People's Perceptions (and Bias) Toward Creative Content generated by AI (ChatGPT-4), Human Experts, and Human-AI Collaboration

Yunhao (Jerry) Zhang, MIT Sloan

Renee Gosline, MIT Sloan

Abstract

With the wide availability of Large Language Models (e.g., ChatGPT-4) nowadays, there are four primary paradigms for Human-AI collaboration: human only, AI only, a human making the final decision with AI output as reference, or an AI making the final decision with human input as reference. In partnership with one of the world's leading consulting firms (over $175 billion USD market capitalization), we enlist professional content creators and ChatGPT-4 to create advertising content for products and persuasive content for campaigns following the paradigms outlined above. When participants were asked to evaluate the quality of the content without any knowledge of the creation paradigm (i.e., they are not informed of AI's potential involvement at all), we find that content generated by ChatGPT-4 making the sole or final decision on the content is perceived as higher quality. This result remains true when participants are partially informed of the content creation process (e.g., they are briefed about the four potential content creation paradigms upfront, but they do not know exactly how each piece of content they evaluate is created.) However, the performance gap between human and AI becomes narrower when participants are informed of the creation paradigm for each type of content; this bias is mostly driven by human favoritism rather than AI aversion. Our findings suggest that knowing the same content is created by a human expert increases its perceived quality, but knowing that AI is involved in the creation process does not reduce its perceived quality. However, even considering this human favoritism, the professional content creators do not outperform ChatGPT-4. We also examine how the results vary with the task categories.

1. Introduction

As Large Language Models such as OpenAI's ChatGPT become increasingly accessible, Generative Artificial Intelligence (GAI) is bound to revolutionize the way human beings work and live. Nevertheless, disruptive technologies are often double-edged swords. Academia and the public have expressed concerns about this new technology (Botha and Pieterse 2020, Clayton 2023, Haupt and Marks 2023, Khan 2023, Li et al. 2023). On the other hand, some have compared GAI to the "printing press", which has the potential to diffuse knowledge, power, and learning widely. Recent research has shown that GAI could enhance labor productivity, for instance, in customer communication (Brynjolfsson et al. 2023) or essay writing (Noy and Zhang 2023). While the cited studies focus on examining the impact of LLMs (i.e., Generative Pre-trained Transformer, or GPT) on the *workers*, our research, to our knowledge, is the first to thoroughly study *consumers'* perception of LLMs. We hereby describe our novel contributions compared to previous relevant research. First, in addition to content generated solely by humans or solely by AI, existing research on people's perception of content generated by LLMs has not examined content produced by human-AI interactions. For example, Ayers et al. 2023 examined how health care professionals would evaluate responses to medical questions generated by physicians vs ChatGPT in an anonymized setting and found that responses generated by ChatGPT received higher quality ratings. Nevertheless, in real life, the responses could be generated by human-AI interactions, or human-in-the-loop. For example, content generators may first obtain a response from ChatGPT as reference before making their final decision (e.g., "augmented human"), or they may enter their response as an input to ChatGPT along with the question prompt, letting ChatGPT edit their response and make the final decision ("augmented AI"). Our research compares the quality of creative content (i.e., advertising content for products and persuasive content for campaigns, details described in Section 2) generated under all four paradigms: (1) Human Expert only (i.e., professional content creators from one of the world's leading consulting firms create the content individually), (2) AI (ChatGPT-4) only, (3) Augmented Human (i.e., a human expert makes the final decision on the output but is given the content first generated by ChatGPT-4 for the task, which they may edit or use as an inspiration), and (4) Augmented AI (i.e., ChatGPT-4 makes the final decision on the output but is given the content first generated by a human expert, which it may edit or use as an inspiration).

Furthermore, our set-up also allows us to shed light on the debate about whether humans or AI should make the final decision in our creative content generation context (McKendrick and Thurai 2022).

We recruit and randomly assign online participants to rate the content quality in different conditions. In the baseline condition, participants are completely unaware of the content generation paradigms at all, thus basing their evaluations solely on the textual output (e.g., they are presented with the content without any mention of humans or AI throughout the study). Our findings show content which ChatGPT-4 solely or ultimately determines the output, is perceived as higher quality. This result is robust in the "uninformed" condition when we add the contextual influence such that participants are informed of the content generation paradigms at the beginning of the survey, but they are unaware of the specific process for each piece.

Additionally, our research also examines consumers' bias towards the content generation paradigms (i.e., given the same piece of content, whether knowing its creator affects people's evaluation). Liu et al. 2022 examined a similar question and found that – when writing emails to console others – the recipients display aversion towards the senders who use AI to write the message. Nevertheless, their study did not involve any emails actually generated by AI, but deceptively revealed and varied the human-generated messages to be either human-generated or AI-mediated. In other words, since content generated by humans and LLMs may have inherently different features and quality, their research is not about how people perceive content generated by LLMs, but how people perceive human-generated content being framed as AI-mediated. Our research examines potential bias towards human experts, AI, or human-AI collaboration without deception. About one-third of the recruited participants are randomly assigned to evaluate content quality in the "informed" condition, in which they are not only informed upfront about the four paradigms as in the "uninformed" condition, they are also explicitly told under which paradigm a piece of content is generated when they evaluate the content. By comparing the baseline and the "informed" condition, we find that people perceive the same piece of content generated solely by a human expert as higher quality if they are aware the content is generated solely by a human expert. However, we do not find this effect among the other three paradigms. Furthermore, this phenomenon is robust when we compare the "uninformed" condition with the

"informed" condition. Therefore, contrary to the prediction by the "algorithm aversion" literature (e.g., Dietvorst et al. 2015, Catello et al. 2019, etc.) that people display aversion towards AI, we do not find aversion (in terms of various measures of quality) towards AI's involvement in generating creative content. Rather, we have evidence for human favoritism. However, even with this bias, content generated by ChatGPT-4 – when it makes the sole or final decision on the output – is still perceived as on par or better than human-generated content. Nevertheless, our results by no means suggest LLMs should completely replace human agents – especially human oversight, which we discuss in Section 5 of our paper.

The remainder of the paper is structured as follows: Section 2 describes how the content used for the study is generated; Section 3 describes the experiment, in which we recruit participants to evaluate content quality; Section 4 describes the results of the experiment; Section 5 discusses the implications of our results and concludes the paper.

2. Content Generation Process

In this section, we describe in detail how the content used for studies is generated. First, we pre-select five retail products (e.g., air fryer, projector, electric bike, emergency kit, and tumbler) from a retail website and five campaign goals that are uncontroversially benign (e.g., stop racism, do more physical exercises, wash hands more often, and eat less junk food). Readers may find screenshots of the instructions given to the professional human content generators in Supplementary Information Section 1. All content generated under the four paradigms is included in SI Section 2.

2.1 Content Generated by Human Expert Individually and Human Expert with Access to Content First Generated by ChatGPT-4

We enlist ten professional content creators from one of the world's best consulting firms to generate advertising content for the products and persuasive content for the campaigns. This research context offers unique advantages that facilitate the exploration of our research questions. First, the ten content creators are industry professionals who have experience with

writing advertising content for corporate products and campaign messages for NGOs.[1] One of our research questions is to compare the quality of content generated by human experts and AI. Engaging top-tier professionals for this purpose grants our study a significant degree of external validity, more so than if we were to recruit subjects from platforms like MTurk or Upwork.[2] Second, according to our liaisons from the firm, these content creators took the tasks very seriously. The department from which we sourced these creators was aware of their higher-level managers' interest in this study, instilling an intrinsic motivation to excel in the task. Moreover, participation was strictly voluntary, ensuring that those involved were genuinely interested in contributing to our research. Our liaisons also facilitated a one-hour meeting with the ten content creators during regular working hours to distribute the tasks via Qualtrics links. These links remained accessible for one day, allowing the creators ample time to engage with the tasks.[3]

Each content creator completed two content generation tasks – one advertising content for one of the five products, and one persuasive content for one of the five campaigns. For the campaign tasks, the experts were provided with the following instructions: "*Your task is to write persuasive content for a campaign in fewer than 100 words. Your goal is to persuade people to change their behavior after seeing your content.*" After each expert had been assigned a campaign, they were asked "*please create persuasive content for a client (e.g., an NGO doing a campaign) to convince people to [perform the action advocated by the campaign] in fewer than 100 words.*"

For the advertising content task, the experts were provided with these instructions: "*The task is to write advertising content for a product in fewer than 100 words without relying on LLMs. Your goal is to get people interested in this product after seeing your content.*" After each expert had been assigned a product, they were given a picture of the product along with a description of the product's features sourced directly from the product description section of a retail website. The experts were then prompted "*please create advertising content for the above product in fewer than 100 words.*" The 100-word limit was chosen as it approximates the amount of text that can be effectively communicated within a 30-second span.

---

[1] When deciding on the task prompts, we worked with our liaisons to ensure the participating content creators had past experience and expertise in this type of tasks.

[2] Anecdotally, we asked our liaison about the cost of recruiting these experts to do the tasks without the partnership. The manager responded "probably not affordable."

[3] This was done also for the purpose of maintaining their anonymity.

Very importantly, the first task the experts completed, be it about the product or the campaign, required the content creators to perform the task without using LLMs.[4] This resembled the typical human content generation paradigm before LLMs became popular. For the second task, however, the content generators were shown the content generated (solely) by ChatGPT-4. Before beginning their second task, they were informed "*we will provide you with the textual content generated by ChatGPT-4, which you may use either as inspiration for your own content or as a first draft that you can edit upon.*

In order to obtain the persuasive / advertising content generated by ChatGPT-4, we provided ChatGPT-4 with a prompt on the campaign topic, and asked it to create persuasive content in fewer than 100 words / we provided ChatGPT-4 with a prompt for the product to be advertised (including the product's name and feature descriptions), and asked it to create an advertisement in fewer than 100 words.

Both the input prompt and the output by ChatGPT-4 will be available to you on the next page."

Then in addition to the aforementioned prompt for the first task, the experts are also presented with content generated solely by ChatGPT-4. The content produced during the second task will be referred to as the "augmented human" content generation paradigm in our paper, distinguishing it from the "human-only" paradigm. This is because, although the final output was determined by the humans, they were assisted by AI.

Overall, we obtained twenty pieces of content generated by ten professional content creators – ten tasks (i.e., five products + five campaigns) multiplied by two content per person (i.e., one with the assistance of ChatGPT-4, and one without.)

2.2 Content Generated by ChatGPT-4 solely and ChatGPT-4 with access to content first generated by a Human Expert

---

[4] This was made explicitly clear by the instructions in our survey as well as our liaison when they held an internal meeting during which they were randomly assigned the tasks.

We paid a $20 monthly subscription fee to obtain access to ChatGPT-4. While each human expert generated two pieces of content, ChatGPT-4 generated twenty pieces of content (ten without access to the content first generated by a human expert, and ten with.) We presented ChatGPT-4 with prompts nearly identical to those given to the human experts. The prompt for the campaign was as follows: "*please create persuasive content for a client (e.g., an NGO doing a campaign) to convince people to [perform the action advocated by the campaign] in fewer than 100 words.*"

The prompt for the products was "*please create advertising content to get people interested in the following product in fewer than 100 words. The product is [the product's name]. For your information, the product has the following features: [the product's features taken from the retail website]."* The outputs generated following these prompts were the AI-generated content we gave to the human experts in their second task as mentioned previously.

When we also gave ChatGPT-4 the content generated solely by a human expert for the same campaign or product, the prompt had the following additional paragraph: "*Below is the advertising / persuasive content generated by a professional content creator from one of the world's best consulting firms, which you may use as inspiration or a first draft you edit upon when generating your own content. '[The content generated by a human expert without access to content generated by ChatGPT-4.]'* " We will refer to the content generated in response to this type of prompt as the "augmented AI" content generation paradigm in our paper. This term distinguishes it from the "AI-only" paradigm because, even though the AI made the final output decision, it had human expert assistance.

3. Methods

3.1 Study Design.

The study protocol and all main analysis are pre-registered at https://aspredicted.org/2qg6y.pdf. Readers may find screenshots of the instructions given to the online participants of this study in SI Section 3.

Participants were randomly assigned to one of the three different conditions: the baseline condition, the "uninformed" condition, and the "informed" condition. In the baseline condition, participants evaluate content quality completely ignorant of the context, meaning they do have any knowledge of the content creation paradigms (i.e., there is no mention of humans, AI, or human-AI collaboration throughout the study. Their judgment of content quality should be solely based on the textual output). In this condition, we are able to measure the quality of the content created under each paradigm without any contextual influence.

