Hypothesis Development

LLMs capabilities in building consumer profiles

[TBA]

Performance difference among models

[TBA]

Common patterns LLMs making mistakes in interpreting consumers' personalities.

[TBA]

LLMs capabilities in preparing personalized offerings.

Generating personalized ads text

Trait-based personalization in advertising aligns content with individual personality traits, enhancing user engagement and ad effectiveness ¥cite{Winter2021}. Studies indicate that personality traits significantly influence user responses to ads, especially in social media contexts (Dodoo & Wen, 2019). For instance, users open to new experiences are more likely to engage with content on platforms like Facebook (Amichai-Hamburger & Vinitzky, 2010). Early research by Hirsh et al. (2012) demonstrated that the Big Five personality traits predict user preferences for tailored product descriptions, confirming that ads resonating with user traits receive more favorable evaluations. Furthermore, the quality of message framing can impact ad effectiveness; poorly framed messages can result in negative evaluations despite personalization ¥cite{hirsh2012personalized, Updegraff2007}. Credibility is a crucial aspect of ad quality, where a higher credibility correlates with reduced ad skepticism and more positive attitudes toward the ad ¥cite{Gaber2019; Tran2017}.

The perception of personalization, irrespective of actual personalization, significantly contributes to the effectiveness of web-based ads ¥cite{Li2016}. Perceived relevance of the ad enhances cognitive processing and increases persuasiveness through mechanisms like self-referencing (Wheeler, Petty, & Bizer, 2005). This perceived relevance leads to a more positive attitude towards the ad, which can be quantitatively assessed through consumer feedback on ad likeability, persuasiveness, and interest in the advertised product or service ¥cite{hirsh2012personalized}. Positive attitudes towards ads are linked to higher engagement intentions ¥cite{Tran2017}. Recent studies using GPT3.5 have confirmed the direct correlation between trait-based personalization and positive attitudes towards ads, particularly for traits like Openness, Conscientiousness, and Extraversion ¥cite{Matz2024}.

Building on this foundation, this study utilizes GPT4 to explore the underlying mechanisms of ad effectiveness, focusing on perceived relevance and the moderating role of credibility. We also extend our investigation to measure direct outcomes such as ad click likelihood. Although not central to our hypotheses, we anticipate finding indirect effects of personalization on perceived credibility and attitude towards ads. We also hypothesize that perceived personalization enhances credibility ¥cite{Tran2017}. Thus, we propose the following frameworks and hypotheses:

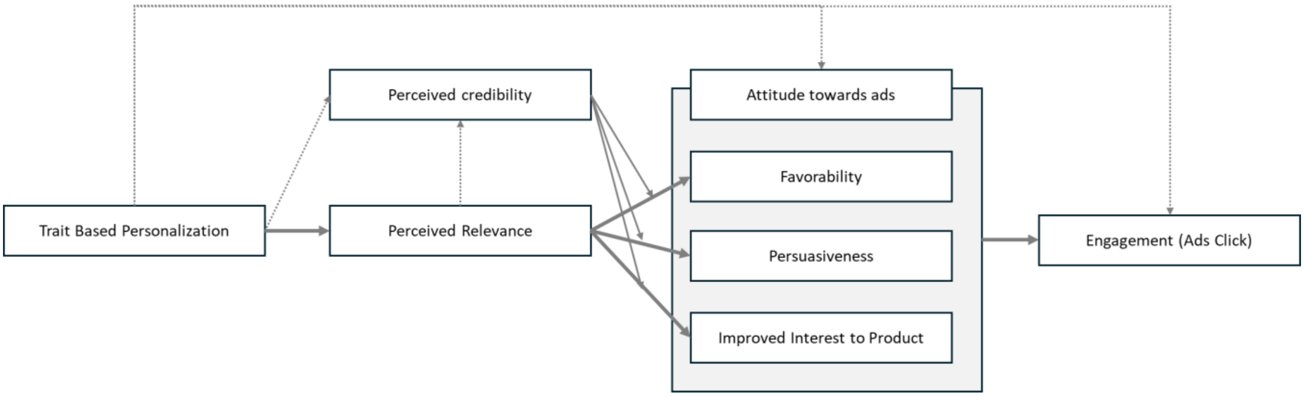

Fig.1: Conceptual Model 1

H6: GPT4 generated personalized ads for each personality trait will significantly affect the following:
H6.1: Generated personalization → Perceived relevance
H6.2: Perceived relevance → Attitude towards ads
H6.3: Perceived credibility moderates → Attitude towards ads
H6.4: Attitude towards ads → Engagement (Ads Click)

Moreover, considering the expansive training data of LLMs like GPT4, which provides a more diverse understanding of human expressions, we hypothesize that a blend of personality traits in personalized ads will outperform ads targeting a single trait.

