Pixel Pioneers

A/B Testing the Impact of Generative AI across Creative Domains

CMU- 95819/ A/B Testing, Design and Analysis

Faraz Ahmad | faraza@andrew.cmu.edu |
Atharva Kulkarni | askulka2@andrew.cmu.edu
Karl Cao | yixiangc@andrew.cmu.edu

Carnegie
Mellon
University

Introduction: What are we testing?

- Measure audience engagement with AI-generated social media posts compared to human-authored posts.

- Evaluate the effectiveness of Generative AI in creating product descriptions that drive interest and conversions.

- Assess the reception and perceived quality of artistic content generated by AI versus human creators.

- Identify potential variations in audience responses based on demographics and preferences.

Hypothesis

- Generative AI-generated content will exhibit superior audience engagement, effectiveness in driving interest and conversions, and perceived quality compared to content created by humans.

Experimental Design

- Treatment Groups: Participants were exposed to either Generative AI-generated content (treatment group) or content created by human authors (control group).

- Randomization: The assignment of participants to treatment and control groups was randomized by 50% to ensure unbiased representation in treatment and control groups. This was programmed into qualtrics.

Outcome Variables of interest include:-

- Engagement metrics (likes, comments, shares)
- Conversion rates (Persuasiveness to buy a product)
- Perceived quality scores, emotional response, artistic detailing (**art detailing, emotional response, clarity and coherence of descriptions, relevance and accuracy of content**)
- Demographics summary

Treatment and Control Groups

Treatment and Control Groups

Apple AirPods (2nd Generation) Wireless Ear Buds • Bluetooth Headphones • with Lightning Charging Case Included • Over 24 Hours of Battery Life • Effortless Setup for iPhone

Visit the Apple Store

4.7 ★★★★☆   607,162 ratings | Search this page

Amazon's Choice in Earbud & In-Ear Headphones by Apple

10K+ bought in past month

-30% $89⁹⁹

List Price: $129.00 ⓘ

Or **$16.62** /mo (6 mo). Select from 1 plan

✓prime Overnight

FREE Returns ⌄

Get a $100 Gift Card: Pay $0.00 $89.99 upon approval for Prime Visa. No annual fee.

About this item

- The sleek, white design is not only stylish but also houses a powerful H1 chip that enables faster wireless connection to your devices. Enjoy seamless switching between devices without any hassle. The Lightning Charging Case, included in the package, ensures your AirPods are always charged and ready to go.
- Immerse yourself in high-quality sound with the Apple AirPods (2nd Generation) Wireless Ear Buds. Designed to keep up with you all day, these Bluetooth headphones offer over 24 hours of battery life for multi-day use on a single charge. With an effortless setup for your iPhone, they are ready to use as soon as you take them out of the case.
- The intuitive design includes touch controls, allowing you to play or skip forward with a double-tap when listening to music or podcasts. Apple AirPods are equipped with high-quality microphones for clear calls and voice assistant interaction, ensuring you stay connected and in control.
- Whether you're at the gym, commuting, or just relaxing, the Apple AirPods are designed to fit comfortably in your ears for long periods. They're the perfect companion for your active lifestyle, delivering the signature Apple sound and convenience that you love. Shop now and elevate your listening experience!

Apple AirPods (2nd Generation) Wireless Ear Buds • Bluetooth Headphones • with Lightning Charging Case Included • Over 24 Hours of Battery Life • Effortless Setup for iPhone

Visit the Apple Store

4.7 ★★★★☆   607,162 ratings | Search this page

Amazon's Choice in Earbud & In-Ear Headphones by Apple

10K+ bought in past month

-30% $89⁹⁹

List Price: $129.00 ⓘ

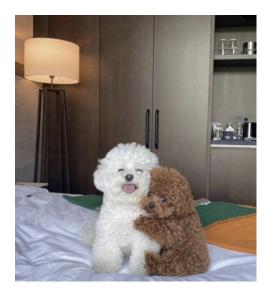Or **$16.62** /mo (6 mo). Select from 1 plan