In the "uninformed" condition, participants rate the content quality with partial knowledge of the content creation paradigms. Participants in this condition are briefed about the four content creation paradigms at the beginning of the survey (e.g., they are informed of the definition of the content generation paradigms and each content they evaluate is generated under one of the four paradigms), but they do not know *exactly* how each piece of content they evaluate is created . This condition resembles real-life situations such as when people are reading a news article. This condition mirrors real-life situations such as reading a news article where readers might be aware that some articles are AI-generated, but they can't be sure whether the article they're reading is human, AI, or human-AI produced. This condition allowed us to measure how people assess content when AI might be involved in the content generation process.

In the "informed" condition, participants rate each content's quality with full knowledge of the content creation paradigms. Not only are participants briefed about the content generation paradigm participants as those in the "uninformed" condition, they also know exactly how each piece of content they evaluate is created. This condition enabled us to determine whether any potential bias towards a specific content generation paradigm exists. By comparing perceived content quality between the baseline and 'informed' conditions, we could examine whether awareness (or lack thereof) of the content generation paradigm for a given piece of content affects people's quality evaluation. For instance, suppose that given the same piece of content produced solely by a human expert, participants perceive the content to be of higher quality when they are aware that it was solely human-generated compared to when they're not aware.

This would clearly indicate a bias in favor of content created under the 'human-only' paradigm. In addition to the comparison with the baseline for all four paradigms, we also compared the 'informed' with the 'uninformed' condition.

3.2 Participants.

All participants were recruited from the research panel platform CloudResearch Connect. A total of 1210 participants entered our survey, and 9 participants failed an attention check and were not allowed to finish the survey. All remaining 1201 participants (50% female, M_age = 38) who finished the survey are included for analyses. The median time to complete the survey was 10.3 minutes. The survey completion fee was $1.5.

3.3 Procedures.

After signing the consent form, all participants initially received the same survey overview. It stated, "*we have generated some advertising content for five different products and some persuasive content for five different campaigns. We want you to evaluate the quality of the text you will read.*" Participants in the "uninformed" and "informed" conditions were further informed about the four content generation paradigms. Those in the baseline condition skipped this step, thus remaining entirely unaware of the content generation paradigms (see SI Section 3 "Page 2 of the instructions" for example). After passing the first attention check, all participants were informed that they would first evaluate the quality of five advertising content for five different products and given a description of the key outcome measures:

(Satisfaction) "*suppose you are the seller of the product, to what extent are you satisfied or unsatisfied with the generated advertising content on a scale from 1 to 7*" with 1 being very unsatisfied and 7 being very satisfied;

(Willingness-to-pay) "*suppose you are the seller of the product and the content has a copyright, what is the maximum amount of money you are willing to pay to use the content as your advertisement? (Assuming you have a budget of $1000, please state your max willingness-to-pay between $1 and $1000)*";

(Interest) "*to what extent you are interested in learning more about the product (e.g., its price, user reviews, complete product descriptions, etc.) on a scale from 1 to 7*" with 1 being not interested at all and 7 being very interested.

Participants then continued to rate the content quality for the five products. They were given a picture of each product and the piece of advertising content generated by one of the four paradigms for the product when they provided responses to the three key outcome measures. Participants in the "informed" condition were additionally informed of under which paradigm the content they were viewing was generated, but participants in the baseline or the "uninformed" condition were not. After evaluating the advertising content, participants needed to pass another attention check. Then they were told that they would next evaluate the quality of five persuasive content for five different campaigns. They were again given a description of the key outcome measures. While the first two outcome measures were the same, the third outcome measure was as follows: "*to what extent you are convinced by the above content to [perform the action advocated by the campaign] on a scale from 1 to 7*" with 1 being not convinced at all and 7 being very convinced. They were given the piece of persuasive content generated by one of the four paradigms for the campaign when they provided responses to the three key outcome measures. Same as before, participants in the "informed" condition were additionally made aware of under which paradigm the content they were viewing was generated, but participants in the baseline or the "uninformed" condition were not. Participants then answered some demographic questions before finishing the study.

To further clarify, we had a between-subjects 3 by 4 design: 3 conditions (baseline, uninformed, informed) and 4 paradigms. Within each of the three conditions, every participant provided quality evaluations for content produced under one of the four paradigms for each of the ten tasks (i.e., five products plus five campaigns). Therefore, with approximately 1200 total participants, we obtained around 100 evaluations for content generated under each paradigm in each of the three conditions.
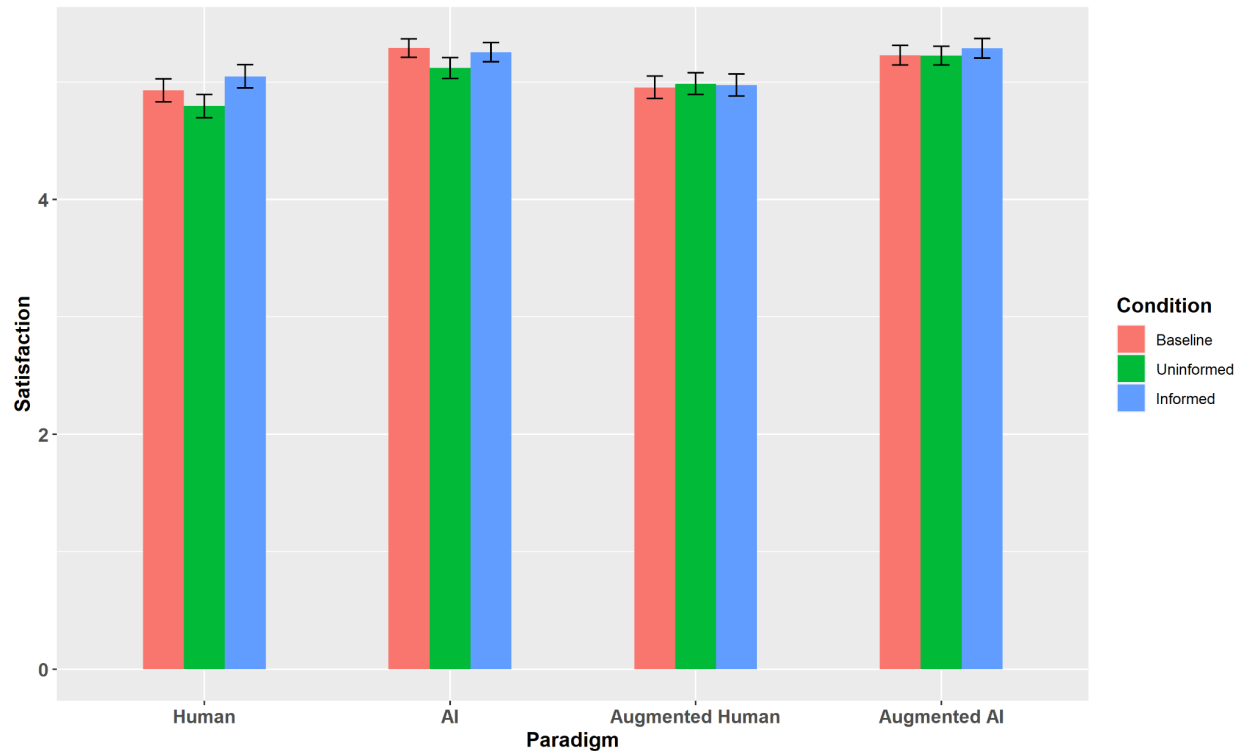
4. Results

Figure 1 *The x-axis is the content generation paradigm: human expert only, AI only, a human expert who finalizes the content first generated by AI, and an AI that finalizes the content first generated by a human expert. The y-axis is the subjects' level of satisfaction pooling all ten contents together for each paradigm. The colors represent the different conditions. The bars indicate 95% confidence intervals.*
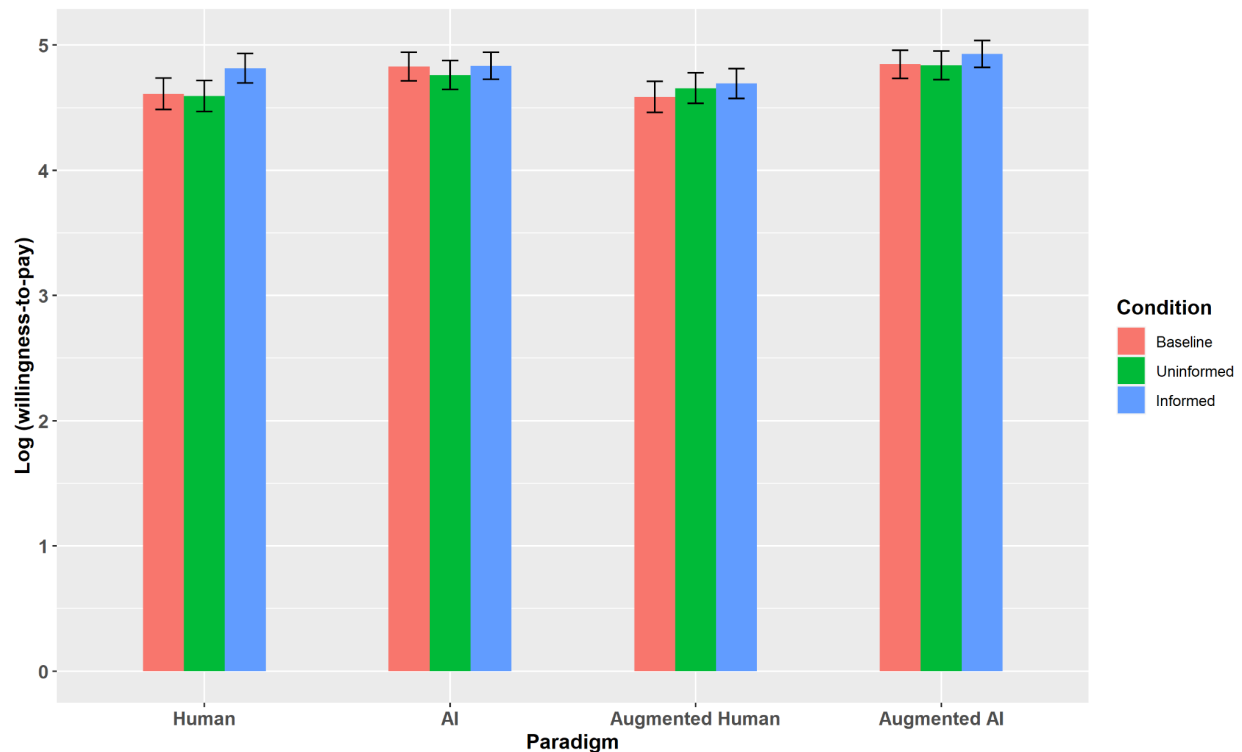
Figure 2 The x-axis is the content generation paradigm: human expert only, AI only, a human expert who finalizes the content first generated by AI, and an AI that finalizes the content first generated by a human expert. The y-axis is the logarithm of subjects' stated willingness-to-pay for the content (pooling all ten contents together for each paradigm). The colors represent the different conditions. The bars indicate 95% confidence intervals.

4.1 Purely Evaluating Content Quality

We first examine participants' stated level of satisfaction for the content generated under different paradigms when their focus is solely on the content quality. As depicted by the baseline condition in Figure 1, participants expressed varying degrees of satisfaction with content produced under different paradigms (one-way ANOVA, p-value = 0.000). Furthermore, on average contents generated solely by ChatGPT-4 resulted in the highest satisfaction level, and it is on par with when ChatGPT-4 finalizes the content first generated by a human expert (5.29 vs 5.23, two-sample t-test, p-value = 0.3). On the other hand, content generated by a human expert resulted in a similar level of satisfaction as content generated by a human expert with access to content first generated by ChatGPT-4 (4.93 vs 4.95, two-sample t-test, p-value = 0.72). Interestingly, content generated when AI makes the final decision on the output resulted in higher satisfaction level compared to contents generated when a human expert makes the final

decision on the output (AI only vs Human Expert Only: 5.29 vs 4.93, two-sample t-test, p-value = 0.000, cohen's d = 0.25; Augmented AI vs Augmented Human Expert: 5.23 vs 4.95, two-sample t-test, p-value = 0.000, cohen's d = 0.19).

Similar patterns emerge when we use participants' willingness-to-pay for the content as a measure of content quality. As depicted by the baseline condition in Figure 2, participants had varying willingness-to-pay for content generated under different paradigms (one-way ANOVA, p-value = 0.001). Furthermore, on average participants' willingness-to-pay was almost the same for content generated solely by AI or content generated when AI finalizes the content first generated by a human expert, (4.83 vs 4.85, two-sample t-test, p-value = 0.83). On the other hand, content generated by human experts and content generated by human experts with access to content first generated by AI had similar willingness-to-pay (4.61 vs 4.59, two-sample t-test, p-value = 0.79). Consistent with the previous results, content generated when AI makes the final decision on the output resulted in higher willingness-to-pay compared to content generated when a human expert makes the final decision on the output (AI only vs Human Expert Only: 4.83 vs 4.61, two-sample t-test, p-value = 0.01, cohen's d = 0.11; Augmented AI vs Augmented Human Expert: 4.85 vs 4.59, two-sample t-test, p-value = 0.003, cohen's d = 0.13). Our results suggest that although the size of the difference is not large, AI seems to have the upper-hand in producing creative content with higher quality.

