
Will Judging a Cookie by the Box Change our Preferences?

Aarushi Somani, Austin Tao, Eleanor Kim, Grayson Meckfessel, and Lauren Murai

Abstract

In the United States, consumerism is a major part of the socioeconomic system and has become a heavily researched field, with studies being conducted in psychology, behavioral economics, and more. Specifically, a lot of research has been done regarding advertisement, branding, and other means of encouraging customers to buy products and to enjoy the product. We intend to quantify how different types of information (verbal versus visual) impact participants' taste preferences with two different off-brand Oreo cookies. We perform three phases of experimentation: the first tests if the order in which people receive cookies has any effect, the second tests how verbal information affects participants, and the third tests how verbal information in addition to visual information (via packaging and branding) affects participants. Through a difference-in-difference analysis, our results show that, on average, there is a statistically significant difference in taste rating between participants who received visual information about the cookies versus those who did not. Through demonstrating this causal effect, this study highlights how impactful seemingly innocuous details such as packaging are.

Introduction

Through our experiment, we hope to quantify the different effects verbal information and visual information have on participants' taste preferences regarding two different off-brand Oreo cookies. Our initial hypothesis was that giving participants verbal information about the quality of the cookies would have a statistically significant effect on their taste preferences, in particular that participants would rate cookies higher when they received verbal information regarding their quality and nutrients as compared to when they receive no information at all. We also hypothesized that visual information would have a smaller effect on taste perception compared to the effects of verbal information. These hypotheses were formed through demonstrated results in the existing literature about information and its efficacy on taste perception.

Literature Review

Schouteten et. al. demonstrated in 2015 that different health labels on cheese significantly impacted both taste perception and health perception of participants [1]. On a 7-point scale (1 being extreme dislike and 7 being extreme like), cheese that was labeled as "light" was rated significantly lower (4.47 ± 1.32 , $p < 0.001$) compared to a label that just mentions "cheese" (5.07 ± 1.20), despite both cheeses being identical. Other studies have done similar on different facets of the product, such as "perceived healthiness" and "perceived enjoyment". In 2015, it was shown that two identical cereals could have an almost 10% difference in perceived healthiness simply by using the label "fruit sugar" as opposed to "sugar" [2].

Experiment Design

In our experiment, the product in question is a knock-off oreo cookie. Originally our hypothesis was that verbally presenting information about the quality of ingredients of some unfamiliar brand “sandwich creme cookie” would cause participants to rate their perception of the cookie higher. To make the design more interesting, we would give participants two different knock-off oreos, have them eat and rate both one cookie at a time, and then take the difference between their ratings. One of the cookies, “Cookie A,” would have higher quality ingredients and be more expensive. The other cookie, “Cookie B,” would have lower quality ingredients and be less expensive. In theory the quality of ingredients is somehow reflected in the taste of the cookies. In the control group, participants would be blind to this information. They would freely rate both cookies without knowing the brand, quality of ingredients, or what the packaging looks like. In the treatment group, participants would be told that Cookie A had higher quality ingredients than Cookie B. Then they would eat and rate both cookies (also without seeing the packaging or brand). Our hypothesis was that the difference in ratings between Cookie A and Cookie B would be higher in the treatment group, when given information about the cookies’ quality, than in the control group, not given any information.

In order to isolate the causal effect of “information”, we needed to consider all the moving pieces in our experimental design that could influence the outcome. For example, what if participants see the packaging or recognize the brand and that’s what causes their preferences to be influenced? What if the first cookie they taste sets a benchmark for how they perceive the second cookie they taste? We also need to account for the variations that might occur in our samples. In order to perform our unpaired t-test analysis, we need to assume not only that our sample is representative of the population in question, but more importantly we need to assume that our treatment and control samples are equal. If the treatment group is disproportionately older than our control group, then we can’t assume comparable ratings across the two groups. For example, older participants may be more influenced by information about ingredients than younger participants.

Accounting for concerns regarding isolating causal effects, three phases of experiments and five separate analyses were designed for our study. In the first phase, we answer the question: *Does the order in which we give participants a cookie affect their ratings?* In the second phase we answer the question: *Does providing information to participants about the cookies’ quality of ingredients (blind to the brand/packaging) affect their ratings?* Then in the third phase we answer the question: *Does providing information to participants about the cookies’ quality of ingredients while showing them the package they come in affect their ratings?* And then finally, in order to isolate the effects of the packaging from providing verbal information: *Does showing the participants the package that the cookies come in (blind to any additional verbal*

