

The Psychology of Color in Advertising: A Survey-Based Experiment

Aishwarya Jayant Rauthan, Gunjan Sharma, Himani Tyagi, Jasmine Gohil, Jenil Shah

I. Introduction

In the ever-evolving age of digital media, understanding the impact of visual elements on user perception is essential for businesses striving to capture and retain audience attention. In this context, our experiment dives deep into the intriguing dependence between the color palette of advertisements and user perception. We sought to explore the distinct effects of black-and-white versus colored ads on how users perceive and engage with content. The choice of color in advertising has long been recognized as a powerful tool, capable of evoking emotions, influencing brand recall, and shaping overall user experience. As businesses navigate the route of visual communication, our study aims to contribute valuable insights into the preferences and responses of users when exposed to different color schemes, providing marketers and advertisers with actionable information to refine their strategies.

The experimental design involved randomization of 119 users into two groups, each exposed to either black and white (treatment group) or colored (control group) advertisements. Through this we aimed to measure the impact of color on user perception, ensuring a focused examination of its role in shaping attitudes and preferences. This report encapsulates the findings of our experiment, shedding light on ways in which color choice within advertisements can influence the digital landscape.

II. Methods

Participants

Since the goal of the experiment was to explore the distinct effects of black-and-white versus colored ads on how users perceive and engage with content, we decided people over the age of 18 would make the best subjects. We also want to check the effect of the age brackets on our experiment. Our initial method to recruit the participants was by posting the Qualtrics survey link in our college group chat, Instagram stories, random other group chats. We also personally messaged our friends and family. People who were interested in participating were required to fill out a Qualtrics survey form that detailed the name, age, country, followed by the digital media advertisements.

There were a few biases that emerged when recruiting only people we knew personally. The first and most obvious is that the sample is not representative of college students across the United States. Since BU is in the northeast and plenty of students in our experiment were from BU, few from New York University and few from University of Maryland. We also had a disproportionate number of people sign up for the study since we personally approached a lot of classmates. Furthermore, the inclusion of participants from outside the United States, while potentially enriching the diversity of perspectives, introduces cultural and contextual variables that may confound the interpretation of our findings within the U.S. college student population. Fortunately, we were able to randomize everyone into the control or treatment

groups to avoid selection bias in the study. We achieved a 50-50 split in the control and treatment group.

Randomization

To decide which participants would be a part of the control or treatment group, we decided to randomize by blocking on certain demographic features to achieve greater statistical power. Using the answers we received from our survey, we blocked participants based on their genders. Upon doing that, we then blocked on their location within each of these groups. Finally, we blocked the participants based on age groups in each category.

Procedure

We began our experiment by creating a survey using Qualtrics, where we created two separate surveys containing the same set of images. The key difference between the control and the treatment group was the colour scheme of these images – control group received coloured images while the treatment group received the same images in black and white format. All of these images were ad campaigns from different companies like – Rado, Halls, Balenciaga, Ikea, Barbie, Pedigree, and Coca-Cola. We sourced the images for these campaigns through open sources like - Pinterest and Google Images.

Based on the appeal of these images, people were asked to rate these images on a scale of 1 to 5 for a comparative analysis. We also made sure to include a variety of questions (eg. How much effort did you feel you had to put in to pay attention to the ad?) to capture various factors and nuances that contribute to visual appeal of a campaign. Additionally, we also added questions related to their demographics as a part of our survey to further provide depth to our analysis. Below is a table detailing the breakdown of the questions.

Order	Question
Q1	What is your name?
Q2	What is your age?
Q3	What is your Gender?
Q4	List of Countries
Q5	How engaging or interesting did you find the ad?
Q6	How likely are you to consider buying the product advertised in the ad?
Q7	How visually appealing did you find the ad?
Q8	How clearly did the ad communicate its message?
Q9	How would you rate the overall impression left by this image?
Q10	How much effort did you feel you had to put in to pay attention to the ad?
Q11	How much did you enjoy the overall style and aesthetics?