4.2 Evaluating Content Quality with Partial Knowledge of AI's Potential Involvement

In the real world, while human-to-human interaction remains the norm, people might sometimes wonder whether they are interacting with a human or an AI agent (e.g., could this message or news article have been written by AI?). Therefore, we examine participants' stated level of satisfaction when they are aware that the AI could *potentially* be involved in the content generation process. The crucial difference between the baseline condition and this "uninformed" condition is that participants in the baseline are completely ignorant of AI's potential involvement in content generation, hence the identity of the content creators is unlikely to be a factor affecting their judgment. However, since participants are not informed of how exactly each content they see is generated, we hereby examine the contextual effect of potential AI involvement on their evaluation.

As depicted by the "uninformed" condition in Figure 1, participants had different levels of satisfaction for content generated under different paradigms (one-way ANOVA, p-value = 0.000). On average, content generated by ChatGPT-4 finalizing the content first generated by a human expert resulted in the highest satisfaction level, and it is on par with when AI solely generated the content (5.23 vs 5.12, two-sample t-test, p-value = 0.08). On the other hand, content generated by human experts finalizing content first generated by AI was better than content generated solely by a human expert (4.99 vs 4.80, two-sample t-test, p-value = 0.006, cohen's d = 0.12). Similar to the baseline condition, content generated when AI makes the final decision on the output resulted in higher satisfaction level compared to content generated when a human expert makes the final decision on the output (AI only vs Human Expert Only: 5.12 vs 4.80, two-sample t-test, p-value = 0.000, cohen's d = 0.21; Augmented AI vs Augmented Human Expert: 5.22 vs 4.99, two-sample t-test, p-value = 0.0001, cohen's d = 0.17).

Furthermore, as depicted by the "uninformed" condition in Figure 2, participants had different willingness-to-pay for content generated under different paradigms (one-way ANOVA, p-value = 0.02). The willingness-to-pay for content is nearly identical when the final decision was made by a human expert, regardless of whether they had access to AI generated content (Human Expert Only vs Augmented Human Expert: 4.59 vs 4.66, two-sample t-test, p-value = 0.47). Similarly, the willingness-to-pay is almost the same for content generated when AI made the final decision on the output, irrespective of whether it had access to the content generated by a human expert (AI Only vs Augmented AI: 4.76 vs 4.84, two-sample t-test, p-value = 0.36). However, the willingness-to-pay for content generated when AI made the final decision on the output is slightly higher (AI only vs Human Expert Only: 4.76 vs 4.59, two-sample t-test, p-value = 0.054, cohen's d = 0.09; Augmented AI vs Augmented Human Expert: 4.84 vs 4.66, two-sample t-test, p-value = 0.03, cohen's d = 0.09).

4.3 Is There Any Bias towards any Content Creation Paradigm?

Baseline vs Informed

We next explore whether individuals show any bias toward any of the content generation paradigms by comparing how participants' subjective content quality evaluations differ when

they are fully aware (the 'informed' condition) versus completely ignorant (i.e., the baseline condition) of the content generation process. We will have evidence for bias if, for example, given the same piece of content, people express greater or less satisfaction or willingness-to-pay if they are informed the content is generated by a human expert. The results are illustrated by a comparison between the "pinkish" bar (baseline) and the "bluish" bar (informed) within each paradigm in Figure 1 and Figure 2. To analyze this more systematically, within each of the four paradigms, we (separately) regress the dependent variable (i.e., level of satisfaction or willingness-to-pay) on the dummy variable indicating the condition (1 = informed, 0 = baseline) with task fixed effects. A positive (negative) coefficient indicates favoritism (aversion) towards a particular content generation paradigm.

First, we find that given the same content generated solely by a human expert, participants felt more satisfied with the content and were willing to pay more if they were informed of the identity of the content creator ($b = 0.09$, p-value = 0.003). Similarly, participants were willing to pay more if they were informed that the content was generated by a human expert ($b = 0.18$, p-value = 0.0000). However, we do not find any effect on the satisfaction level or willingness-to-pay for the other content generation paradigms (AI: $b\_satisfaction = -0.05$, p-value = 0.35, $b\_wtp = -0.01$, p-value = 0.87; Augmented Human: $b\_satisfaction = -0.004$, p-value = 0.91, $b\_wtp = 0.1$, p-value = 0.23; Augmented AI: $b\_satisfaction = 0.08$, p-value = 0.12, $b\_wtp = 0.12$, p-value = 0.10). Therefore, we do not have any evidence for aversion towards AI or the involvement of AI. Instead, we have evidence for human favoritism in our context.[5]

4.4 Persuasive Content for Campaigns vs Advertising Content for Products

---

[5] See SI Section 4 for a comparison of the uninformed and the informed condition.
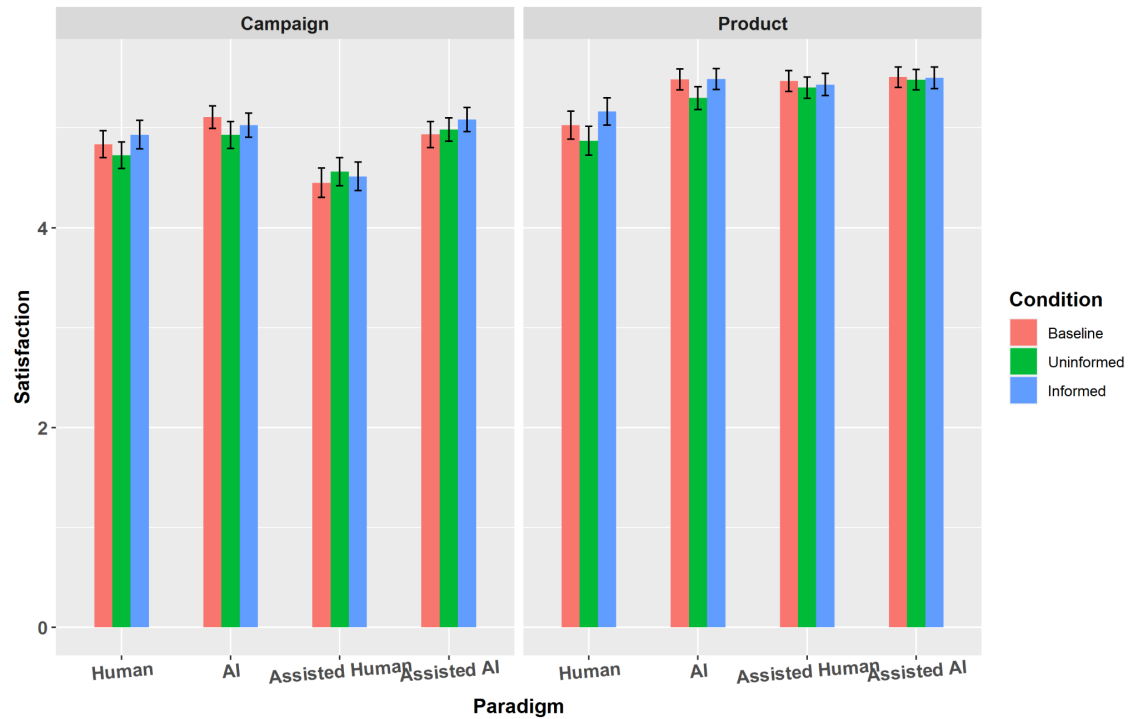
Figure 3 *The x-axis is the content generation paradigm. The y-axis is the subjects' level of satisfaction pooling the five contents together for each paradigm given a task category. The left panel depicts persuasive contents generated for five campaigns, and the right panel depicts advertising contents generated for five products. The colors represent the different conditions. The bars indicate 95% confidence intervals.*
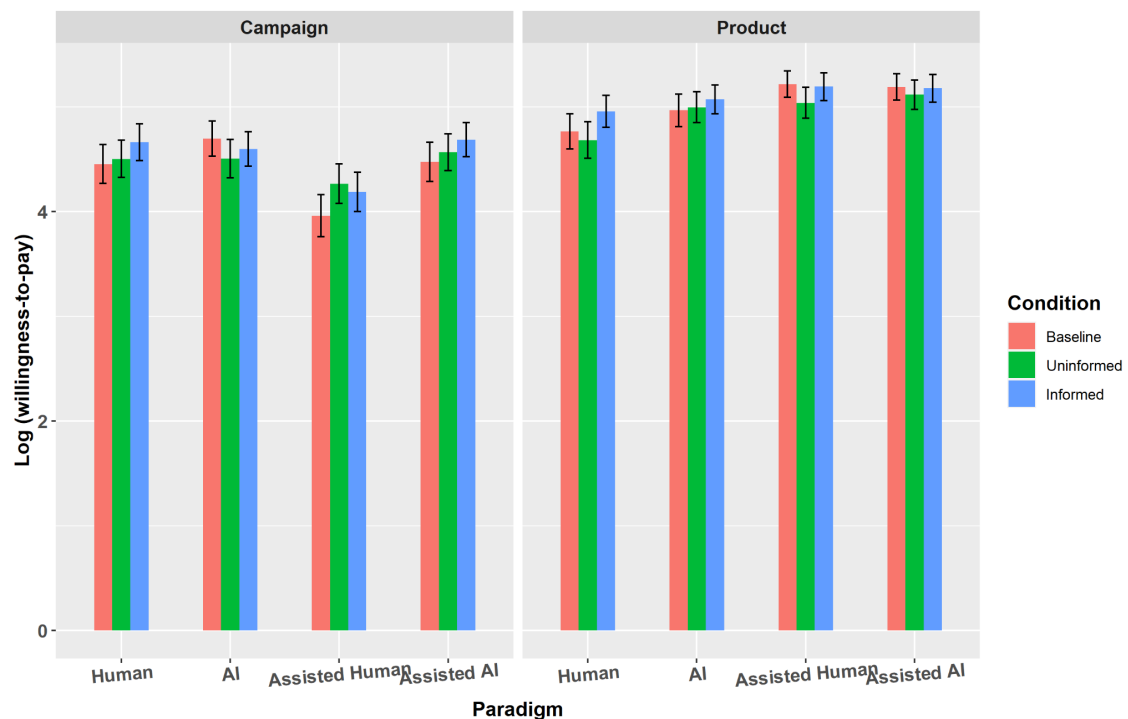
Figure 4 The x-axis is the content generation paradigm. The y-axis is the logarithm of subjects' willingness-to-pay pooling the five contents together for each paradigm given a task category. The left panel depicts persuasive contents generated for five campaigns, and the right panel depicts advertising contents generated for five products. The colors represent the different conditions. The bars indicate 95% confidence intervals.
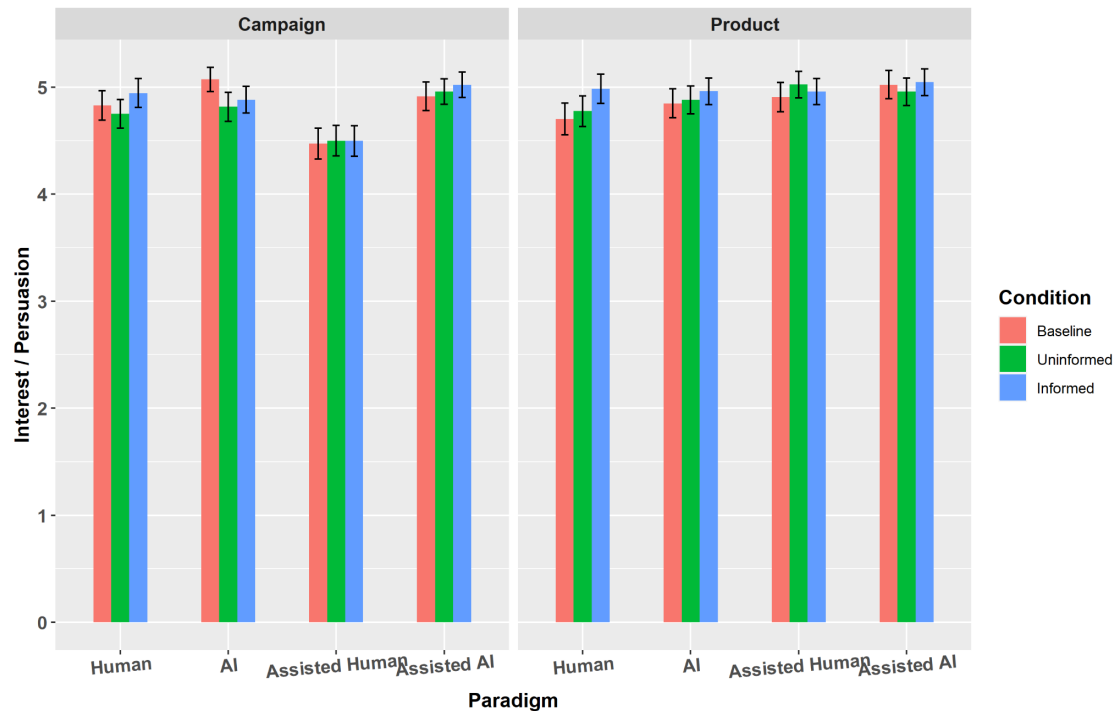
Figure 5 The x-axis is the content generation paradigm. The y-axis for the left panel (pooling persuasive contents generated for five campaigns) is the extent to which participants are persuaded by the persuasive content. The y-axis for the right panel (pooling advertising contents generated for five products) is the extent to which participants are interested in learning more about the product after seeing the advertising content. The colors represent the different conditions. The bars indicate 95% confidence intervals.