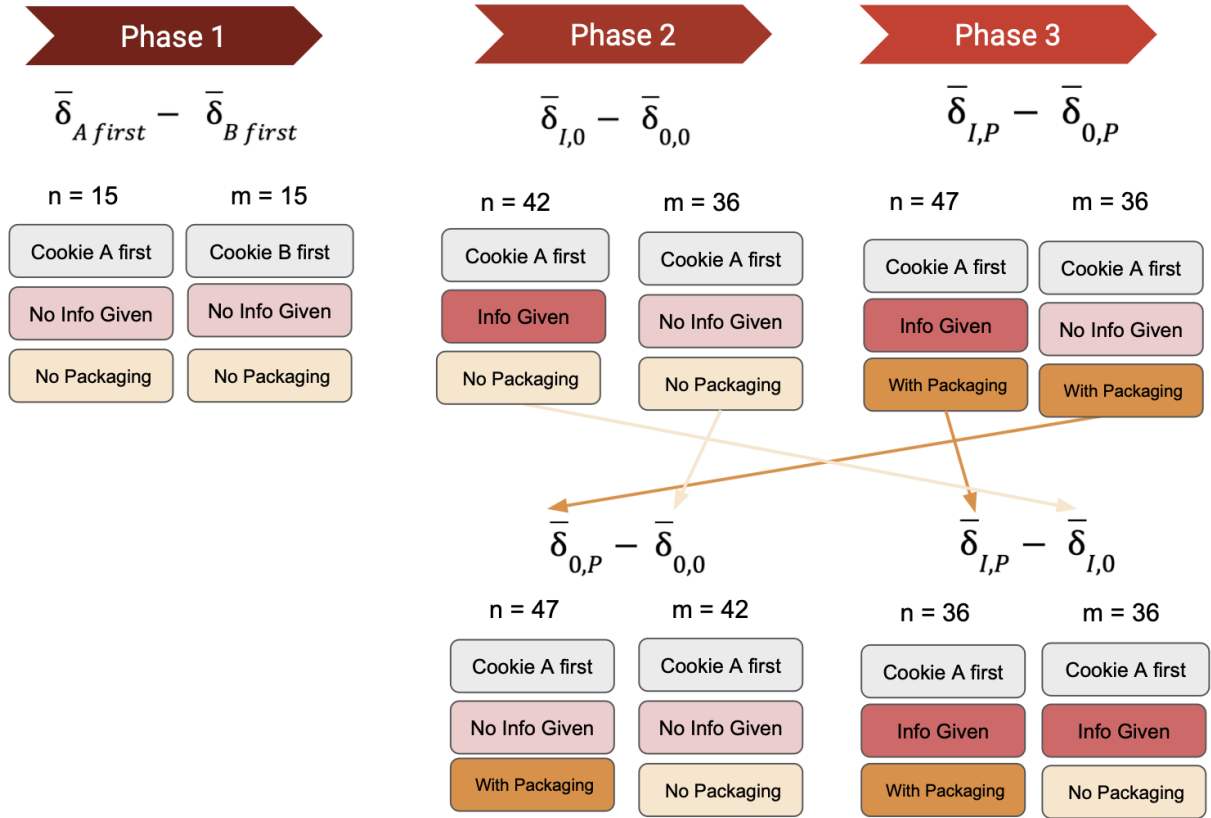
information) affect their ratings? The fourth question recycles data we collected in the second and third phase by cross comparing the treatment groups from phases 2 and 3 and then separately comparing the control groups from phases 2 and 3.

Many of our participants were the students of our group members who were graduate student instructors. During our lab sections we randomly assigned students to a treatment and control group during one of our phases. This meant that we were sampling from a population of college-aged undergraduate statistics and economics students. Across the three phases, we have six experimental groups total (three phases, treatment and control groups). We randomly assigned our lab sections or participants to one of our phases so that each student participated in our experiment once.

In our analysis we collected data on participants' ratings for each cookie. Each participant had to eat Cookie A and rate out of 10 the taste, texture, and perceived quality of Cookie A, then do the same for Cookie B. Let's call the respective ratings for Cookie A and Cookie B in a single category x^A and x^B . Then we calculate the observed difference in ratings between Cookie A and Cookie B; let's call this $\delta = x^A - x^B$. So then our unpaired t-test for difference in means compares the mean rating difference in the treatment $\bar{\delta}_T$ with the mean rating difference in the control $\bar{\delta}_C$. Below is an example of the notation for the mean treatment rating difference for the sample that received verbal information about the cookies but did not see the packaging (treatment group in phase 2) .

$$\overline{\left\{ \left| x_{I,0,i}^A - x_{I,0,i}^B \right| \right\}}_{i=1}^n = \bar{\delta}_{I,0}$$

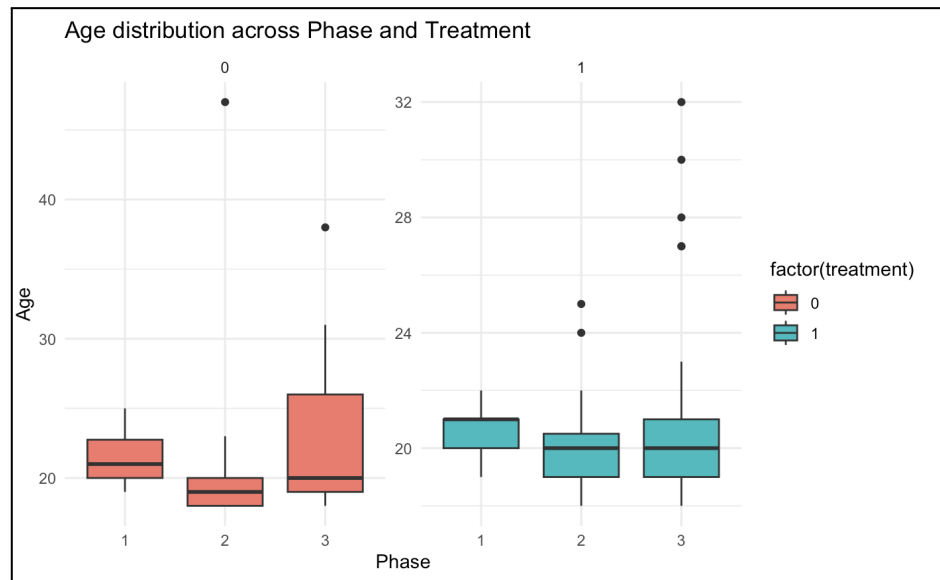
The structure of our three experimental phases, including the additional two analyses for question four, are outlined below. Note that I denotes that information was given (else 0), P denotes that packaging was shown (else 0), and n and m are the sample sizes of the control and treatment group respectively. Initially the motive behind our experiment was to lure participants into our study with cookies. This proved to be an effective way of accumulating a total of 191 participants across our study. In our power analysis, we calculated that we would need a sample size of 170 participants in order to achieve a high power of about 80% under the more conservative scenario considered.



