evaluation. This reduces reliance on costly, time-intensive human studies prone to biases like order effects and participant fatigue.

RQ2: Which messaging components are effective, and how do personality types and scenarios influence effectiveness? Effective messages include key risk factors and success stories, with personality types significantly influencing resonance. For example, Extreme Opposition (P1) prefers minimal, evidence-based information, while Neutral (P3) and Compliant (P5) favor comprehensive details. Tailoring messages to personality traits enhances their effectiveness and provides valuable guidance for public campaigns.

6 Discussions, Implications, and Limitations

Our findings emphasize the value of data-driven design, allowing teams to test messages across personality types to optimize outreach. By leveraging LLMs, organizations can conduct cost-effective evaluations alongside traditional studies, reducing time and expenses. Additionally, strategies can dynamically adapt based on systematic testing, minimizing reliance on intuition and improving the precision of communication efforts. LLMs efficiently prototype insights comparable to human studies, supporting early-stage evaluations and aligning with prior HCI work [16]. They address individual differences by simulating diverse traits, overcoming traditional challenges like limited participant numbers and non-diverse samples [19, 27], fostering personalized and inclusive solutions.

While our study provides valuable insights, it is essential to acknowledge its limitations. Despite their sophistication, LLMs have inherent limitations in understanding context and nuance [44]. Our five personality types are necessarily simplifications, and the current study focused on a specific context. Future work should continue refining LLM to capture the complexity of human responses better, explore more nuanced personality models, and investigate cross-cultural applications and variations. Potential research directions could include integrating real-time data to dynamically adjust message strategies, exploring applications in other domains, and developing ethical frameworks for the responsible use of AI in public communication.

7 Conclusion

Our study highlights the effectiveness of persona-based message evaluation using LLMs, accelerating the optimization of communication strategies across diverse audiences. This approach, albeit impactful in itself, offers significant potential for HCI interfaces to deliver more personalized, adaptive user experiences by tailoring interactions based on nuanced understanding of user personas. As AI-driven personalization evolves, HCI systems can achieve greater engagement and usability, provided ethical considerations and real-world validation remain central to their design and deployment.

Acknowledgments

This material is based upon work supported by NSF under Grant NRT-HDR 2125295 and IIS-2237585. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] Mohammad Abolnejadian, Sharareh Alipour, and Kamyar Taeb. 2024. Leveraging ChatGPT for Adaptive Learning through Personalized Prompt-based Instruction: A CS1 Education Case Study. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 521, 8 pages. https://doi.org/10.1145/3613905.3637148

[2] Icek Ajzen. 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50, 2 (1991), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T Theories of Cognitive Self-Regulation.

[3] Ananya Bhattacharjee, Jiyau Pang, Angelina Liu, Alex Mariakakis, and Joseph Jay Williams. 2023. Design Implications for One-Way Text Messaging Services that Support Psychological Wellbeing. 30, 3, Article 34 (June 2023), 29 pages. https://doi.org/10.1145/3569888

[4] James Brand, Ayelet Israeli, and Donald Ngwe. 2023. *Using LLMs for Market Research.* Harvard Business School Marketing Unit Working Paper 23-062. Harvard Business School. Available at SSRN: https://ssrn.com/abstract=4395751 or http://dx.doi.org/10.2139/ssrn.4395751.

[5] L. Y. Chan, N. El-Haddad, B. Freeman, B. J. O'Hara, L. Woodland, and B. Harris-Roxas. 2021. A case study of an sms text message community panel survey and its potential for use during the covid-19 pandemic. *JMIR Formative Research* 5 (2021), e28929. Issue 11. https://doi.org/10.2196/28929

[6] Darius Coelho, Nikita Gupta, Eric Papenhausen, and Klaus Mueller. 2022. Patterns of social vulnerability-an interactive dashboard to explore risks to public health on the us county level. In *2022 Workshop on Visual Analytics in Healthcare (VAHC)*. IEEE, 01–09.

[7] Conjointly. 2024. Conjointly. https://www.conjointly.com Accessed: 2024-09-12.

[8] Thomas H Costello, Gordon Pennycook, and David G Rand. 2024. Durably reducing conspiracy beliefs through dialogues with AI. *Science* 385, 6714 (2024), eadq1814.

[9] Li Deng, Shiqian Luo, Qiang Fang, and Jun Xu. 2023. Intertemporal decision-making as a mediator between personality traits and self-management in type 2 diabetes: a cross-sectional study. *Frontiers in Psychology* 14 (Jul 2023), 1210691. https://doi.org/10.3389/fpsyg.2023.1210691

[10] Paramveer S. Dhillon, Somayeh Molaei, Jiaqi Li, Maximilian Golub, Shaochun Zheng, and Lionel Peter Robert. 2024. Shaping Human-AI Collaboration: Varied Scaffolding Levels in Co-writing with Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA)

(CHI '24). Association for Computing Machinery, New York, NY, USA, Article 1044, 18 pages. https://doi.org/10.1145/3613904.3642134

[11] Felix Eggers, Henrik Sattler, Thorsten Teichert, and Franziska Völckner. 2021. Choice-based conjoint analysis. In *Handbook of market research*. Springer, 781–819.