One might wonder if the primary results differ when we separately analyze persuasive content for campaigns and advertising content for products. While persuasive content for campaigns might necessitate a deeper understanding of human psychology and more creativity, advertising content in our case tends to be more standardized, given that a significant portion of the text consists of product feature descriptions.

In the baseline and the "uninformed" condition for both task categories, participants either felt at least as satisfied, or even more satisfied, with the content generated with AI's involvement than

without. They were either willing to pay an equivalent amount, or more, for the content generated, and they became either equally interested or more interested in the product or persuaded to support the campaign when AI made the sole or final decision on the output. (The results are obtained by comparing the "pinkish" bars across the paradigms in Figure 3 through Figure 5, see SI Section 5 for detailed statistics).

Interestingly, in the baseline condition, although the performance gap between human experts and AI is similar between the two categories, the gap between augmented human experts and augmented AI is smaller in content generated for products than for campaigns. For example, we separately regress the dependent variable (level of satisfaction or log willingness-to-pay) on the content generation paradigm (0 = human, 1 = AI; or 0 = augmented human, 1 = augmented AI), category (0 = campaign, 1 = product), and their interaction, with task fixed effects. The negative coefficient of the interaction term suggests a reduction in the gap in level of satisfaction and willingness-to-pay between augmented human experts and augmented AI in the product category relative to the campaign category ($b\_satisfaction$ = -0.44, p-value = 0.04; b_wtp = -0.53, p-value = 0.02). Furthermore, this effect is driven by an increase in perceived quality of content generated by the "augmented human" paradigm in the product category, rather than a decrease in the "augmented AI" paradigm. Nevertheless, although the interaction is not statistically significant, the performance gap between augmented human experts and augmented AI is smaller in content generated for advertising content for products than for campaigns in the "uninformed" condition ($b\_satisfaction$ = -0.35, p-value = 0.10; b_wtp = -0.24, p-value = 0.26).

Last but not least, we observe favoritism towards content generated solely by human experts for both task categories, but do not find any aversion towards content generated with AI's involvement. In addition, the level of human favoritism is the same for the two task categories[6] (detailed statistics are provided in SI Section 6). To summarize the results in this section, although the primary outcomes are qualitatively similar when we examine the two task categories separately, we do observe that the performance gap between the 'augmented human' and 'augmented AI' paradigms is smaller when they are creating content for products. This is due

---

[6] Within a content generation paradigm, we regress the DVs on the condition (1 = informed, 0 = baseline), task category (1 = product, 0 = campaign), and their interaction, with task fixed effects. The coefficient of the interaction term for satisfaction and log willingness-to-pay is 0.02 (p-value = 0.76) and -0.04 (p-value = 0.63)

to the fact that human experts are more likely to adopt the texts written by AI in this case because a large fraction of the text is standard product description.

5. Discussion and Conclusion

Although our results suggest ChatGPT-4 outperforms human experts in generating advertising content for products and persuasive content for campaigns, and it has the potential to reduce human labor on standard creative content generation, we by no means suggest GAI should completely displace human workers, especially human oversight. For example, Bai et al. 2023 suggests AI can persuade humans on political issues. In our contexts, we carefully choose the products and campaigns to be harmless. Nevertheless, human oversight is still needed to ensure the content produced by GAI is appropriate in more sensitive topics, and inappropriate content is never distributed. Furthermore, our study does not examine non-textual creative content (e.g., graphical and audio), which are also popular means of communication. The performance between human experts and GAI in these domains remains to be explored by future research. Nevertheless, our results indeed serve as evidence that GAI can benefit capital owners and consumers by raising productivity (e.g., it takes ChatGPT-4 a matter of seconds to produce the content of on par or higher quality than the human experts in our context) and lowering prices (e.g., the monthly subscription fee for ChatGPT-4 is $20) (Acemoglu and Restrepo 2018, Agrawal et al. 2019, Acemoglu and Restrepo 2020).

Our result also contributes to the discussion on algorithm aversion (Dietvorst et al. 2015) vs algorithm appreciation (Logg et al. 2019) in the domain of GAI. Instead of algorithm aversion, we demonstrate human favoritism as a form of bias – simply knowing a piece of content being generated by human experts increases the perceived quality of the content. However, we do not find evidence of algorithm aversion in our context (i.e., knowing a piece of content being generated with AI's involvement does not lower the perceived quality of the content). This result is somewhat surprising given that Castello et al. 2019 clearly shows that people display aversion towards AI in subjective task contexts such as evaluation of creative content (e.g., evaluating joke funniness). This is probably because GAI is able to demonstrate its superior capabilities as

participants have the opportunity to see how good the content is (Zhang and Gosline 2022). To our knowledge, our research is the first to document consumers' perception of creative content generated by industry professionals and LLMs and people's bias (favoritism) toward content generated solely by human experts. Future research could further investigate the performance and ethics of LLMs, and refine the human-in-the-loop protocol.

Reference

Acemoglu, Daron and Pascual Restrepo, "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," American Economic Review, 2018, 108 (6), 1488–1542.

Acemoglu, Daron and Pascual Restrepo , "Robots and Jobs: Evidence from US Labor Markets," Journal of Political Economy, 2020, 128 (6), 2188–2244.

Agrawal, Ajay, Joshua S Gans, and Avi Goldfarb, "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction," Journal of Economic Perspectives, 2019, 33 (2), 31–50.

Ayers, John W., et al. "Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum." JAMA Internal Medicine (2023).

Botha, Johnny, and Heloise Pieterse. "Fake news and deepfakes: A dangerous threat for 21st century information security." ICCWS 2020 15th International Conference on Cyber Warfare and Security. Academic Conferences and publishing limited. 2020.

Clayton, James. "Sam Altman: CEO of OpenAI calls for US to regulate artificial intelligence." BBC News, 17 May 2023, https://www.bbc.com/news/world-us-canada-65616866, Accessed May 2023.

Haupt, Claudia E., and Mason Marks. "AI-Generated Medical Advice—GPT and Beyond." JAMA 329.16 (2023): 1349-1350.

Khan, Gretel. "Will AI-generated images create a new crisis for fact-checkers? Experts are not so sure." Reuters Institute, 11 April 2023, https://reutersinstitute.politics.ox.ac.uk/news/will-ai-generated-images-create-new-crisis-fact-checkers-experts-are-not-so-sure, Accessed May 2023.

Li, Ron, Andre Kumar, and Jonathan H. Chen. "How Chatbots and Large Language Model Artificial Intelligence Systems Will Reshape Modern Medicine: Fountain of Creativity or Pandora's Box?." JAMA Internal Medicine (2023).

Liu, Yihe, et al. "Will AI console me when I lose my pet? Understanding perceptions of AI-mediated Email writing." Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022.

McKendrick, J., and A. Thurai. "AI Isn't Ready to Make Unsupervised Decisions." Harvard Business Review. https://hbr. org/2022/09/ai-isnt-ready-to-make-unsupervised-decisions (2022).

Noy, Shakked and Whitney Zhang, "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," Available at SSRN 4375283, 2023.

Zhang, Yunhao and Gosline, Renée, Understanding Algorithm Aversion: When Do People Abandon AI After Seeing It Err? (December 9, 2022). MIT Sloan Research Paper No. 6846-22, Available at SSRN: https://ssrn.com/abstract=4299576 or http://dx.doi.org/10.2139/ssrn.4299576

Supplementary Information

1. Instructions for the Content Creators to Create the Content for the Main Study

Q821

Overview

ChatGPT-4 is the state-of-art large language model (LLM) developed by OpenAI. We are partnering with Accenture to study various aspects of human-LLM collaborations (for example, the potential impact of incorporating AI input into the content generation process).

As expert content generators, your experience is paramount. This is an academic study and you are completing the first part (not meant for actual marketing distribution to the public), in which you will be asked to generate some advertising contents for a product or some persuasive contents for a campaign. Though you will not be participating in the second part of the academic study, your content will be included as representative text to study participants. No identifying information will be asked or shared.

You will see further task instructions on the next page.

○ I have carefully reviewed the above information

--- Page Break --

Q893

We would like you to complete two tasks.

In the first task, we want you to create advertising content for a product without relying on a large language model. You may search for information or inspirations online or offline, **but please do not use language models such as ChatGPT-4 to help with the task.**

In the second task, we want you to create persuasive contents for a campaign. **This time, we will provide you with the textual content generated by ChatGPT-4**, which you may use as your inspiration or a draft that you edit upon.

For both tasks, your responses are required to be textual (e.g., no graphical content).

Page 1 of the instructions given to the professional content creators: Overview

☐ Q886

You will see your first task on the next page. The task is to write advertising content for a product in fewer than 100 words without relying on LLM.

Your goal is to get people interested in this product after seeing your content.

Page 2 of the instructions given to the professional content creators: The Goal for Advertising Content

Product: Ready America 72 Hour Emergency Kit

Features:

Emergency Kit For Disaster Preparedness: Stay Prepared For Emergency Situations Everywhere You Go. Ready America Emergency Survival Kits Sustain 2 People For 3 Days (72 Hours). Basic Essential Emergency Supplies Recommended By The American Red Cross

Includes 33-Piece First Aid Kit: Comes With First Aid Essentials For Minor Injuries Like Bandages And Wound Cleaning Solution. Take Care Of Medical Emergencies With Easy To-Use Supplies. Great For A Car Emergency Kit Or Boat Accessories

Food, Water, Emergency Blankets: Everything You Need In Case Of An Emergency. 2 Survival Food Bars (2400 Calories) And Water Pouches With A 5-Year Shelf Life, Survival Blankets, Disposable Dust Masks, Nitrile Gloves, Whistle, Ponchos, Pocket Tissues

Two Safety Light Sticks: In Case You'Re Outdoors Or In The Wilderness With No Light, This Emergency Bag Includes Two Safety Light Sticks. Equipped With Survival Gear For Hunting, Boating, Workplace, Camping, Backpacking, Fishing, Hiking, Travel, Home

Product3

Please create advertising content for the above product in fewer than 100 words

Page 3 of the instructions given to the professional content creators: The content creators all first complete a task without access to ChatGPT-4's content. For those who first complete a task about the product, they are given a picture and a description of the product and asked to create advertising content for the product. The products and campaigns vary for different content creators. For those who first complete a task about the campaign, they are given a task prompt revealing the campaign (see below).

Q1870

You will see your second task on the next page. Your task is to write persuasive content for a campaign in fewer than 100 words. Your goal is to persuade people to change their behavior after seeing your content.

We will provide you with the textual content generated by ChatGPT-4, which you may use either as inspiration for your own content or as a first draft that you can edit upon.

In order to obtain the persuasive content generated by ChatGPT-4, we provided ChatGPT-4 with a prompt on the campaign topic, and asked it to create persuasive content in fewer than 100 words. Both the input prompt and the output by ChatGPT-4 will be available to you on the next page.

Both the input prompt and output by ChatGPT-4 will be available to you on the next page.

Page 4 of the instructions given to the professional content creators: They are briefed about the second task, and they will be presented with the content generated by ChatGPT-4.

Racism_AI

Please create persuasive content for a client (e.g., an NGO doing a campaign) to convince people to stop racism in fewer than 100 words.

Below is the persuasive content created by ChatGPT-4:

Embrace unity and take a stand against racism! Our diverse world is built on the richness of varied cultures, backgrounds, and experiences. By dismantling prejudice and promoting understanding, we can foster a society where everyone is treated with fairness, respect, and dignity. Together, let's challenge discrimination and empower ourselves to celebrate our differences. It starts with you – have open conversations, educate yourself, and be an ally. Join the movement for a more inclusive, compassionate world, where we can all thrive, united in our humanity. Remember, love conquers hate – always.