In order to conduct this unpaired t-test, we must assume our control and treatment groups are sampled from the same population. In the survey we administer to collect data on participants' preferences, we also ask them a series of questions that can help us determine whether or not we can assume our samples are equal across various categories. We ask participants questions on age, dietary preferences, taste preferences, how hungry they are. The table below demonstrates that we cannot conclude that plots demonstrate the difference between the samples in some of the surveyed areas, compared across phases. We can see that there are no outstanding differences in our samples, as all p-values are > 0.05 (See Table 1). The first five variables are categorical variables; p-values are calculated with a chi-squared test. A t-test compares age between treatment and control (See box plots on next page).

Table 1: P-values for difference between treatment and control

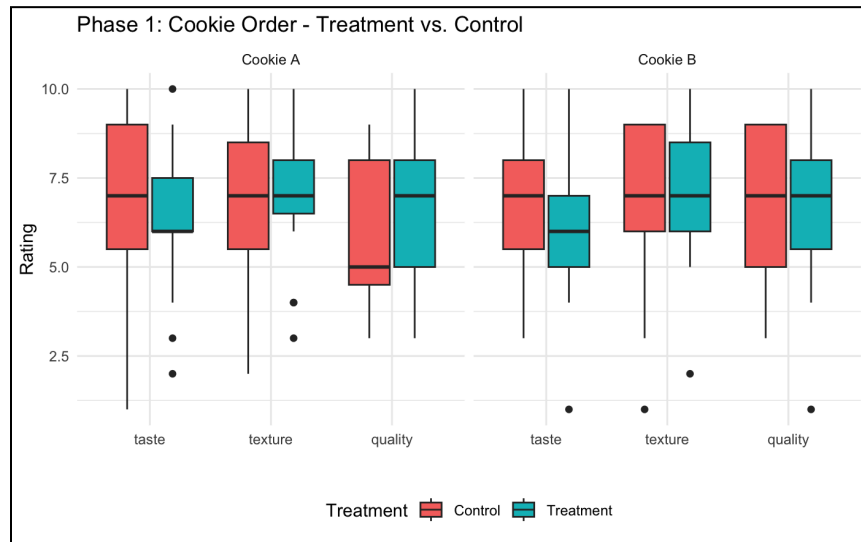
	sweet_savory	hunger	vegan	dairy	exercise	age
Phase 1	0.2722536	0.9868822	1.0000000	0.5954020	0.1918036	0.1271328
Phase 2	0.1204053	0.9352029	0.9381056	0.6450083	0.1347465	0.9239979
Phase 3	0.2618257	0.7268042	0.8135148	0.1825082	0.0921306	0.4469166



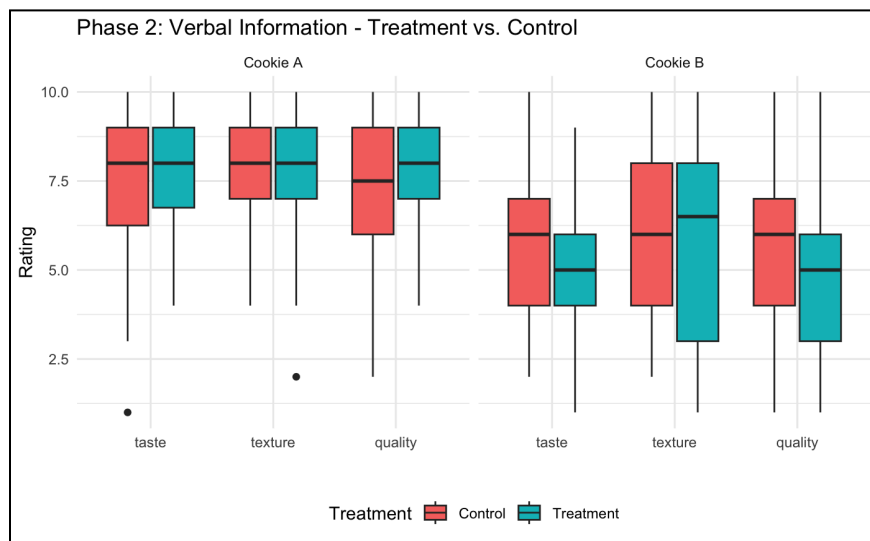
Results

In this section, we present a comprehensive analysis of the results derived from our experimental study on consumer preferences. Phase 1 of the experiment scrutinizes the possible effects of the order in which the cookies are presented, while Phase 2 delves into the influence of verbal information alone. Phase 3 extends this exploration by incorporating the visual element of packaging into the information provided. This analysis serves not only to confirm or refute our hypotheses but also to enhance our understanding of the interplay between sensory information and consumer judgment.