✓prime Overnight

FREE Returns ⌄

Get a $100 Gift Card: Pay $0.00 $89.99 upon approval for Prime Visa. No annual fee.
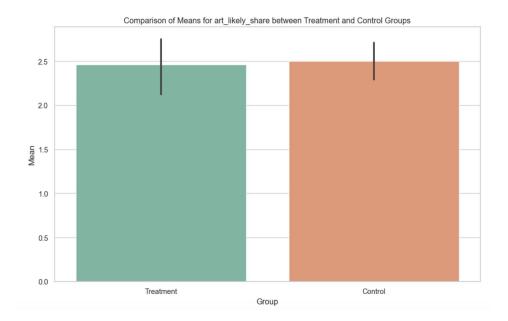
About this item

- HIGH-QUALITY SOUND — Powered by the Apple H1 headphone chip, AirPods (2nd generation) deliver rich, vivid sound.
- EFFORTLESS SETUP — After a simple one-tap setup, AirPods are automatically on and always connected. They sense when they're in your ears and pause when you take them out. And sound seamlessly switches between your iPhone, Apple Watch, Mac, iPad, and Apple TV.
- VOICE CONTROL WITH SIRI — Just say "Hey Siri" for assistance without having to reach for your iPhone.
- 24-HOUR BATTERY LIFE — More than 24 hours total listening time with the Charging Case.
- AUDIO SHARING — Easily share audio between two sets of AirPods on your iPhone, iPad, iPod touch, or Apple TV.
- LEGAL DISCLAIMERS — This is a summary of the main product features. See "Additional information" to learn more.
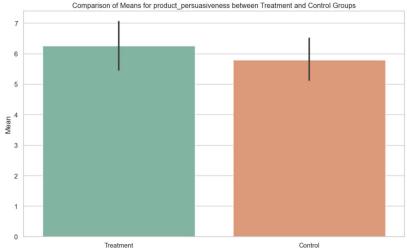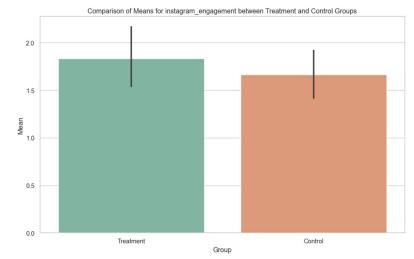
Treatment and Control Groups

Caption:

Friends fur-ever

Besties for the resties

Mean Comparison and T-tests

Regression Analysis - Art

```
                          OLS Regression Results
==============================================================================
Dep. Variable:         art_likely_share   R-squared:                       0.001
Model:                              OLS   Adj. R-squared:                 -0.021
Method:                   Least Squares   F-statistic:                   0.04364
Date:                Wed, 28 Feb 2024    Prob (F-statistic):              0.835
Time:                        17:13:40    Log-Likelihood:                 -49.340
No. Observations:                  48    AIC:                             102.7
Df Residuals:                      46    BIC:                             106.4
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const             2.5000      0.141     17.727      0.000       2.216       2.784
Group_Treatment  -0.0417      0.199     -0.209      0.835      -0.443       0.360
==============================================================================
Omnibus:                        6.785   Durbin-Watson:                   1.275
Prob(Omnibus):                  0.034   Jarque-Bera (JB):                6.936
Skew:                          -0.909   Prob(JB):                       0.0312
Kurtosis:                       2.599   Cond. No.                        2.62
==============================================================================
```