Page 3 of the instructions given to the professional content creators: The content creators all complete the second task with access to ChatGPT-4's content. For those whose first task is about a product, the second task will be about one of the five campaigns. For those whose first task is about a campaign, their second task will be about one of the five products. They are given a picture and a description of the product and asked to create advertising content for the product.

2. Content Generated by Human, AI, Augmented AI, and Augmented Human

Task	Human	AI	Augmented AI	Augmented Human
Emergency Kit	Gain peace of mind knowing you are taking your emergency preparedness efforts to the next level with Ready	Be prepared for the unexpected with the Ready America 72 Hour Emergency Kit! Designed for two people, this kit	Be prepared for the unexpected with the Ready America 72 Hour Emergency Kit! As recommended by the American	Stay prepared and confident with Ready America's 72 Hour Emergency Kit. Designed for two people, this

America Emergency Survival Kits. With basic essential emergency supplies and survival gear recommended by the American Red Cross, you'll have everything you need for two people to sustain three days (72 hours) in one backpack, including a 33-piece first aid kit, food bars, water, blankets and light sticks. Perfect for hunting, hiking, fishing, boating, backpacking, camping, travel or even at home. Be prepared for whatever comes your way with Ready America Emergency	provides essential emergency supplies for 3 days, as recommended by the American Red Cross. Inside, find a 33-piece first aid kit for minor injuries, 2400-calorie survival food bars, water pouches with a 5-year shelf life, emergency blankets, dust masks, nitrile gloves, a whistle, ponchos, and pocket tissues. Stay safe outdoors or at home with two safety light sticks, perfect for hunting, boating, camping, and more. Don't get caught unprepared - make Ready America your go-to emergency solution!	Red Cross, this kit is designed for two people and provides essential emergency supplies for three days. Inside, you'll find a 33-piece first aid kit for minor injuries, two 2,400-calorie survival food bars, water pouches with a 5-year shelf life, emergency blankets, dust masks, nitrile gloves, a whistle, ponchos, pocket tissues, and two safety light sticks. Stay safe at home or outdoors when hunting, boating, camping, and more. Don't get caught unprepared - make Ready America your go-to emergency solution!	comprehensive kit includes essential emergency supplies recommended by the American Red Cross. From a 33-piece first aid kit to food bars, water pouches, blankets, and light sticks, you'll have everything necessary to face unexpected situations. Ideal for outdoor adventures, travel, or home use, this emergency kit ensures you're ready for anything. Don't leave safety to chance – equip yourself with Ready America's 72 Hour Emergency Kit today.

	Survival Kits.			
Electric bike	The eBike Designed for the Smart City! Whatever your urban environment throws at you, the Heybike Cityscape Electric Bike has you covered. Travel up to 40 miles per charge thanks to its 36V battery. A 350w rear-drive motor lets you cruise at up to 19mph, making even the longest commutes a breeze. Experienced cyclists will fall in love with the 7-speed transmission while new riders ease into cycling with dual shock absorbers and 3 adjustable pedal	Ride into the future with the Heybike Cityscape Electric Bike - the ultimate 350W Electric City Cruiser! Experience the freedom of a 40-mile range, thanks to a modular 36V, 10Ah battery, and reach top speeds of 19mph. Upgrade your commute with our Step-Through geometry frame and wide swept-back handlebar, ensuring a comfortable ride on 26" puncture-resistant tires, Shimano 7-speed, and dual shock absorber. Choose from 3 working modes to	Open doors to open roads. The 250W Electric City Cruiser Bicycle is built for adventure. Pre-assembled and ready to ride, your journey is everything that's ahead. REVOLUTIONARY FREEDOM- Thanks to a modular 36V and 10AH battery you can reach speeds of 19 mph with a 40-mile range and 3 working modes to choose from. SMOOTHER SAILING- Shimano 7-speed, dual shock absorber and 26" puncture resistance tires make for an easy ride. SAFE AND SOUND- Dual	Introducing the Heybike Cityscape Electric Bike – the ultimate urban cruiser tailored for your city adventures! With its powerful 350W motor and long-range 36V, 10Ah battery, you can travel up to 40 miles per charge at speeds of up to 19mph. The step-through frame and swept-back handlebar ensure a comfortable ride, while the 26" puncture-resistant tires, Shimano 7-speed gears, and dual shock absorbers guarantee a smooth experience. Choose from

	assist modes. From a relaxing ride to a cardio workout, the HeyBike Cityscape is the eBike you've been looking for.	suit your journey, and trust in the dual disc brakes and lighting system for unbeatable safety. Arriving mostly pre-assembled, just hop on and let the adventure begin!	disc brakes and lighting system ensure safety with a better, more geometric frame and wide swept-back handlebar.	throttle or three pedal-assist modes for effortless commuting or a workout. Pre-assembled for convenience, the Cityscape also features an integrated rear rack for cargo. Transform your city rides today!
Tumbler	The Hydro Flask Reusable Mug. The Mug for drop off. The Mug for drop ins. The Mug you won't drop.	Introducing the Hydro Flask Stainless Steel Reusable Mug! Sip in style with a splash-proof, Closeable Press-In Lid, perfect for on-the-go enjoyment. Its Soft Touch Exterior ensures comfort and grip, while the innovative TempShield double-wall vacuum insulation	Meet the only mug you'll ever need. The Hydro Flask Stainless Steel Reusable Mug is soft on the outside for a comforting grip, with 18/8 pro-grade stainless steel on the inside so you'll never taste the metal. Its TempShield double-wall vacuum insulation keeps liquids at	Introducing the Hydro Flask Stainless Steel Reusable Mug: your ultimate companion for sips and spills. Featuring a closeable Press-In Lid that slides open for effortless drinking and straw access, while keeping splashes at bay. The Soft Touch Exterior ensures comfortable grip

		keeps your drinks hot or cold for hours. Crafted from 18/8 pro-grade stainless steel, taste the purity without flavor transfer. Live worry-free with a BPA-Free & Toxin-Free design. Experience the ultimate sustainable companion for your favorite beverages. Cheers to a greener future!	the perfect temperature for hours – piping hot, icy cold, or wherever you like it in-between. The snug lid stops spills and splashes at home or on-the-go. It slides open for sips or straws. The mug is made of toxin-free materials (no BPAs!). You enjoy a safe and delicious drinking experience, every time.	and the TempShield double-wall vacuum insulation keeps your drinks hot or cold for hours. Made with 18/8 pro-grade stainless steel, enjoy pure taste without flavor transfer. BPA-free and toxin-free, sip with confidence. Say hello to the Hydro Flask Reusable Mug - easy to hold, and hard to drop.
Air Fryer	Enjoy yummy goodness every day—with 70% less calories from fat. Make crispy, golden fixings—guilt-free —with just a flip of the switch. With the #1 selling XL Power	Discover the PowerXL Air Fryer, your all-in-one kitchen solution from the #1 best-selling air fryer brand! This 7-qt multi-cooker features built-in commercial steaming	Discover the PowerXL Air Fryer, an all-in-one kitchen solution from the #1 best-selling air fryer brand! Perfect for everything from vegetables and shellfish to	Discover the PowerXL Air Fryer – the ultimate kitchen gadget for guilt-free indulgence. Enjoy crispy, golden delights with 70% less fat, thanks to its innovative

	Brand Air Fryer, you can cook tender fall-off-the-bone ribs and wings with a crispy crust. Steam tender vegetables. Bake bread, toast bagels, or reheat pretzels. Air fry technology and 1-touch cooking presets mean you can press a button and forget about it. Serve a family-sized meal or just plate for yourself. Either way, cooking has never been easier. And with our dishwasher-safe parts, clean-up is a breeze. Just what you'd expect.	technology, perfect for vegetables, shellfish, and bakery-style treats. Enjoy 10-in-1 LED digital 1-touch cooking presets and temperature control up to 400°. Savor crispy, tender bites with up to 70% fewer calories from fat. The PowerXL Air Fryer includes a 7-qt fry basket, steamer/air fryer rack, and recipe book, with dishwasher-safe parts for easy cleanup. Elevate your culinary experience with the versatile, health-conscious PowerXL Air Fryer!	bakery-style treats. Create delicious meals with up to 70% fewer calories from fat. Features include: 7-quart capacity with built-in commercial steaming technology. --10 one-touch digital cooking presets, and temperature control up to 400°. --Includes a fry basket, steamer/air fryer rack, and recipe book, with dishwasher-safe parts for easy cleanup. Elevate your culinary experience with the versatile, health-conscious PowerXL Air Fryer!	air-fry technology. The versatile 1700-watt, 7-qt air fryer features 10-in-1 LED digital one-touch cooking presets for everything from air frying and steaming to baking and reheating. Achieve perfect results with tender steamed veggies, bakery-style bread, and crispy ribs or wings – all with fewer calories. The PowerXL Air Fryer comes with a fry basket, steamer rack, and recipe book, and its dishwasher-safe parts make cleanup a breeze. Revolutionize your cooking experience today!

| Projector | What if your mobile phone could project cinema-quality video onto a 10-foot screen at the same resolution as modern flat-screen TVs? But what would you look at while you're watching the movie, you say? We've got good news. The WEMAX Go Advanced projector is the size of a phone and has the power of a movie theater. Best of all: it can handle ambient light, project onto most any angle, and has endless battery power. Very handy in the office, an offsite event, or any on-the-go adventure. The | Introducing the WEMAX Go Advanced Portable Smart Laser Projector – innovation and convenience in the palm of your hand! Utilizing ALPD laser technology and TI DLP for vivid, bright images up to 120 inches, this ultra-portable projector is perfect for on-the-go presentations. With a built-in rechargeable battery lasting 1.5 hours, you can present without power outlets or cables. Weighing just 1.7 lbs and under 1-inch thick, it's the ideal travel companion. Say goodbye to complex setups with | Present with confidence on the go—introducing the WEMAX Go Advanced Portable Smart Laser Projector. Less than 1" thick and just 1.7 pounds, it goes everywhere with you—and with 1.5 hours of viewing time, it'll last through your meeting even when a charger isn't handy. You'll get high-end cinema-quality ALPD laser technology with industry-first portability. Wow your viewers even in ambient light with full HD using over 16.7M distinct colors and 600 ANSI lumens of brightness at up to 120" projection size. Advanced | Experience cinema magic anytime, anywhere with the WEMAX Go Advanced Portable Smart Laser Projector. Harnessing ALPD and DLP technologies, this pocket-sized powerhouse delivers stunning, vivid visuals up to 120 inches, even in ambient light. With a rechargeable battery, enjoy 1.5+ hours of uninterrupted presentations or entertainment—simply extend with a power bank for limitless viewing. Weighing a mere 1.7 lbs, it's perfect for the office, conferences, or on-the-go adventures. Smart |

	only other thing you'll need to bring is the popcorn.	auto-calibration features, making every presentation a breeze. Experience the future of projection with WEMAX Go Advanced!	auto-calibration technology means you'll be up and running in seconds every time.	screen calibrations ensure hassle-free setup, so you can focus on what matters—just don't forget the popcorn!
Recycle	What if you made a choice today that saved the life of a single tree? Instead of getting chopped down, chipped, and shipped to a paper mill, your tree stands tall for 1,000 years. Each day, it cleans carbon dioxide from the air and produces enough oxygen for a family of four. It shelters generations of wildlife, protects the soil with its roots, and lowers air temperature	Join the recycling revolution and make a difference for our planet! Recycling conserves precious resources, reduces waste, and combats pollution. It's a small action with a big impact, helping to preserve our environment for future generations. Don't wait; start recycling today and become an agent of change. Together, we can create a cleaner, greener, and more	A letter to the next generation. We're sorry that we didn't do more. We're sorry that we didn't speak up. We're sorry that we didn't stand up. We're sorry that so many gave up. If only we realized the impact we could have had, if only we knew the life you could have had, with only a few small changes. If we could turn back time, and do it all again, just know that we would do	Imagine making a simple decision today that could save a tree and make a lasting impact on our environment. By recycling paper, cardboard, and other reusable materials, you can help preserve trees that provide oxygen for families, homes for wildlife, and shade that cools our air. Your choice to recycle not only reduces waste but also combats climate change by