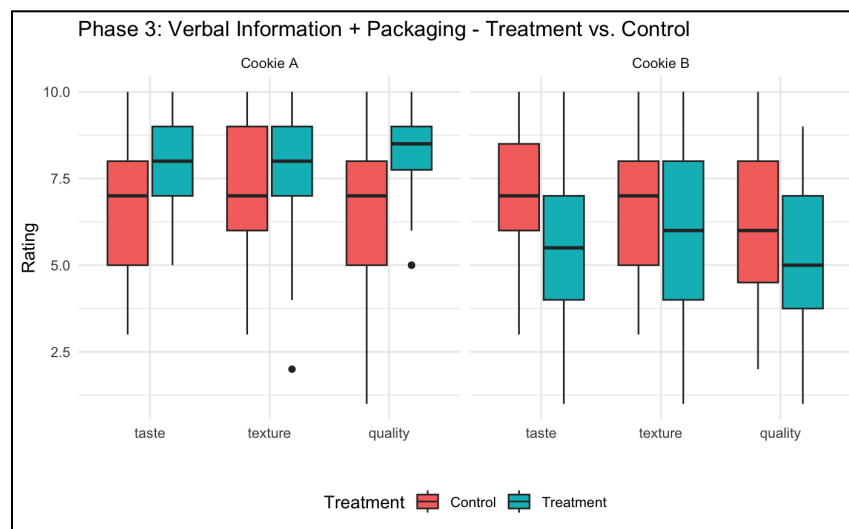
In Phase 1 of our experiment, the box plots revealed the differences in participant responses between the control and treatment groups across various attributes of Cookie A and Cookie B. Notably, the median taste rating for Cookie A in the treatment group was slightly lower at 6 compared to 7 in the control group, with a narrower interquartile range (IQR) and several outliers indicating more varied individual preferences. Similarly, the median quality rating for Cookie A significantly improved in the treatment group, moving from 5 in the control group to 7 in the treatment group, with both scenarios maintaining a consistent spread in the IQR.



In Phase 2 of our experiment, we found that for Cookie A, both groups maintained a consistent median rating of 8 for taste and texture, with slight variations in the interquartile range (IQR) indicating a compact consensus among participants' perceptions in the treatment group compared to the control. The quality attribute also showed a closely aligned median score, with the treatment group nudging slightly higher to a median of 8. Conversely, Cookie B's ratings illustrated a divergent trend, with the treatment group recording a lower median taste rating of 5 compared to 6 in the control group, and a median texture rating that slightly improved to 6.5 in the treatment group from 6 in the control. This phase demonstrated tighter ranges of ratings across most attributes, especially in the treatment group, reflecting a narrower spread of opinions.



In Phase 3 of our experiment, Cookie A showed a marked improvement in the treatment group for all attributes, with taste and texture achieving a median of 8 and 8 respectively, and quality peaking at 8.5, compared to medians of 7 in the control group. The interquartile range was notably tighter in the treatment group, especially for quality, suggesting a more uniform assessment among participants. For Cookie B, the treatment group's ratings for taste increased slightly to 5.5 from 7 in the control group, while texture ratings decreased to 6. The quality attribute also saw a decrease in the median rating in the treatment group to 5, from 6 in the control. Overall, Phase 3 results reflected a more polarized view of the cookies in the treatment group, with higher ratings for Cookie A and lower for Cookie B compared to the control group.