H7: GPT4 generated personalized ads leveraging blended personality traits will have a larger effect size on the following:
H7.1: Generated personalization → Perceived relevance
H7.2: Perceived relevance → Attitude towards ads
H7.3: Perceived credibility moderates → Attitude towards ads
H7.4: Attitude towards ads → Engagement (Ads Click)

Providing tailored communication as service agent

Leading brands and government agencies are increasingly adopting intelligent agent technologies, such as chatbots, to enhance operational efficiencies and provide on-demand customer services ¥cite{Shumanov2021}. According to similarity-attraction theory, personalized communication is effective because individuals prefer interacting with entities—whether people or computers—that resemble themselves in personality ¥cite{Shumanov2021}. Research indicates that users attribute personalities to computers based on cues from the interface and prefer interactions that reflect congruent personality traits (Nass & Moon, 2000). This preference extends to consumer engagement in services, suggesting that personalized response language can significantly influence customer interactions.

Studies have shown that trust, satisfaction, and positive emotions are critical drivers of customer engagement ¥cite{deOliveiraSantini2020}. We integrate these elements under the umbrella of 'attitude towards AI agents'. Unlike personalized advertising, there is no established precedent that perceived personalization is essential for effective personalization in conversational contexts. Thus, we hypothesize that personalization subconsciously influences attitudes towards AI chatbots.

Previous research indicates that, similar to human interactions, consumers engage more with chatbots that exhibit personality traits akin to their own, such as extraversion or introversion, which leads to increased engagement and service purchases ¥cite{Shumanov2021}. We measure customer engagement by the duration a customer spends interacting with a chatbot, following previous studies ¥cite{Shumanov2021}. Duration is an established performance indicator online and is recognized as a key metric for evaluating repeat visits and enhancing stock returns (Danaher, Mullarkey, & Essegaier, 2006).

Building on these insights, this study aims to explore the influence of the broader range of Big Five personality traits on customer engagement with AI chatbots and to evaluate GPT4's capability in delivering personalized communication. We also investigate how trait-based personalization affects customer engagement by examining its impact on perceived relevance and attitudes toward AI agents through trust, satisfaction, and positive emotions. This leads to our formulated hypotheses:

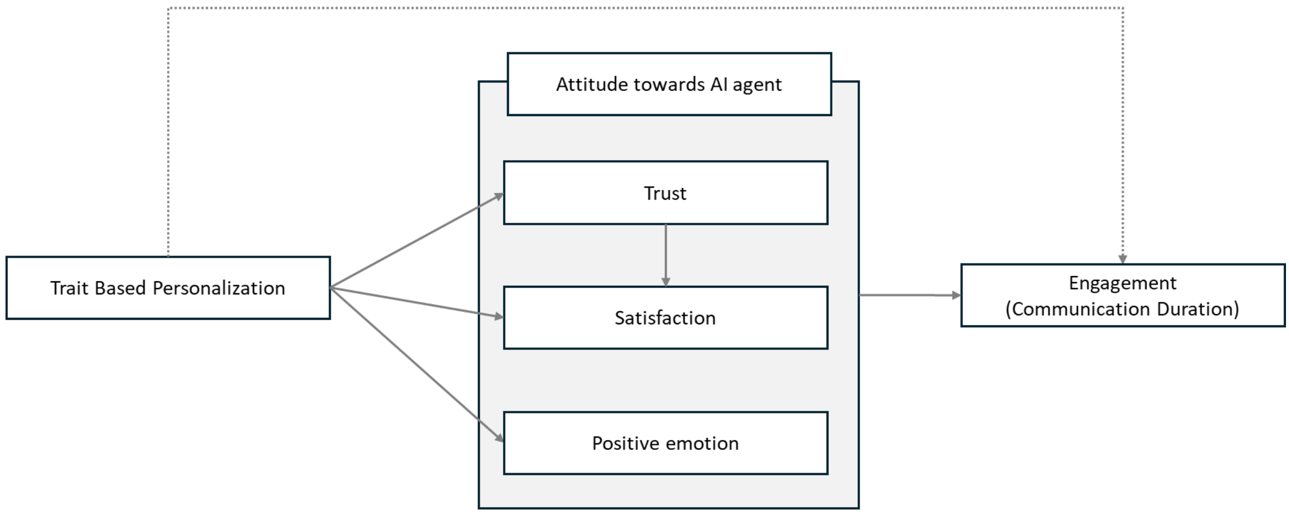

Fig.2: Conceptual Model 2

H8: GPT4 generated personalized conversation via AI chatbot agents tailored to individual personality traits will significantly influence the following:
H8.1: Trait-based personalization → Attitude towards AI agent
H8.2: Attitude towards AI agent → Engagement (Likely duration spent with AI chatbot)

In line with the hypotheses on ad personalization (H7), we further propose:

H9: GPT4 generated personalized interactions for blended personality traits in AI chatbots will exhibit a larger effect size compared to personalization targeting a single personality trait:
H9.1: Trait-based personalization → Attitude towards AI agent
H9.2: Attitude towards AI agent → Engagement (Likely duration spent with AI chatbot)

Method

Study 1

[TBA]

Data Collection

[TBA]

Procedures and Measurements

Study 1a

[TBA]

Study 1b

[TBA]

Study 2

Study 2 explores GPT4's capability to generate tailored advertising content and dynamically adjust communication tones on online shopping platforms, tailored to each of the Big Five personality traits. This involves two components: tailored advertising texts (Study 2a) and chatbot interactions (Study 2b).