	with its shade. Recycle paper, cardboard, and other reusable materials instead of putting them in the trash. The future of the world depends on the choices we all make today.	sustainable world for all. Remember, it's not just about us – it's about our children, our wildlife, and the Earth we all share. Recycle, because every little bit counts!	it differently. We'd recycle, for you & for them. We'd recycle, for you to have a chance.	minimizing carbon dioxide levels. Take action now - recycling is a small step with immense benefits. The future of our planet depends on the collective choices we make today.
Do More Physical Exercises	HEADLINE It's not about better fitting pants—it's everything to do with your heart. BODY/ARTICLE There's a way to decrease the obesity rate in America, and it doesn't take an expensive gym membership or exhaustive hours of training. It starts with your heart. Raising your heart rate doesn't just slim down our waist,	Get moving and unlock the power of physical exercise! Boost your energy, improve your mood, and enhance your overall well-being by incorporating regular workouts into your daily routine. Experience the life-changing benefits, from reduced stress and better sleep to increased strength and weight	The benefits of physical exercise go further than just skin deep. From a boost in energy to improvements in your mood, you can enhance nearly every facet of your well-being just by staying active and making time for physical activity in your daily routine. From reduced stress to better sleep to improved posture, regular	Transform your life by simply moving more! It's not just about fitting into your favorite outfit - it's about improving your heart health and overall well-being. Regular physical exercise can lower blood pressure, reduce stress, decrease inflammation, and help prevent diabetes, leading to a longer, healthier life. You

	studies show physical activity can lower blood pressure and stress, reduce inflammation and slow the development of diabetes, which contribute to a longer life. Here are low-impact activities to increase heart health. 1. Swimming 2. Walking 3. Low-intensity interval training 4. Bicycling 5. Weightlifting	management. It's not just about looking good; it's about feeling fantastic from the inside out. There's an activity for everyone – walk, run, dance, or lift, just get started. Take the first step towards a healthier, happier life today. Remember, a little progress each day adds up to big results!	exercise and physical activity provide countless benefits. Even weight management and increased strength if swimsuit season is, in fact, your biggest priority. There's an activity for everyone, from low-impact to high-intensity! Take your first steps towards a healthier you! Get out and get moving!	don't need an expensive gym membership or grueling workout sessions. Start with these low-impact activities to boost your heart health and happiness: 1.Swimming 2.Walking 3. Low-intensity interval training 4. Bicycling 5. Weightlifting Embrace the power of movement and unlock a healthier you!
Stop Racism	Racism still exists in our modern world, and it hurts us all. But hate can never end hate, and separating ourselves into angry camps will never bring us all	Embrace unity and take a stand against racism! Our diverse world is built on the richness of varied cultures, backgrounds, and experiences. By dismantling	Would you rather live in a world fueled by love or hate? Perhaps that's what we should ask ourselves as we consider the pervasive global problem of	Together, we can overcome the destructive force of racism. It may persist in today's society, but we have the power to change that. Division and hatred will not

	together. We want to end racism in America today, tomorrow, and forever. By finding common ground where people can share without fear and listen without judgment, we believe that racism can forever be a relic of the past. Join us as we embark on a bold new journey to unite future generations and create a world free from an old affliction.	prejudice and promoting understanding, we can foster a society where everyone is treated with fairness, respect, and dignity. Together, let's challenge discrimination and empower ourselves to celebrate our differences. It starts with you – have open conversations, educate yourself, and be an ally. Join the movement for a more inclusive, compassionate world, where we can all thrive, united in our humanity. Remember, love conquers hate – always.	systemic racism. Our world is built on the richness of diverse cultures, backgrounds and experiences. By dismantling prejudice, we can foster a society where everyone is treated with fairness, respect and dignity. But we all must do our part to openly challenge discrimination – have direct conversations, educate ourselves and be allies. Join the movement for a more inclusive, compassionate world, where we can all thrive, united in our humanity. Love can conquer hate.	lead to unity, so we must choose understanding and empathy. Let's end racism, today and always, by fostering an environment where open dialogue thrives, without fear or judgment. Stand with us as we forge a path towards a future where generations can live in harmony, leaving racism behind as a distant memory. Embrace our shared humanity and help create a world free from prejudice.

| Wash their hands often | If you could make one simple, fast, effective, and free choice every day that would improve your health, wouldn't you try it? Then simply wash your hands. All it takes is a little soap, water, and less than 30 seconds of your time, and you can drastically reduce the spread of bacteria, viruses, and microbes, keeping you and others around you healthy. Demonstrating this to children and encouraging them to do it before and after eating, using the bathroom, and riding the bus can help stem the spread of illness worldwide. | Clean hands save lives! Washing your hands regularly is a simple yet powerful way to protect yourself and those around you from illness. By making hand hygiene a priority, you can effectively combat the spread of harmful germs and viruses. It's easy, quick, and a cornerstone of good health. So, remember to lather up with soap and water for at least 20 seconds, especially before meals and after touching public surfaces. Let's all commit to healthier habits, and together, we can create a safer, cleaner world for | It's easy to save a life: Wash your hands.   It's simple. Just follow these 3 steps before and after eating and after touching public surfaces: 1)   Use soap and water to create a lather. 2)   Rub your hands together for at least 20 seconds. 3) Rinse and dry. That's it! Do it now. And start saving lives. | Protect your health and the well-being of those around you with one quick, easy, and cost-free action: washing your hands. Using just soap and water for 20-30 seconds, you can significantly decrease the spread of germs, safeguarding your health and the health of others. Teach children the importance of handwashing before and after meals, restroom breaks, and public transportation use. This simple habit is a powerful tool in preventing the spread of illnesses globally. Choose handwashing, and contribute to a healthier world for all. |

		everyone. Wash your hands, because prevention starts with you!		
Eat less junk food	Real food tastes better. It's better for your body and the planet. You don't have to be a five-star chef to move beyond junk food—you just have to consider all your options and make the healthier choice more often. Apples? Delicious! Salad? Your new best friend. Roast chicken? Easier than you think and so, so worth it. And hey, when you do eat those nachos or scarf down that cheesecake, enjoy every moment of it, because guilt	Transform your life by choosing healthier alternatives to junk food! Consuming excess junk food can lead to weight gain, fatigue, and long-term health issues. Opt for nutritious meals packed with vitamins, minerals, and antioxidants to feel energized, maintain a healthy weight, and reduce your risk of chronic diseases. Embrace the power of wholesome ingredients and discover the delicious flavors	Junk food makes us feel good…almost immediately. Just thinking about it makes our mouths water. So why is it called junk food? Because you're being tricked. And not in a fun magician kind of way. No, junk food tricks your brain and body into thinking you're better off. It triggers all the feel-good chemicals without giving you the actual good stuff you need to feel better tomorrow when you wake up…or five years from now. So,	Savor the true flavors of life! Prioritize your well-being and the environment by opting for wholesome, natural foods. With countless nutritious alternatives, ditching junk food doesn't require culinary expertise—just a commitment to making better choices. Relish the crunch of apples, the freshness of salads, or the warmth of a home-cooked roast chicken. Treat yourself guilt-free to the

	tastes terrible—and you deserve the best.	of fruits, vegetables, lean proteins, and whole grains. Remember, a balanced diet is the key to a happier, healthier you. Make the switch today – your body and mind will thank you!	maybe try eating less junk food, and see how you feel. We think you'll feel it, and this time in a good way.	occasional indulgence, knowing you're on the path to a healthier lifestyle. Choose quality over junk and elevate your taste buds and overall health to new heights!

3. Instructions for the Online Participants to Rate the Content Quality

Informed Consent

Your decision to complete this survey is voluntary. There is no way for us to identify you. The only information we will have, in addition to your responses, is the time at which you completed the survey. The results of the research may be presented at scientific meetings or published in scientific journals. Going to the next page indicates that you are at least 18 years of age and agree to complete this survey voluntarily.

Q894

Please also verify the following

[ ] I'm not a robot — reCAPTCHA Privacy - Terms

--------- Page Break ---------

workid

Please enter your WorkerID (we will delete this information after approving your work)

[   ]

--------- Page Break ---------

Q821

Overview

We have generated some advertising content for five different products and some persuasive content for five different campaigns. We want you to evaluate the quality of the text you will read.

○ I have carefully reviewed the above information

Page 1 of the instructions given to the online participants: Overview

Information about Content Generators (please read carefully before proceeding)

The textual content for the products and campaigns you are about to evaluate were generated under one of the four paradigms:

- AI only,
- A Human Expert only,
- AI-generated first, with final edits by human expert, or,
- Human Expert-generated first, with final edits by AI.

AI Only: For each task, the textual content is generated by ChatGPT-4 developed by OpenAI, which is the state-of-the-art large language models trained to produce human-like textual contents. For example, we gave a product's name and its feature descriptions to the AI and asked it generate advertising content. Or we gave the prompt of a campaign to the AI and asked it to generate persuasive content. The output by ChatGPT-4 was taken directly without further human-edit.

Human Expert Only: Each textual content is generated by a human expert without the help of ChatGPT-4 or other large language models. The human experts are recruited from one of the world's best consulting firms. Each human expert was given the same information as the AI, and their generated content was taken directly without further editing.

AI-generated first, with final edits by human expert: In addition to the information provided in the "human expert only" condition, the professional content creators were also given the content generated by ChatGPT-4 (i.e., the exact output by the AI in the "AI Only" condition), which they may use as an inspiration or a first draft that they can edit upon. In other words, in this condition, a human expert will have access to the content generated by ChatGPT-4 but humans will finalize the output for each task.

Human Expert-generated first, with final edits by AI: In addition to the information provided in the "AI only" condition, ChatGPT-4 was also given the content generated by one of the professional content creators who had completed the task (i.e., the exact output by a human expert in the "Human Expert Only" condition), which ChatGPT-4 may edit. In other words, in this condition, the AI will have access to the content generated by a human expert but the AI will finalize the output for each task.

○ I have carefully reviewed the above information

Page 2 of the instructions given to the online participants: Those who are in the baseline condition skipped this page. Only those in the "uninformed" and "informed" condition saw this information.

lan_check

Language Check: What is the eleventh month of a year called?

○ December

○ October

○ July

○ November

○ June

○ January

Page 3 of the instructions given to the online participants: Attention Check

Instructions

You will first evaluate advertising content generated for five different products. When generating the advertising content, a content generator was given the product's name and a description of its features.

After seeing the advertising content for each product, we want to know

(1) to what extent you are interested in learning more about the product (e.g., its price, user reviews, complete product descriptions, etc.)

(2) suppose you are the seller of the product, to what extent you are satisfied with the generated advertising content

(3) suppose you are the seller of the product and the content has a copyright, what is the maximum amount of money you are willing to pay to use the content as your advertisement. (Assuming you have a $1000-budget for this content and please state the maximum amount you are willing to pay between $1 and $1000. Please enter a numerical value without the $ sign).

Each advertising content should contain fewer than 100 words. Please read the content carefully.

○ I have reviewed the above information

Page 4 of the instructions given to the online participants in the baseline condition

Instructions

You will first evaluate advertising content generated for five different products. When generating the advertising content, each content generator (*either an AI, or a human expert, or a human expert who finalizes content first generated by AI, or an AI that finalizes content first generated by a human expert*) was given the product's name and a description of its features.

After seeing the advertising content for each product, we want to know

(1) to what extent you are interested in learning more about the product (e.g., its price, user reviews, complete product descriptions, etc.)

(2) suppose you are the seller of the product, to what extent you are satisfied with the generated advertising content

(3) suppose you are the seller of the product and the content has a copyright, what is the maximum amount of money you are willing to pay to use the content as your advertisement. (Assuming you have a $1000-budget for this content and please state the <u>maximum</u> amount you are willing to pay between $1 and $1000. Please enter a numerical value without the $ sign).

Each advertising content should contain fewer than 100 words. Please read the content carefully.

○ I have reviewed the above information

Page 4 of the instructions given to the online participants in the uninformed condition

Instructions

You will first evaluate advertising contents generated for five different products. When generating the advertising content, each content generator (*either an AI, or a human expert, or a human expert who finalizes content first generated by AI, or an AI that finalizes content first generated by a human expert*) was given the product's name and a description of its features.

After seeing the advertising content for each product, we want to know

(1) to what extent you are interested in learning more about the product (e.g., its price, user reviews, complete product descriptions, etc.)

(2) suppose you are the seller of the product, to what extent you are satisfied with the generated advertising content

(3) suppose you are the seller of the product and the content has a copyright, what is the <u>maximum</u> amount of money you are willing to pay to use the content as your advertisement. (Assuming you have a $1000-budget for this content and please state the <u>maximum</u> amount you are willing to pay between $1 and $1000. Please enter a numerical value without the $ sign).

Each advertising content should contain fewer than 100 words. Please read the content carefully. We will also tell you under which paradigm a content is created.

○ I have reviewed the above information

Page 4 of the instructions given to the online participants in the informed condition. Note that they are told they will be informed under which paradigm a content is created

Q1122

The task will start on the next page. Please read the content and questions carefully!