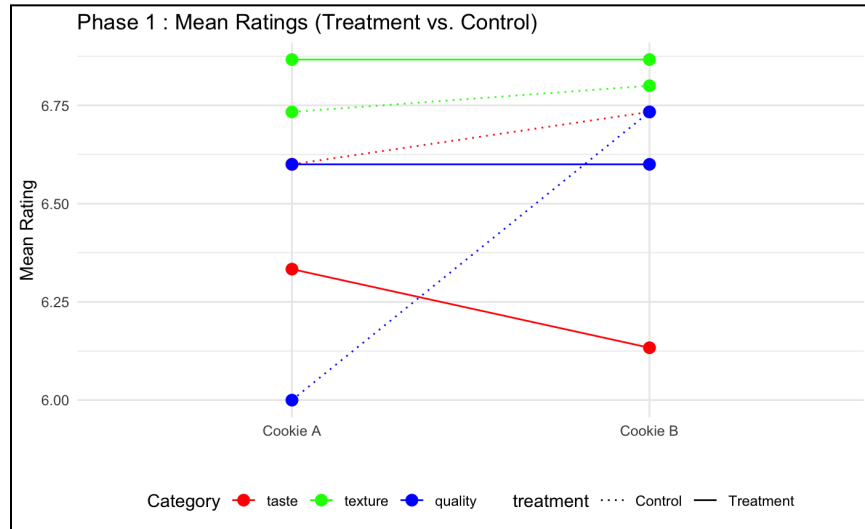


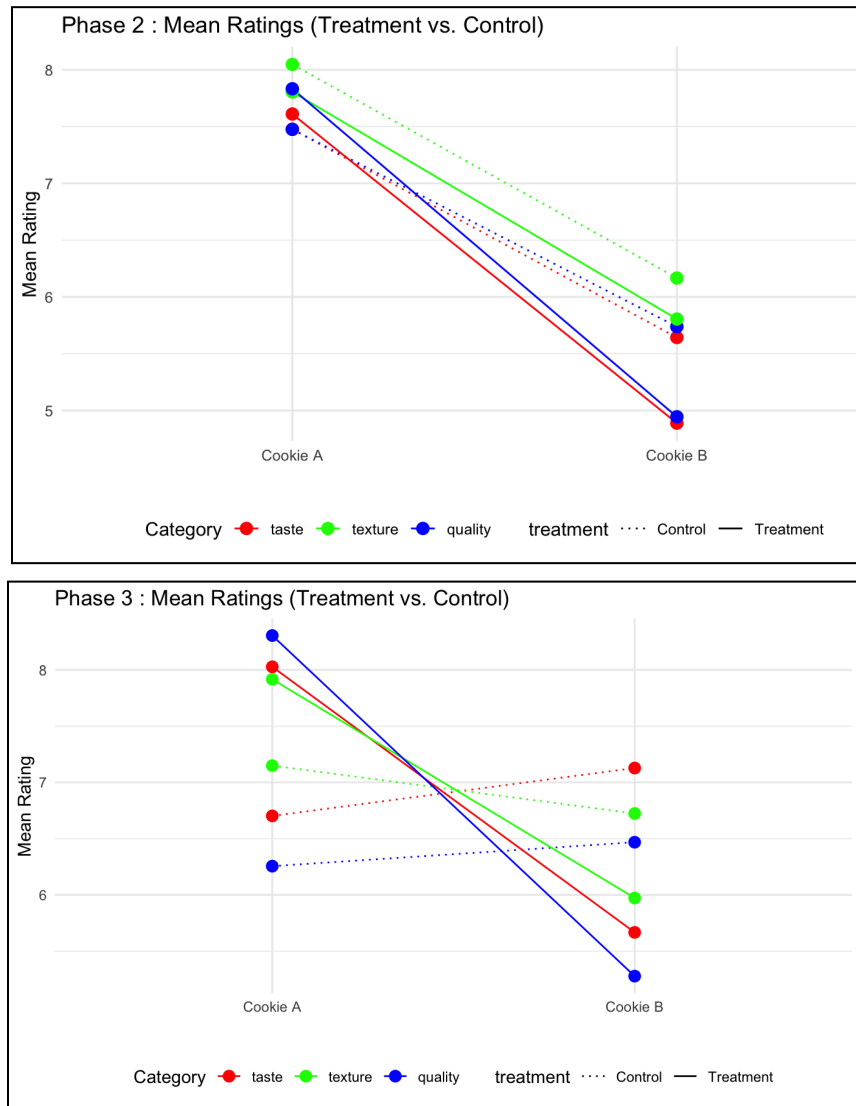
In pursuit of a more robust understanding of our visual findings, we employed statistical testing as outlined in our experimental methodology to substantiate the assumptions drawn from our box plot observations. As depicted in Table 2, p-values were calculated to test the null hypotheses associated with each experimental phase. In Phases 1 and 2, the high p-values across taste, texture, and quality suggest that there is no significant difference between the treatment and control groups, indicating that the order of cookie presentation and verbal information alone may not have a notable effect on participant ratings. However, in Phase 3, the low p-values for texture (0.0429), quality (0.0004), and taste (0.0008) suggest that the combined effect of verbal information and visual packaging cues significantly influences consumer ratings, challenging the null hypotheses and underscoring the importance of packaging in consumer perception.

Table 2: P-values for difference between treatment and control

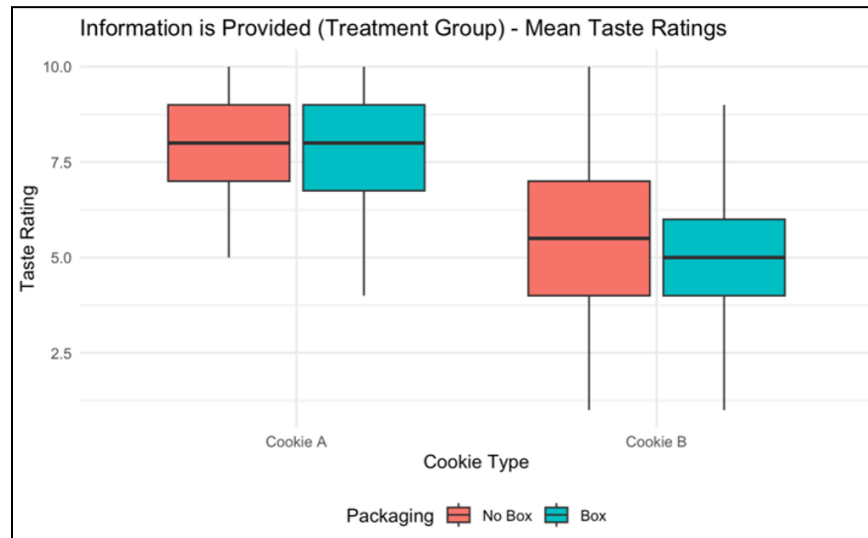
Phase	P_value	Question	Null
1	0.9465	texture	cookie order doesn't matter
1	0.2859	quality	cookie order doesn't matter
1	0.6901	taste	cookie order doesn't matter
2	0.8643	texture	info doesn't matter
2	0.0847	quality	info doesn't matter
2	0.1840	taste	info doesn't matter
3	0.0429	texture	info with box doesn't matter
3	0.0004	quality	info with box doesn't matter
3	0.0008	taste	info with box doesn't matter