Data Collection

We recruited 110 participants via Prolific Academic. After excluding six participants for failing attention checks, 104 remained in the sample. These participants had an average age of 37.2 years (SD = 13.2), with an equal gender distribution (50% female).

Procedures and Measurements

Study 2a

We created ten advertisements, each featuring a photo of an iPhone and text generated by GPT-4 tailored to the high and low ends of a personality trait, such as extraversion and introversion. The text for each ad was crafted to reflect the motivational concerns associated with one of the Big Five personality dimensions. Participants evaluated the ads using two 11-point bi-polar scales, assessing perceived relevance, credibility, favorability, persuasiveness, product interest, and likelihood of clicking on the ads. The responses for favorability, persuasiveness, and increased product interest were averaged to gauge overall ad effectiveness ($\alpha = .94$).

Finally, Participants completed an established measure of the Big Five personality traits (BFI-2S, 63), which asks participants to rate their agreement with a set of 30 statements. Responses were recorded on a 7-point scale ranging from 1 = Strongly Disagree to 7 = Strongly Agree. With Cronbach's alphas ranging from 0.78 to 0.87, the scale reliabilities were found to be good (Openness = 0.82, Conscientiousness = 0.82, Extraversion = 0.83, Agreeableness = 0.78 and Neuroticism = 0.87). Participants also responded to a series of socio-demographic questions, including age, gender, ethnicity, employment status and education.

We performed linear regression analyses to test if people prefer ads that are congruently framed by GPT-4, assessing the influence of each Big Five trait and control variables such as age, gender, and education on the continuous ratings for each trait, addressing H6.

Study 2b

Similarly, we designed ten chatbot interactions for an e-commerce website, manipulating each to highlight a different Big Five personality trait. We generated two sets of responses for common customer queries, such as "Can you help me track my order?" and "Do you have this shirt in other colors?" tailored to high and low scores in each trait. Participants rated each interaction using similar bi-polar scales as in Study 2a, focusing on perceived relevance, trust, satisfaction, positive emotion, and increased AI chatbot usage. Scores for trust, satisfaction, and positive emotion were combined to form an overall positive attitude towards the AI chatbot ($\alpha = .94$).

We conducted linear regression analyses to evaluate the impact of trait-based personalization in chatbot interactions, similar to the procedure in Study 2a, aiming to validate our conceptual model of LLMs' capability in delivering personalized conversations, addressing H8.

Study 3

Study 3 investigates GPT-4's ability to create tailored advertising content and adjust communication tones for online shopping platforms based on a blended model of the Big Five personality traits. Participants from Study 2 were reclassified into two clusters using k-means clustering based on their Big Five personality scores. After excluding outliers and overlapping samples, the final participant count was 90. Distances from each participant to the cluster means on each personality trait served as independent variables, representing each participant's alignment with one of two blended personality profiles.

Data Collection

We recruited participants already profiled in previous studies, ensuring a known personality baseline before the study began. From the initial clusters derived, after excluding outliers and those unwilling or unable to participate, 82 participants remained. The demographic profile included an average age of 37.2 years (SD = 13.2), with 50% identifying as female.

Procedures and Measurements

Study 3a

We constructed two sets of advertisements, each featuring an iPhone image and accompanying text. Each set was targeted at one of the two personality clusters. We used GPT-4 to generate tailored text for each cluster, with one version reflecting the target cluster's personality blend and the other representing the reverse. Ads were crafted by instructing GPT-4 to generate texts that would appeal to a specific set of trait scores, for example, "19 on Extraversion, 24 on Agreeableness, etc., with a maximum score of 42." Each ad version was assessed using two 11-point bipolar scales, with participants rating on perceived relevance, credibility, favorability, persuasiveness, product interest, and ad click likelihood.

Linear regression analyses were conducted to compare responses across clusters, adjusting for sociodemographic factors, followed by hypothesis testing as per the proposed conceptual model in Figure 1, addressing hypothesis H7.

Study 3b

We constructed two scenarios and two versions of common online chat conversations with a chatbot agent on an e-commerce website, following the same methodological approach as in Study 3a. For each conversation scenario, such as “Can you help me track my order?” and “Do you have this shirt in other colors?”, we directed GPT-4 to craft responses considering a customer's blended scores on the Big Five personality traits, e.g., "19 on Extraversion, 24 on Agreeableness, 26 for Conscientiousness, 29 for Neuroticism, 14 for Openness with the maximum possible score for each trait being 42". This tailored approach aimed to personalize the chatbot’s responses to match the distinct personality profiles of the clusters and their reverse scores.

Participants evaluated each chatbot conversation using the same assessment items as in Study 2b, which included ratings on perceived relevance, trust, satisfaction, positive emotion, and possible increased duration of engaging with the AI chatbot. We analyzed these responses using linear regression to assess the impact of personality-tailored communication on user engagement and satisfaction, paralleling the analyses conducted in Study 3a. This comprehensive evaluation aimed to validate our conceptual model for LLMs’ effectiveness in delivering personalized chatbot conversations to address hypothesis H9.