For distribution of these surveys, we shared the link to our Qualtrics form on various social media platforms like Instagram and WhatsApp to reach out to as many people as possible. Through Qualtrics' in-built functionality, we were able to randomize between the control group and the treatment group while using a common survey link.

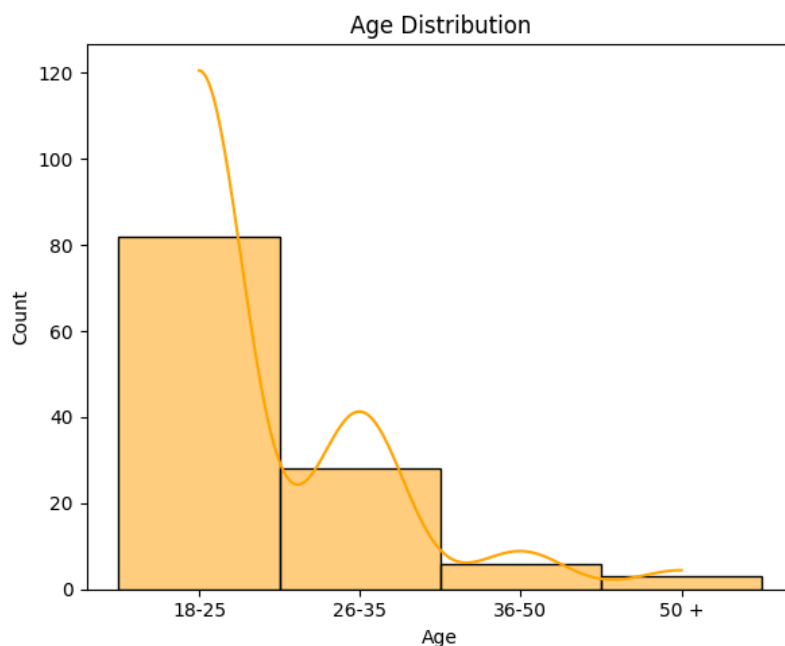
To make sure that the experiment would run smoothly, we ran a pilot study of the experiment with a few friends to check if the randomization was working fine. For this part, we sent the link to a handful of people and asked them what form they received. This way, we found out if the randomization was working properly by communicating to these people regarding the form they received and seeing if it was a decent distribution. By trying a pilot run for this experiment, we discovered that the randomization between control and treatment group was working fine, thus ensuring us of a smooth run of the main experiment.

III. Data Analysis

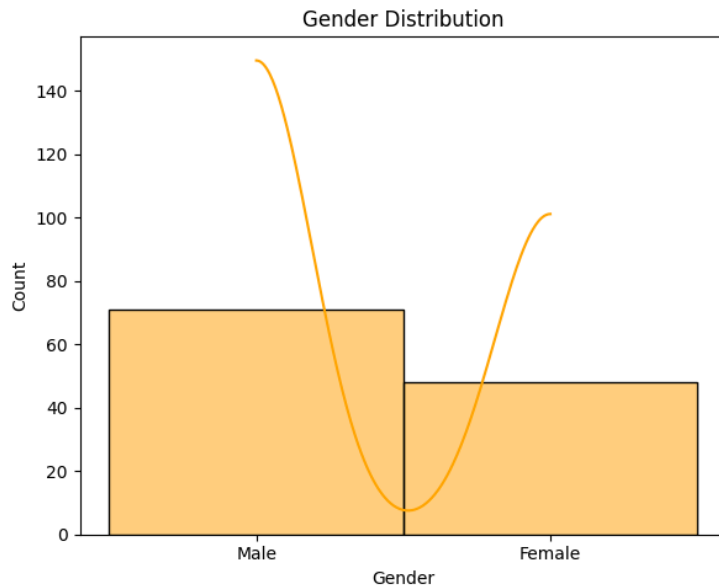
	Treatment	Control
Number of Participants	61	58

The experiment involved 119 participants (61 + 58) with an almost balanced distribution between the treatment and control groups.