The difference-in-differences plots for Phases 1, 2, and 3 serve as visual corroboration for the statistical findings indicated by the p-values. In Phase 1, the mean ratings show negligible divergence between treatment and control, paralleling the non significant p-values and suggesting that the order of cookie tasting does not substantially affect ratings. As we transition to Phase 2, the plots demonstrate a downward trend in mean ratings from Cookie A to Cookie B, more pronounced in the treatment group. The trend becomes most evident in Phase 3, where the mean ratings for all categories in the treatment group display notable deviation from the control. The lines showing the cookie ratings in the third figure below are very steep for the treatment group but flat for the control group. This visual trend aligns with the significantly lower p-values, solidifying the assertion that the combination of verbal information and packaging markedly influences consumer perceptions. These plots thus visually represent and support the statistically significant conclusions drawn from the experimental data, reinforcing the causal relationship suggested by the numerical evidence.

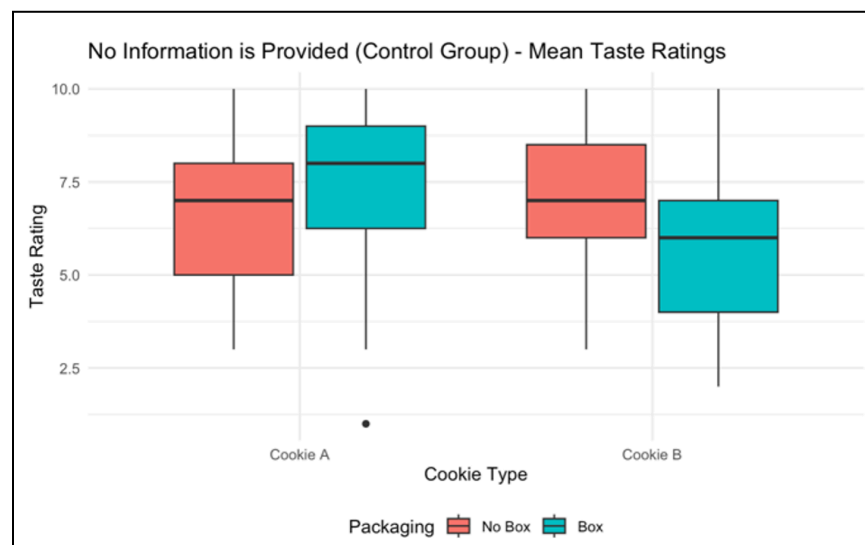




To further our analysis, we applied a difference-in-differences approach paired with t-tests, now comparing only the treatment group and the control group with each other across phases 2 and 3. We began by examining the mean ratings of taste to identify any patterns and differences associated with verbal information and the use of original packaging. The first box plot graph below shows the mean differences in taste ratings in the treatment group, where in both groups they received verbal information. We first observe that for both phases, Cookie A is rated higher and similarly compared to Cookie B. For Cookie A, packaging does not seem to significantly impact the median taste rating, but it appears to reduce variability slightly, as seen from the marginally tighter IQR. For Cookie B, the presence of packaging also does not have a major impact on the median rating but does seem to reduce variability, similar to Cookie A.



Now moving to our initial control group analysis, where no verbal information was provided, the box plots below show the taste ratings between phase 2 and phase 3. The medians for Cookie A and Cookie B in phase 2 are closely aligned. However, a notable observation is the substantial differences between mean taste ratings in phase 3. We see that there is a more pronounced difference in median ratings between the two Cookie A and Cookie B in phase 3 (blue boxplots) than phase 2 (red ones). Based on these observations, this possibly suggests a causal effect of packaging on preferences. To test this, we utilized paired t-tests to assess the difference in mean ratings between the two phases under information/no information and packaging/no packaging.



In order to evaluate the potential impact of packaging on taste preferences as observed in the box plot analysis, we constructed a specialized function in R to perform paired t-tests. Our analysis aims to determine whether the observed differences in taste ratings were statistically significant.

R Code: Paired t-tests

```
# Function to conduct t-tests on the difference in mean difference between two phases
conduct_t_tests_mean_diff_between_phases <- function(phase1_data, phase2_data) {
  phase1_mean_diff <- mean(phase1_data$taste_a - phase1_data$taste_b)
  phase2_mean_diff <- mean(phase2_data$taste_a - phase2_data$taste_b)

  t_test_result <- t.test(phase1_data$taste_a - phase1_data$taste_b,
                          phase2_data$taste_a - phase2_data$taste_b,
                          mu = phase2_mean_diff - phase1_mean_diff)

  return(t_test_result$p.value)
}
```

In the control groups for both phase 2 and phase 3, our paired t-tests revealed a statistically significant difference in mean taste ratings, with a p-value of less than 0.0001. This finding suggests that packaging alone may have a substantial influence on taste perception. Conversely, for the treatment group—which received both verbal information and visual cues—the analysis yielded a p-value of 0.2583. This result indicates that the presence of original packaging did not significantly affect taste ratings when combined with verbal information. These p-values are detailed in Table 3, which follows.