Regression Analysis - Instagram Captions

```
                          OLS Regression Results
==============================================================================
Dep. Variable:     instagram_engagement   R-squared:                       0.014
Model:                              OLS   Adj. R-squared:                 -0.007
Method:                   Least Squares   F-statistic:                    0.6765
Date:                  Wed, 28 Feb 2024   Prob (F-statistic):              0.415
Time:                        17:13:40   Log-Likelihood:                 -50.102
No. Observations:                    48   AIC:                             104.2
Df Residuals:                        46   BIC:                             107.9
Df Model:                             1
Covariance Type:              nonrobust
==============================================================================
                     coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const              1.6667      0.143     11.632      0.000       1.378       1.955
Group_Treatment    0.1667      0.203      0.822      0.415      -0.241       0.575
==============================================================================
Omnibus:                        4.885   Durbin-Watson:                   2.089
Prob(Omnibus):                  0.087   Jarque-Bera (JB):                2.589
Skew:                           0.321   Prob(JB):                        0.274
Kurtosis:                       2.061   Cond. No.                         2.62
==============================================================================
```

Regression Analysis - Product Description

```
                           OLS Regression Results
==============================================================================
Dep. Variable:     product_engagement_interest   R-squared:                  0.007
Model:                                      OLS   Adj. R-squared:            -0.015
Method:                           Least Squares   F-statistic:               0.3239
Date:                         Wed, 28 Feb 2024   Prob (F-statistic):         0.572
Time:                                 17:13:40   Log-Likelihood:            -34.502
No. Observations:                           48   AIC:                         73.00
Df Residuals:                               46   BIC:                         76.75
Df Model:                                    1
Covariance Type:                     nonrobust
==============================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const             2.1667      0.104     20.928      0.000       1.958       2.375
Group_Treatment   0.0833      0.146      0.569      0.572      -0.211       0.378
==============================================================================
Omnibus:                       1.403   Durbin-Watson:                 1.353
Prob(Omnibus):                 0.496   Jarque-Bera (JB):              0.915
Skew:                          0.336   Prob(JB):                      0.633
Kurtosis:                      3.084   Cond. No.                      2.62
==============================================================================
```

Analysis & Insights – Product Description

Statistical Comparison: No significant differences detected between AI and human-generated content.

User Perception: Similar quality perceived in AI and human descriptions regarding key aspects like persuasiveness and creativity.

Quality Metrics: AI matches humans on accuracy, engagement, and authenticity in product descriptions.

Potential Impact: Users may not distinguish between AI and human-created content in e-commerce settings.

Analysis & Insights – Instagram Captions

Creativity & Uniqueness: AI captions rated lower than human counterparts, affecting user engagement.

Accuracy Importance: Regression shows creativity and accuracy are key to engagement.

Mitigating Engagement Drop: Enhanced AI accuracy and creativity could offset lower engagement rates.

Strategic Consideration: Human touch is critical for high-engagement Instagram content.

Analysis & Insights – Artist Drawings

AI vs. Human Art: While explicit data is lacking, AI art may rival human art in technical aspects..

Creativity & Resonance: AI might not yet capture the deep creative essence and emotional resonance found in human-created art.

Engagement Factors: Engagement with AI-generated art could vary, mirroring trends seen in other content types.

Strategic Consideration: Human touch is critical for high-engagement Instagram content.

Recommendations & Conclusion:

- **Product Description** - Given the comparable performance, it's recommended to use AI-generated descriptions to supplement human efforts, especially in high-volume scenarios. However, for products requiring unique or emotionally resonant descriptions, human input remains invaluable.

- **Instagram Captions** - Focus on human creativity for engagement-critical posts, use AI-generated captions as a supplementary tool, and invest in AI development to improve creativity and accuracy, aiming to narrow the engagement gap between AI-generated and human-generated captions.

- **Artist Drawings -** Utilize AI for bulk content production and initial drafts; incorporate human creativity for final touches on high-impact content.

Improvements for the experiment

- Depending on the nature of the content, **longer exposure periods** may be needed to capture more nuanced reactions

- Evaluate the **possibility of increasing the sample size** to enhance the statistical power of the experiment. A larger sample size can provide more robust insights into the impact of Generative AI on user engagement.

- Include **more variability in the demographics** to ascertain if people without technical expertise are able to distinguish between human and AI generated content.

- Involve **different styles, formats, or themes** to understand if certain content aspects perform better than others