○ I have reviewed the above information

Page 5 of the instructions given to the online participants

Q1908

If p5 Is Equal to 1
And study Is Not Equal to 2

Below is the advertising content:

"What if your mobile phone could project cinema-quality video onto a 10-foot screen at the same resolution as modern flat-screen TVs? But what would you look at while you're watching the movie, you say? We've got good news. The WEMAX Go Advanced projector is the size of a phone and has the power of a movie theater. Best of all: it can handle ambient light, project onto most any angle, and has endless battery power. Very handy in the office, an offsite event, or any on-the-go adventure. The only other thing you'll need to bring is the popcorn."

○ I have reviewed the above advertising content

Q1065

If p5 Is Equal to 1
And study Is Equal to 2

Below is the advertising content created by a human expert:

"What if your mobile phone could project cinema-quality video onto a 10-foot screen at the same resolution as modern flat-screen TVs? But what would you look at while you're watching the movie, you say? We've got good news. The WEMAX Go Advanced projector is the size of a phone and has the power of a movie theater. Best of all: it can handle ambient light, project onto most any angle, and has endless battery power. Very handy in the office, an offsite event, or any on-the-go adventure. The only other thing you'll need to bring is the popcorn."

Page 6 of the instructions given to the online participants who saw human generated content for the product. Note those in the informed condition also saw the identity of the content creator, but those in the other two conditions were not.

If p5 Is Equal to 2
And study Is Not Equal to 2

Below is the advertising content:

"Introducing the WEMAX Go Advanced Portable Smart Laser Projector – innovation and convenience in the palm of your hand! Utilizing ALPD laser technology and TI DLP for vivid, bright images up to 120 inches, this ultra-portable projector is perfect for on-the-go presentations. With a built-in rechargeable battery lasting 1.5 hours, you can present without power outlets or cables. Weighing just 1.7 lbs and under 1-inch thick, it's the ideal travel companion. Say goodbye to complex setups with auto-calibration features, making every presentation a breeze. Experience the future of projection with WEMAX Go Advanced!"

○ I have reviewed the above advertising content

Q1066

▾ 🔁 Display this question

If p5 Is Equal to 2
And study Is Equal to 2

Below is the advertising content created by an AI:

"Introducing the WEMAX Go Advanced Portable Smart Laser Projector – innovation and convenience in the palm of your hand! Utilizing ALPD laser technology and TI DLP for vivid, bright images up to 120 inches, this ultra-portable projector is perfect for on-the-go presentations. With a built-in rechargeable battery lasting 1.5 hours, you can present without power outlets or cables. Weighing just 1.7 lbs and under 1-inch thick, it's the ideal travel companion. Say goodbye to complex setups with auto-calibration features, making every presentation a breeze. Experience the future of projection with WEMAX Go Advanced!"

○ I have reviewed the above advertising content

Page 6 of the instructions given to the online participants who saw AI generated content for the product.

Q1910

Display this question

If p5 Is Equal to 3
And study Is Not Equal to 2

Below is the advertising content:

"Present with confidence on the go—introducing the WEMAX Go Advanced Portable Smart Laser Projector. Less than 1" thick and just 1.7 pounds, it goes everywhere with you—and with 1.5 hours of viewing time, it'll last through your meeting even when a charger isn't handy. You'll get high-end cinema-quality ALPD laser technology with industry-first portability. Wow your viewers even in ambient light with full HD using over 16.7M distinct colors and 600 ANSI lumens of brightness at up to 120" projection size. Advanced auto-calibration technology means you'll be up and running in seconds every time."

○ I have reviewed the above advertising content

Q1067

Display this question

If p5 Is Equal to 3
And study Is Equal to 2

Below is the advertising content created by a human expert who edits and finalizes the content first generated by the AI:

"Present with confidence on the go—introducing the WEMAX Go Advanced Portable Smart Laser Projector. Less than 1" thick and just 1.7 pounds, it goes everywhere with you—and with 1.5 hours of viewing time, it'll last through your meeting even when a charger isn't handy. You'll get high-end cinema-quality ALPD laser technology with industry-first portability. Wow your viewers even in ambient light with full HD using over 16.7M distinct colors and 600 ANSI lumens of brightness at up to 120" projection size. Advanced auto-calibration technology means you'll be up and running in seconds every time."

○ I have reviewed the above advertising content

Page 6 of the instructions given to the online participants who saw "augmented human" generated content for the product.

> ▾ ⤷ Display this question
>
> If p5 Is Equal to 4
> And study Is Not Equal to 2

Below is the advertising content:

) "Experience cinema magic anytime, anywhere with the WEMAX Go Advanced Portable Smart Laser Projector. Harnessing ALPD and DLP technologies, this pocket-sized powerhouse delivers stunning, vivid visuals up to 120 inches, even in ambient light. With a rechargeable battery, enjoy 1.5+ hours of uninterrupted presentations or entertainment—simply extend with a power bank for limitless viewing. Weighing a mere 1.7 lbs, it's perfect for the office, conferences, or on-the-go adventures. Smart screen calibrations ensure hassle-free setup, so you can focus on what matters—just don't forget the popcorn!"

○ I have reviewed the above advertising content

Q1068

> ▾ ⤷ Display this question
>
> If p5 Is Equal to 4
> And study Is Equal to 2

Below is the advertising content created by an AI that edits and finalizes the content first generated by a human expert:

"Experience cinema magic anytime, anywhere with the WEMAX Go Advanced Portable Smart Laser Projector. Harnessing ALPD and DLP technologies, this pocket-sized powerhouse delivers stunning, vivid visuals up to 120 inches, even in ambient light. With a rechargeable battery, enjoy 1.5+ hours of uninterrupted presentations or entertainment—simply extend with a power bank for limitless viewing. Weighing a mere 1.7 lbs, it's perfect for the office, conferences, or on-the-go adventures. Smart screen calibrations ensure hassle-free setup, so you can focus on what matters—just don't forget the popcorn!"

○ I have reviewed the above advertising content

Page 6 of the instructions given to the online participants who saw "augmented AI" generated content for the product.

To what extent are you interested in <u>learning more about the product</u> (e.g., its price, user reviews, complete product descriptions, etc.) on a scale from 1 to 7?

1 - Not interested at all	2 - Uninterested	3 - Somewhat uninterested	4 - Neither uninterested nor interested	5 - Somewhat interested	6 - Interested	7 - Very interested
○	○	○	○	○	○	○

sat5

Suppose you are the seller of the product, to what extent are you <u>satisfied or unsatisfied</u> with the generated advertising content on a scale from 1 to 7?

1 - Very unsatisfied	2 - Unsatisfied	3 - Somewhat unsatisfied	4 - Neutral	5 - Somewhat satisfied	6 - Satisfied	7 - Very satisfied
○	○	○	○	○	○	○

pay5

Suppose you are the seller of the product and the content has a copyright, what is the <u>maximum amount of money you are willing to pay</u> to use the content as your advertisement? (Assuming you have a budget of $1000, please state your max willingness-to-pay between $1 and $1000.)

Page 6 of the instructions given to the online participants: the variables of interest.

att check

att_check

Attention Check: 1+1 = ?

○ 1

○ 2

○ 3

○ 4

○ 5

○ 6

Page 7 of the instructions given to the online participants: attention check.

☐ Q1047

Instructions

Next you will evaluate persuasive contents generated for five different campaigns. After seeing the persuasive content for each campaign, we want to know

(1) to what extent you are convinced by the content to change your behavior

(2) suppose you are the campaign manager, to what extent you are satisfied or unsatisfied with the generated persuasive content

(3) suppose you are the campaign manager and the content has a copyright, what is the maximum amount of money you are willing to pay to use the content for persuasion. (Assuming you have a $1000-budget for this content and please state the <u>maximum</u> amount you are willing to pay between $1 and $1000. Please enter a numerical value without the $ sign).

Each persuasive content should contain fewer than 100 words. Please read the content carefully.

○ I have reviewed the above information

Page 8 of the instructions given to the online participants in the baseline condition.

▾ 🔁 Display this question

If study Is Equal to 1

Instructions

Next you will evaluate persuasive contents generated for five different campaigns. Similar to the advertising content, each persuasive content was also generated under one of the four paradigms (*either an AI, or a human expert, or a human expert who finalizes content first generated by AI, or an AI that finalizes content first generated by a human expert*).

After seeing the persuasive content for each campaign, we want to know

(1) to what extent you are convinced by the content to change your behavior

(2) suppose you are the campaign manager, to what extent you are satisfied or unsatisfied with the generated persuasive content

(3) suppose you are the campaign manager and the content has a copyright, what is the maximum amount of money you are willing to pay to use the content for persuasion. (Assuming you have a $1000-budget for this content and please state the maximum amount you are willing to pay between $1 and $1000. Please enter a numerical value without the $ sign.).

Each advertising content should contain fewer than 100 words. Please read the content carefully.

○ I have reviewed the above information

Page 8 of the instructions given to the online participants in the uninformed condition.

Instructions

Next you will evaluate persuasive contents generated for five different campaigns. Similar to the advertising content, each persuasive content was also generated under one of the four paradigms (*either an AI, or a human expert, or a human expert who finalizes content first generated by AI, or an AI that finalizes content first generated by a human expert*).

After seeing the persuasive content for each campaign, we want to know

(1) to what extent you are convinced by the content to change your behavior

(2) suppose you are the campaign manager, to what extent you are satisfied or unsatisfied with the generated persuasive content

(3) suppose you are the campaign manager and the content has a copyright, what is the maximum amount of money you are willing to pay to use the content for persuasion. (Assuming you have a $1000-budget for this content and please state the maximum amount you are willing to pay between $1 and $1000. Please enter a numerical value without the $ sign).

Each advertising content should contain fewer than 100 words. Please read the content carefully. We will also tell you under which paradigm a content is created.

○   I have reviewed the above information

Page 8 of the instructions given to the online participants in the informed condition.

▾ 🔂 Display this question

If p10 Is Equal to 1
And study Is Not Equal to 2

Below is the persuasive content:

"Real food tastes better. It's better for your body and the planet. You don't have to be a five-star chef to move beyond junk food—you just have to consider all your options and make the healthier choice more often. Apples? Delicious! Salad? Your new best friend. Roast chicken? Easier than you think and so, so worth it. And hey, when you do eat those nachos or scarf down that cheesecake, enjoy every moment of it, because guilt tastes terrible—and you deserve the best."

○ I have reviewed the above persuasive content

Q1086

▾ 🔂 Display this question

If p10 Is Equal to 1
And study Is Equal to 2

Below is the persuasive content created by <u>a human expert</u>:

"Real food tastes better. It's better for your body and the planet. You don't have to be a five-star chef to move beyond junk food—you just have to consider all your options and make the healthier choice more often. Apples? Delicious! Salad? Your new best friend. Roast chicken? Easier than you think and so, so worth it. And hey, when you do eat those nachos or scarf down that cheesecake, enjoy every moment of it, because guilt tastes terrible—and you deserve the best."

○ I have reviewed the above persuasive content

Page 9 of the instructions given to the online participants who saw human generated content.

☐ Q1942

▼ 🔗 Display this question

If p10 Is Equal to 2

And study Is Not Equal to 2

⌐

"Transform your life by choosing healthier alternatives to junk food! Consuming excess junk food can lead to weight gain, fatigue, and long-term health issues. Opt for nutritious meals packed with vitamins, minerals, and antioxidants to feel energized, maintain a healthy weight, and reduce your risk of chronic diseases. Embrace the power of wholesome ingredients and discover the delicious flavors of fruits, vegetables, lean proteins, and whole grains. Remember, a balanced diet is the key to a happier, healthier you. Make the switch today – your body and mind will thank you!"

○ I have reviewed the above persuasive content

Q1087

▼ 🔗 Display this question

If p10 Is Equal to 2

And study Is Equal to 2

Below is the persuasive content created by an AI:

"Transform your life by choosing healthier alternatives to junk food! Consuming excess junk food can lead to weight gain, fatigue, and long-term health issues. Opt for nutritious meals packed with vitamins, minerals, and antioxidants to feel energized, maintain a healthy weight, and reduce your risk of chronic diseases. Embrace the power of wholesome ingredients and discover the delicious flavors of fruits, vegetables, lean proteins, and whole grains. Remember, a balanced diet is the key to a happier, healthier you. Make the switch today – your body and mind will thank you!"

○ I have reviewed the above persuasive content

Page 9 of the instructions given to the online participants who saw AI generated content.