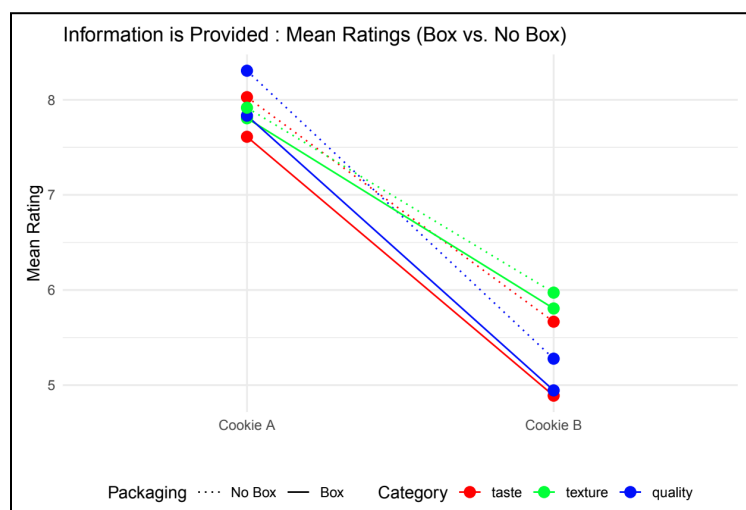
Table 3: P-values for difference between Phase 2 and 3 Results

Null	P_value
Box with no info doesn't matter	0.0000
Box with info doesn't matter	0.2583

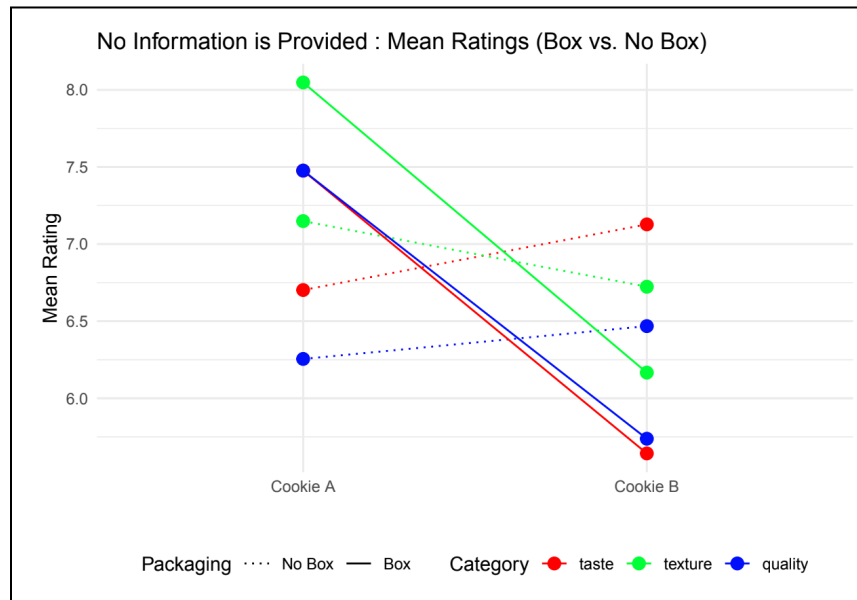
Our statistical analysis indicates a strong effect of visual cues on the control group's cookie preferences. The t-test for the control group between the mean taste rating differences for Phase 2 (no box, no information) and Phase 3 (original box, no information) yielded a highly significant p-value less than 0.0001. This shows that the introduction of the original packaging in Phase 3 significantly influenced the taste ratings of the cookies. When verbal information accompanied the original packaging, the change in mean taste rating differences from Phase 2 to Phase 3 was not statistically significant, with a p-value of 0.2583. This suggests that the additional verbal information provided to participants, along with the cookies' original packaging, did not alter their preferences significantly from when they received verbal cues

alone. Therefore, the combination of both verbal and visual cues in the treatment group did not have a distinguishing effect on cookie preference as compared to verbal information alone. This outcome highlights the influence of packaging on consumer perceptions, particularly when no other information is provided.

To achieve this, we implemented a difference-in-differences analysis. We utilize difference in differences plots to provide a visual and analytical comparison of the trends between the treatment and control groups. By plotting the mean ratings for taste, texture, and quality for each cookie under different conditions, we were able to observe and infer potential causal relationships. Looking at the difference in differences plot below, which is strictly comparing the treatment group and their ratings for Cookie A vs Cookie B where they received verbal information about the ingredients and descriptions of the cookie, but in phase 2 did not see the original box and phase 3 the participant did see the original box. Notice that for both phases, whether or not they saw the box or not, the mean ratings generally follow the same downward slope. We do not observe any large differences in how the participants rated them.