Demographics of the Participants



While the majority of participants in our survey fall within the 18-25 age range, we strived for a fair distribution across age groups, ensuring the inclusion of younger adults and some representation of slightly older individuals.



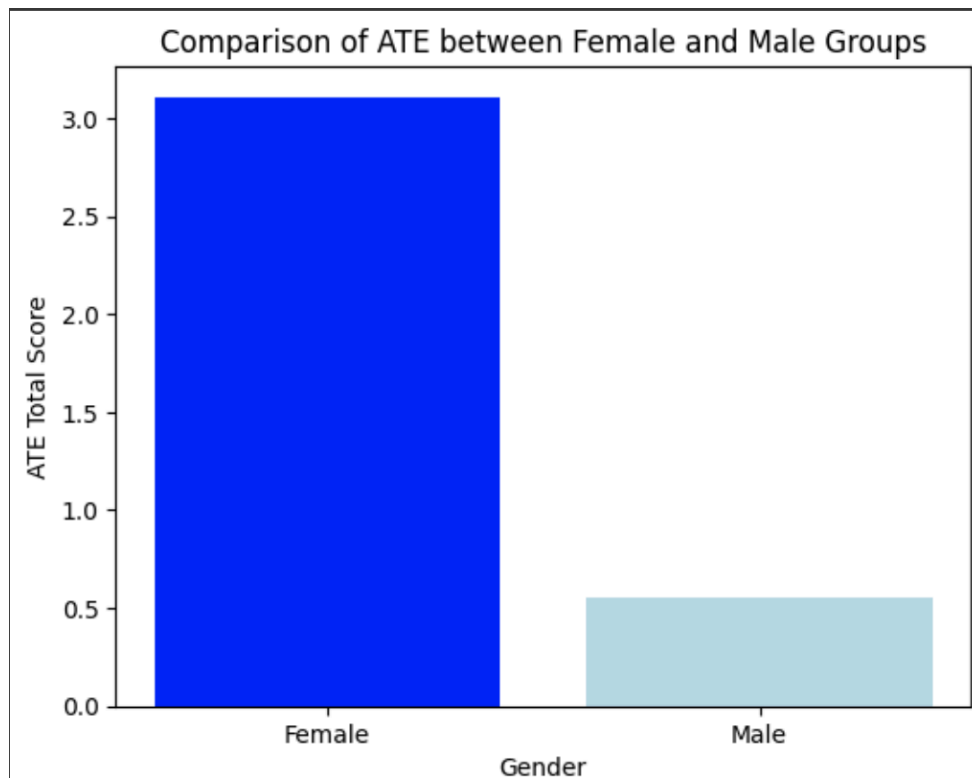
The Gender Distribution depicts that the data that male participants outnumber their female counterparts, indicating a demographic skew towards males in our survey responses. This strategy was intended to include a broad spectrum of gender identities, mirroring our approach to age diversity, which emphasized inclusivity of not only younger adults but also those from moderately older age cohorts.

IV. Average Treatment Effect and Regression

Estimated Average Treatment Effect

The estimated Average Treatment Effect is 1.669, this finding suggests that seeing color ads (control group) was associated with a slightly higher average total score compared to seeing black and white ads (treatment group). The interpretation needs to be nuanced by considering statistical significance, effect size, and individual variability.

Conditional Average Treatment Effect



Interpretation:

Male

The average Treatment Effect (ATE) on total scores for males indicates that, on average, males in the treatment group scored 0.5529 points higher than males in the control group. This positive ATE suggests a potential positive impact of the treatment on the total scores for males.

Female

The average Treatment Effect (ATE) on total scores for females indicates that, on average, females in the treatment group scored 3.1084 points higher than females in the control group. This positive ATE suggests a substantial positive impact of the treatment on the total scores for females, potentially indicating a more pronounced effect compared to males.