[12] K. P. Frey and A. H. Eagly. 1993. Vividness can undermine the persuasiveness of messages. *Journal of Personality and Social Psychology* 65 (1993), 32–44. Issue 1. https://doi.org/10.1037/0022-3514.65.1.32

[13] Ilona Fridman, Nicole Lucas, Debra Henke, and Christina K Zigler. 2020. Association Between Public Knowledge About COVID-19, Trust in Information Sources, and Adherence to Social Distancing: Cross-Sectional Survey. *JMIR Public Health Surveill* 6, 3 (15 Sep 2020), e22060. https://doi.org/10.2196/22060

[14] Samuel D Gosling, Peter J Rentfrow, and William B Swann Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality* 37, 6 (2003), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

[15] Donald P Green and Alan S Gerber. 2008. *Get out the vote: How to increase voter turnout*. Brookings Institution Press.

[16] Yuanning Han, Ziyi Qiu, Jiale Cheng, and RAY LC. 2024. When Teams Embrace AI: Human Collaboration Strategies in Generative Prompting in a Creative Design Task. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 176, 14 pages. https://doi.org/10.1145/3613904.3642133

[17] Jeffrey T Hancock, Mor Naaman, and Karen Levy. 2020. AI-Mediated Communication: Definition, Research Agenda, and Ethical Considerations. *Journal of Computer-Mediated Communication* 25, 1 (January 2020), 89–100. https://doi.org/10.1093/jcmc/zmz022

[18] Christina Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing Community-Based Collaborative Design: Towards More Equitable Participatory Design Engagements. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 216 (Nov. 2019), 25 pages. https://doi.org/10.1145/3359318

[19] Craig Hedge, Georgina Powell, and Petroc Sumner. 2018. The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior research methods* 50 (2018), 1166–1186.

[20] Eric Horvitz. 1999. Principles of mixed-initiative user interfaces *(CHI '99)*. Association for Computing Machinery, New York, NY, USA, 159–166. https://doi.org/10.1145/302979.303030

[21] Richard Johnson and Bryan Orme. 2003. *Getting the Most from CBC*. Sawtooth Software Research Paper Series. Sawtooth Software. https://sawtoothsoftware.com/resources/technical-papers/getting-the-most-from-cbc

[22] Katikapalli Subramanyam Kalyan. 2023. A survey of GPT-3 family large language models including ChatGPT and GPT-4. *Natural Language Processing Journal* (2023), 100048.

[23] Noreen Kamal, Sidney Fels, Michael Fergusson, Jenny Preece, Dan Cosley, and Sean Munson. 2013. Designing social media for change. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 3183–3186. https://doi.org/10.1145/2468356.2479642

[24] Jiho Kim, Ray C. Flanagan, Noelle E. Haviland, ZeAi Sun, Souad N. Yakubu, Edom A. Maru, and Kenneth C. Arnold. 2024. Towards Full Authorship with AI: Supporting Revision with AI-Generated Views. arXiv:2403.01055 [cs.HC] https://arxiv.org/abs/2403.01055

[25] Matthew W Kreuter, Fiona C Bull, Eddie M Clark, and Debra L Oswald. 1999. Understanding how people process health information: A comparison of tailored and nontailored weight-loss materials. *Health Psychology* 18, 5 (1999), 487–494. https://doi.org/10.1037/0278-6133.18.5.487

[26] C. Lin, V. Imani, N. R. Majd, Z. Ghasemi, M. D. Griffiths, K. Hamilton, M. S. Hagger, and A. H. Pakpour. 2020. Using an integrated social cognition model to predict covid-19 preventive behaviours. *British Journal of Health Psychology* 25 (2020), 981–1005. Issue 4. https://doi.org/10.1111/bjhp.12465

[27] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How weird is CHI?. In *Proceedings of the 2021 chi conference on human factors in computing systems*. 1–14.

[28] Jordan J Louviere. 2000. *Stated choice methods: analysis and applications*. Cambridge University Press.

[29] Wendy E. Mackay. 1991. Triggers and barriers to customizing software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New Orleans, Louisiana, USA) *(CHI '91)*. Association for Computing Machinery, New York, NY, USA, 153–160. https://doi.org/10.1145/108844.108867

[30] Sandra C Matz, Michal Kosinski, Gideon Nave, and David J Stillwell. 2017. Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the national academy of sciences* 114, 48 (2017), 12714–12719.