▾ 🔗 Display this question

If p10 Is Equal to 3
And study Is Not Equal to 2

Below is the persuasive content:

"Junk food makes us feel good...almost immediately. Just thinking about it makes our mouths water. So why is it called junk food? Because you're being tricked. And not in a fun magician kind of way. No, junk food tricks your brain and body into thinking you're better off. It triggers all the feel-good chemicals without giving you the actual good stuff you need to feel better tomorrow when you wake up...or five years from now. So, maybe try eating less junk food, and see how you feel. We think you'll feel it, and this time in a good way."

○ I have reviewed the above persuasive content

Q1088

▾ 🔗 Display this question

If p10 Is Equal to 3
And study Is Equal to 2

Below is the persuasive content created by <u>a human expert who edits and finalizes the content first generated by the AI</u>:

"Junk food makes us feel good...almost immediately. Just thinking about it makes our mouths water. So why is it called junk food? Because you're being tricked. And not in a fun magician kind of way. No, junk food tricks your brain and body into thinking you're better off. It triggers all the feel-good chemicals without giving you the actual good stuff you need to feel better tomorrow when you wake up...or five years from now. So, maybe try eating less junk food, and see how you feel. We think you'll feel it, and this time in a good way."

○ I have reviewed the above persuasive content

Page 9 of the instructions given to the online participants who saw "augmented human"content.

Below is the persuasive content:

"Savor the true flavors of life! Prioritize your well-being and the environment by opting for wholesome, natural foods. With countless nutritious alternatives, ditching junk food doesn't require culinary expertise—just a commitment to making better choices. Relish the crunch of apples, the freshness of salads, or the warmth of a home-cooked roast chicken. Treat yourself guilt-free to the occasional indulgence, knowing you're on the path to a healthier lifestyle. Choose quality over junk and elevate your taste buds and overall health to new heights!"

○ I have reviewed the above persuasive content

Q1089

Below is the persuasive content created by an AI that edits and finalizes the content first generated by a human expert:

"Savor the true flavors of life! Prioritize your well-being and the environment by opting for wholesome, natural foods. With countless nutritious alternatives, ditching junk food doesn't require culinary expertise—just a commitment to making better choices. Relish the crunch of apples, the freshness of salads, or the warmth of a home-cooked roast chicken. Treat yourself guilt-free to the occasional indulgence, knowing you're on the path to a healthier lifestyle. Choose quality over junk and elevate your taste buds and overall health to new heights!"

○ I have reviewed the above persuasive content

Page 9 of the instructions given to the online participants who saw "augmented AI" content.

int10

To what extent are you convinced by the above content to eat less junk food on a scale from 1 to 7?

1 - Not convinced at all	2 - Not convinced	3 - Somewhat not unconvinced	4 - Neutral	5 - Somewhat convinced	6 - Convinced	7 - Very convinced
○	○	○	○	○	○	○

sat10

Suppose you are the campaign manager, to what extent are you satisfied or unsatisfied with the generated persuasive content on a scale from 1 to 7?

1 - Very unsatisfied	2 - Unsatisfied	3 - Somewhat unsatisfied	4 - Neutral	5 - Somewhat satisfied	6 - Satisfied	7 - Very satisfied
○	○	○	○	○	○	○

pay10

Suppose you are the campaign manager and the content has a copyright, what is the maximum amount of money you are willing to pay to use the content for persuasion? (Assuming you have a budget of $1000, please state your max willingness-to-pay between $1 and $1000.)

Page 9 of the instructions given to the online participants: the variables of interest.

4. Additional Results 1: Uninformed vs Informed

We also compare participants' evaluation of content quality between the "uninformed" and the "informed" condition. The results are illustrated by a comparison between the "pinkish" bar (baseline) and the "bluish" bar (informed) within each paradigm in Figure 1 and Figure 2. The regression analysis is the same as previously mentioned except we change the coding for the dummy variable (1 = informed, 0 = uninformed). First, we find that given the same content generated solely by a human expert, participants felt more satisfied with the content and were willing to pay more if they were informed of the content was created by a human expert ($b\_satisfaction$ = 0.23, p-value = 0.0000; b_wtp = 0.21, p-value = 0.01). Although participants were slightly more satisfied with the content generated solely by AI when they were informed (b = 0.14, p-value = 0.01), their willingness-to-pay did not increase significantly (b = 0.08, p-value = 0.11). We also do not find any effect on the satisfaction level or willingness-to-pay for the remaining two content generation paradigms (Augmented Human: $b\_satisfaction$ = -0.03, p-value = 0.62, b_wtp = 0.02, p-value = 0.81; Augmented AI: $b\_satisfaction$ = 0.07, p-value = 0.18, b_wtp = 0.11, p-value = 0.14). Therefore, we again have evidence of human favoritism after resolving uncertainty around the identity of the content creator. Nevertheless, evidence for AI favoritism is ambiguous because although the positive effect on perceived satisfaction after resolving uncertainty is marginally significant, there is no effect on willingness-to-pay.

5. Additional Results 2: Separate Analysis on Quality Evaluation for the Advertising Content and Persuasive Content

In this section, we compare the perceived quality of content generated for campaigns and products separately. We first examine participants' stated level of satisfaction for the persuasive content for campaigns generated under different paradigms. As depicted by the baseline condition in Figure 3, participants expressed varying degrees of satisfaction with content produced under different paradigms (one-way ANOVA, p-value = 0.000). On average,

advertising content generated solely by ChatGPT-4 resulted in higher satisfaction level than content generated solely by a human expert (5.11 vs 4.83, two-sample t-test, p-value = 0.002, cohen's d = 0.19). Content generated by "augmented AI" also outperformed content generated by "augmented human" (4.93 vs 4.45, two-sample t-test, p-value = 0.000, cohen's d = 0.31). Advertising content generated solely by ChatGPT-4 resulted in higher satisfaction level than content generated solely by a human expert (5.48 vs 5.03, two-sample t-test, p-value = 0.000, cohen's d = 0.33). Advertising content generated by "augmented AI" also outperformed content generated by "augmented human" (5.51 vs 5.47, two-sample t-test, p-value = 0.61).

Similar patterns emerge when we use participants' willingness-to-pay for the content as a measure of content quality. As depicted by the baseline condition in Figure 4, participants had different willingness-to-pay for content generated under different paradigms (one-way ANOVA, p-value = 0.000, for both categories). Furthermore, on average, participants' willingness-to-pay for campaign content was slightly higher for content generated solely by AI than solely by a human expert, (4.70 vs 4.45, two-sample t-test, p-value = 0.06, cohen's d = 0.12). On the other hand, persuasive content generated by "augmented AI" achieved higher willingness-to-pay than content generated by "augmented human" (4.48 vs 3.96, two-sample t-test, p-value = 0.000, cohen's d = 0.23). On average, participants' willingness-to-pay for advertising content was the same for content generated solely by AI and solely by a human expert, (4.97 vs 4.77, two-sample t-test, p-value = 0.09, cohen's d = 0.11). On the other hand, advertising content generated by "augmented AI" also had the same willingness-to-pay as content generated by "augmented human" (5.21 vs 5.19, two-sample t-test, p-value = 0.76).

Similar patterns also emerge when we examine the level of interest after seeing the advertising content and the degree of persuasion after seeing persuasive content as measures of content quality. As depicted by the baseline condition in Figure 5, participants had different level of interest for content generated under different paradigms (one-way ANOVA, p-value for campaigns = 0.000, p-value for products = 0.01). Interestingly, for the advertising content, there was no difference between content generated by AI and humans. The level of interest in the product after seeing advertising content solely by a human expert is 4.70 vs 4.85 after seeing content generated solely by AI (two-sample t-test, p-value = 0.15). The level of interest in the

product after seeing advertising content by "augmented human" is 4.91 vs 5.02 after seeing content generated by "augmented AI" (two-sample t-test, *p*-value = 0.22). However, the degree of persuasion after seeing persuasive content generated solely by AI for campaigns is still stronger than that by a human expert (5.07 vs 4.82, two-sample t-test, *p*-value = 0.008, cohen's *d* = 0.17). The degree of persuasion after seeing persuasive content generated by "augmented AI" for campaigns is still stronger than that by "augmented human" (4.91 vs 4.47, two-sample t-test, *p*-value = 0.000, cohen's *d* = 0.28).

We repeat the above analysis for the "uninformed" condition. As depicted by the "uninformed" condition in Figure 3, on average, advertising content generated solely by ChatGPT-4 resulted in higher satisfaction level than content generated solely by a human expert (4.93 vs 4.72, two-sample t-test, *p*-value = 0.03, cohen's *d* = 0.13). Content generated by "augmented AI" also outperformed content generated by "augmented human" (4.98 vs 4.56, two-sample t-test, *p*-value = 0.000, cohen's *d* = 0.28). Advertising content generated solely by ChatGPT-4 resulted in higher satisfaction level than content generated solely by a human expert (5.30 vs 4.87, two-sample t-test, *p*-value = 0.000, cohen's *d* = 0.29). Advertising content generated by "augmented AI" also outperformed content generated by "augmented human" (5.48 vs 5.4, two-sample t-test, *p*-value = 0.30).

Next we examine participants' willingness-to-pay for the content as a measure of content quality. As depicted by the baseline condition in Figure 4, on average, participants' willingness-to-pay for campaign content was the same for content generated solely by AI and solely by a human expert, (4.51 vs 4.50, two-sample t-test, *p*-value = 0.98). On the other hand, persuasive content generated by "augmented AI" achieved higher willingness-to-pay than content generated by "augmented human" (4.57 vs 4.27, two-sample t-test, *p*-value = 0.02, cohen's *d* = 0.14). On average, participants' willingness-to-pay for advertising content was higher for content generated solely by AI than solely by a human expert, (5.00 vs 4.68, two-sample t-test, *p*-value = 0.007, cohen's *d* = 0.17). On the other hand, advertising content generated by "augmented AI" had the same willingness-to-pay as content generated by "augmented human" (5.12 vs 5.04, two-sample t-test, *p*-value = 0.45).

Last but not least, we examine the level of interest after seeing the advertising content and the degree of persuasion after seeing persuasive content as measures of content quality. As depicted by the baseline condition in Figure 5, for the advertising content, there was no difference between content generated solely by AI and humans. The level of interest in the product after seeing advertising content solely by a human expert is 4.88 vs 4.78 after seeing content generated solely by AI (two-sample t-test, p-value = 0.29). The level of interest in the product after seeing advertising content by "augmented human" is 5.03 vs 4.96 after seeing content generated by "augmented AI" (two-sample t-test, p-value = 0.46). Furthermore, the degree of persuasion after seeing persuasive content generated solely by AI for campaigns is also the same as content generated solely by a human expert (4.82 vs 4.75, two-sample t-test, p-value = 0.50). However, the degree of persuasion after seeing persuasive content generated by "augmented AI" for campaigns is still stronger than that by "augmented human" (4.96 vs 4.50, two-sample t-test, p-value = 0.000, cohen's d = 0.30).

6. Additional Results 3: Examining Potential Bias for the Advertising Content and Persuasive Content Separately

Within a content generation paradigm, we regress the DVs on the condition (1 = informed, 0 = baseline), task category (1 = product, 0 = campaign), and their interaction, with task fixed effects. We report the coefficient on the interaction term to examine any potential bias. A positive (negative) coefficient suggests people perceive the content generated under the paradigm as higher (lower) quality after knowing the identity of the content creator. Overall, we observe human favoritism, especially in the content generated for products.

We first report the coefficient for content generated for the **campaigns**.

Paradigm	DV	Coefficient	Cluster-Robust Standard Error	P-value
Human	Satisfaction	0.08	0.04	0.07
Human	WTP	0.20	0.08	0.01

AI	Satisfaction	-0.10	0.08	0.29
AI	WTP	-0.10	0.08	0.20
Augmented Human	Satisfaction	0.04	0.04	0.34
Augmented Human	WTP	0.20	0.12	0.09
Augmented AI	Satisfaction	0.15	0.08	0.07
Augmented AI	WTP	0.22	0.12	0.07

Table 1 *The coefficient is the coefficient of the interaction term in the regression. Cluster-robust standard errors and the p-values are also reported.*

We first report the coefficient for content generated for the **products**.

Paradigm	DV	Coefficient	Cluster-Robust Standard Error	P-value
Human	Satisfaction	0.10	0.04	0.02
Human	WTP	0.16	0.04	0.000
AI	Satisfaction	0.003	0.03	0.91
AI	WTP	0.09	0.07	0.23
Augmented Human	Satisfaction	-0.05	0.08	0.57
Augmented Human	WTP	-0.01	0.08	0.86

Augmented AI	Satisfaction	0.01	0.04	0.82
Augmented AI	WTP	0.02	0.06	0.73

Table 1 *The coefficient is the coefficient of the interaction term in the regression. Cluster-robust standard errors and the p-values are also reported.*