Regression Analysis

1. Regression of the main effect on outcomes:

Dependent variable: total_score	
(1)	
Intercept	21.931 ^{***} (0.527)
Treatment_Group	-1.669 [*] (0.854)
Observations	119
R ²	0.032
Adjusted R ²	0.023
Residual Std. Error	4.650 (df=117)
F Statistic	3.816 [*] (df=1; 117)
Note:	* p<0.1; ** p<0.05; *** p<0.01

Our analysis reveals that participants exposed to black and white ads (Treatment Group) scored 1.669 points lower on the total score compared to those who saw color ads (Control Group). This effect is not statistically significant (p-value < 0.1).

The R-squared value is 0.032, indicating that the independent variable 'Treatment_Group' explains approximately 3.2% of the variance in 'total_score'. The residual error is 4.650, indicating the typical amount of error or variability in the model's predictions.

In simpler terms, seeing black and white ads seems to be associated with a decrease in overall score compared to color ads.

2. Regression of the main effect with good control variables

Dependent variable: total_score		
	(1)	(2)
Gender[T.Male]		-0.951 (0.869)
Intercept	21.931 ^{***} (0.527)	22.456 ^{***} (0.677)
Treatment_Group	-1.669 [*] (0.854)	-1.585 [*] (0.862)
Observations	119	119
R ²	0.032	0.042
Adjusted R ²	0.023	0.025
Residual Std. Error	4.650 (df=117)	4.646 (df=116)
F Statistic	3.816 [*] (df=1; 117)	2.564 [*] (df=2; 116)
Note:	* p<0.1; ** p<0.05; *** p<0.01	

The decrease in standard error and increase in R-squared after controlling for the gender covariate indicate that the model has become more precise and explanatory. This suggests

that the inclusion of the gender covariate has improved the accuracy and comprehensiveness of the regression analysis.

Additionally, the inclusion of gender as a covariate did not substantially change the treatment effect estimate, but it increased the precision of the estimate, resulting in a more precise treatment effect estimate compared to the previous regression.

Power Analysis

Cohen's D of 0.355

Based on Cohen's d, seeing black and white ads in the experiment seems to be associated with a small decrease in total score compared to color ads. It suggests that black and white ads might have a slightly negative impact on the total score compared to color ads, but the effect size is relatively small.

Power T-test - 48.33%

Power indicates the probability of a statistical test correctly rejecting the null hypothesis (no difference between groups) when there truly is a difference of the observed size (Cohen's d). In this case, a power of 0.4833 suggests a moderate power. This means there's a 48.33% chance the t-test would have detected the observed effect size (small) as statistically significant, assuming it truly exists in the population. It also suggests there's a significant chance the t-test might have failed to detect a real effect due to the sample size or the small effect size itself.

V. Limitations

We found out there might be several different ways to improve our analysis.

The sample size of 120 participants may constrain the ability to detect subtle color effects, diminishing the study's statistical power and generalizability to a broader population. Moreover, relying on self-reported data introduces subjectivity, as participant responses may be influenced by memory, attention, or individual interpretations. The survey format, particularly if participants are aware of their involvement in an experiment, poses a risk of bias, potentially altering their behavior and responses. Additionally, the study's focus on color effects without considering product type or overall ad design limits the contextual understanding of how colors may interact with different advertising elements. Finally, conducting experiments in a controlled environment may not fully capture the complexities of real-world advertising exposure, affecting the external validity of the study. Acknowledging and addressing these limitations is crucial for a nuanced interpretation of the study's outcomes and for informing future research in the field.

VI. Conclusion

The analysis we conducted demonstrated that the psychology of color in advertising plays a significant role in advertisements. The A/B testing revealed that the treatment group shows higher impressions for the Barbie ad, mostly ranging from moderate to high (3 to 5 on the

scale). The regression analysis further confirmed that participants exposed to black and white ads (Treatment Group) scored 1.669 points lower on the total score compared to those who saw color ads (Control Group), and the effect is not statistically significant ($p\text{-value} < 0.1$), indicating a reliable difference between the groups. The addition of covariates such as Overall impression, style, enjoyment, and visual appeal increased the precision of the results. This indicates a more robust understanding of the factors influencing ad engagement. Overall, the study underscores the importance of integrating image color in advertisements to attract and engage the demographic effectively.

APPENDIX A