[31] Robert R McCrae and Paul T Costa Jr. 2008. The five-factor theory of personality. In *Handbook of personality: Theory and research* (3 ed.), Oliver P John, Richard W Robins, and Lawrence A Pervin (Eds.). The Guilford Press, 159–181.

[32] A. K. M. Nuhil Mehdy, Michael D. Ekstrand, Bart P. Knijnenburg, and Hoda Mehrpouyan. 2021. Privacy as a Planned Behavior: Effects of Situational Factors on Privacy Perceptions and Plans. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization* (Utrecht, Netherlands) *(UMAP '21)*. Association for Computing Machinery, New York, NY, USA, 169–178. https://doi.org/10.1145/3450613.3456829

[33] Marta Micheli and Alberto Valese. 2024. Modeling user personality traits from aesthetic preference on multiple images. In *Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization* (Cagliari, Italy) *(UMAP '24)*. Association for Computing Machinery, New York, NY, USA, 189–194. https://doi.org/10.1145/3627043.3659568

[34] Oded Nov and Ofer Arazy. 2013. Personality-targeted design: theory, experimental procedure, and preliminary results. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13)*. Association for Computing Machinery, New York, NY, USA, 977–984. https://doi.org/10.1145/2441776.2441887

[35] Bryan K Orme. 2010. *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research* (2nd ed.). Research Publishers LLC, Madison, Wis.

[36] Stefan Palan and Christian Schitter. 2018. Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17 (2018), 22–27.

[37] Don Peppers and Martha Rogers. 1993. *The one to one future: Building relationships one customer at a time*. Crown Business.

[38] S. Pink, M. N. Stagnaro, J. Chu, J. S. Mernyk, J. G. Voelkel, and R. Willer. 2023. The effects of short messages encouraging prevention behaviors early in the covid-19 pandemic. *Plos One* 18 (2023), e0284354. Issue 4. https://doi.org/10.1371/journal.pone.0284354

[39] J. Roozenbeek, C. R. Schneider, S. Dryhurst, J. R. Kerr, A. L. J. Freeman, G. Recchia, A. M. v. d. Bles, and S. v. d. Linden. 2020. Susceptibility to misinformation about covid-19 around the world. *Royal Society Open Science* 7 (2020), 201199. Issue 10. https://doi.org/10.1098/rsos.201199

[40] Ivan Sekulić, Mohammad Alinannejadi, and Fabio Crestani. 2024. Analysing Utterances in LLM-Based User Simulation for Conversational Search. *ACM Trans. Intell. Syst. Technol.* 15, 3, Article 62 (May 2024), 22 pages. https://doi.org/10.1145/3650041

[41] Veronica Thomas, Marisol Chávez, Erica Browne, and Alexandra Minnis. 2020. Instagram as a tool for study engagement and community building among adolescents: A social media pilot study. *DIGITAL HEALTH* 6 (03 2020), 205520762090454. https://doi.org/10.1177/2055207620904548

[42] Sarah Theres Völkel, Ramona Schödel, Daniel Buschek, Clemens Stachl, Quay Au, Bernd Bischl, Markus Bühner, and Heinrich Hussmann. 2019. Opportunities and challenges of utilizing personality traits for personalization in HCI. *Personalized Human-Computer Interaction* 31 (2019).

[43] Emily Wall, Cindy Xiong, and Yea-Seul Kim. 2022. Vishikers' guide to evaluation: Competing considerations in study design. *IEEE Computer Graphics and Applications* 42, 3 (2022), 29–38.

[44] Huichen Will Wang, Jane Hoffswell, Victor S Bursztyn, Cindy Xiong Bearfield, et al. 2024. How Aligned are Human Chart Takeaways and LLM Predictions? A Case Study on Bar Charts with Varying Layouts. *IEEE Transactions on Visualization and Computer Graphics* (2024).

[45] Wen Wu, Li Chen, and Yu Zhao. 2018. Personalizing recommendation diversity based on user personality. *User Modeling and User-Adapted Interaction* 28, 3 (Aug. 2018), 237–276. https://doi.org/10.1007/s11257-018-9205-x

[46] Rui Zhang, Wen Duan, Christopher Flathmann, Nathan McNeese, Bart Knijnenburg, and Guo Freeman. 2024. Verbal vs. Visual: How Humans Perceive and Collaborate with AI Teammates Using Different Communication Modalities in Various Human-AI Team Compositions. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 437 (Nov. 2024), 34 pages. https://doi.org/10.1145/3686976

[47] Z. Zhang. 2024. Advancements and challenges in ai-driven language technologies: from natural language processing to language acquisition. *Applied and Computational Engineering* 57 (2024), 146–152. Issue 1. https://doi.org/10.54254/2755-2721/57/20241325