The difference-in-differences plot presented below focuses on the control group's ratings for Cookie A versus Cookie B, where neither verbal information nor original packaging was provided. In Phase 3, denoted by solid lines, participants experienced both cookies in their original packaging. The steeper slope in this phase underscores the substantial impact of packaging on taste evaluations. This suggests that the visual presentation associated with the high-quality packaging of Cookie A significantly enhanced its perceived taste quality compared to Cookie B. Conversely, in Phase 2, represented by the dotted lines, where both cookies were presented without any original packaging or additional verbal information, the differences in ratings are noticeably less steep. The smaller slope shows that without the influence of the original boxes, participants' taste perceptions of the two cookies were more closely aligned.



The difference in slopes between the treatment and control conditions within the difference framework provides evidence of a possible causal effect of packaging on taste perception. The steeper slope in the phase 3 control group implies that the visual cues from the packaging had a more significant influence when not accompanied by verbal information, thereby suggesting a primary role of visual cues in influencing consumer preferences. The steeper slope observed in the phase 3 control group aligns with the highly significant p-value from the t-tests, reinforcing the notion that packaging profoundly influences taste ratings. The integration of these findings within our statistical framework highlights the relation between visual and verbal information in consumer perception.

Discussion

Our comprehensive experimental study, engaging with various elements of consumer psychology, primarily investigated how different forms of product information influence consumer preferences for 2 chocolate oreo knock-off cookies. The experimental design spanned five distinct phases, each tailored to dissect the interplay between verbal information, visual branding, and consumer perception.

The results from the first phase showed no significant order effects, allowing us to confidently assert that any subsequent findings regarding preference were not confounded by the anchoring effect of the first cookie tasted. Several concerns raised while doing the experiment in class were the effects of setting a baseline with the first cookie given, but by running phase 1, we eliminated the effect of time and order.

Phases two and three were pivotal, examining the impact of verbal information about the quality of ingredients and the additional influence of visual branding (cookie packaging). Our findings here were particularly revealing; while verbal information alone (phase two) did not significantly sway consumer preferences, the combination of verbal information and visual cues (packaging) in phase three markedly influenced taste perceptions, underscoring the potent synergy between what is told and what is seen.

The fourth and fifth phases of the experiment involved cross-comparative analyses designed to isolate the effects of packaging from verbal information by comparing both treatment and control groups across phases two and three. The analysis from phase four, comparing treatment groups, did not yield significant differences, suggesting that the verbal information's impact remained consistent across settings. However, phase five revealed significant differences between control groups, indicating that packaging alone substantially influences consumer preferences even without accompanying verbal information.

The Power of Visual Cues

The results from our study vividly illustrate that visual branding, coupled with verbal information, plays a critical role in shaping consumer preferences. This finding aligns with existing research in marketing psychology, suggesting that consumers are not merely rational actors but are significantly influenced by sensory and emotional appeals embedded within branding elements. The significant effect of packaging on taste preference, even in the absence of verbal cues (as demonstrated in phase five), reiterates the importance of visual cues in consumer decision-making processes.

Implications for Consumer Awareness

These insights are especially pertinent in the context of modern grocery shopping, where consumers are bombarded with a myriad of product options and marketing messages. Our study suggests that consumers, often unknowingly, may be swayed more by the appearance and presentation of products than by the intrinsic quality of the products themselves. The participants in our study did not have a significant change in preferences when only given nutritional verbal information about the quality of the cookies without revealing the brand. They had a significant impact when they saw the box of the brand and the packaging themselves. This underscores the need for consumers to be more discerning and critical of the persuasive tactics employed through packaging and branding. It also suggests a potential avenue for consumer education, emphasizing the importance of looking beyond the surface and making informed choices based on the product's substantive qualities rather than its aesthetic appeal.

Limitations and Further Research

While our findings are robust within the context of our experimental design, they are not without limitations. The sample predominantly consisted of college-aged statistics students, which may not be fully representative of the general consumer population. Future studies could look to replicate these findings across a more diverse demographic to generalize the conclusions more broadly to other food items and compare nutrients rich vs junk food. With time and resource constraints, we could only play around with packaging/branding and information about the product, but this could be extended to studying the human psyche and seeing how people react when information is available - for example, a possible experiment is to switch the cookie boxes or the information given and see if that sways people's preferences. Additionally, exploring the long-term effects of repeated exposure to branding and information on consumer preferences could yield deeper insights into consumer behavior dynamics.

In conclusion, our study provides compelling evidence that both verbal and visual information impacts consumer preferences, with visual cues playing a particularly significant role. As we navigate the complex landscape of consumer goods, this knowledge encourages both consumers and marketers to consider the profound effects of packaging and branding on decision-making processes. In an era where consumer awareness is more critical than ever, such insights are invaluable in guiding more informed and deliberate consumption choices.

References

- [1] Schouteten, J. J., De Steur, H., De Pelsmaeker, S., Lagast, S., De Bourdeaudhuij, I., & Gellynck, X. (2015). Impact of Health Labels on Flavor Perception and Emotional Profiling: A Consumer Study on Cheese. *Nutrients*, 7(12), 10251–10268.
<https://doi.org/10.3390/nu7125533>
- [2] Sütterlin, B., & Siegrist, M. (2015). Simply adding the word “fruit” makes sugar healthier: The misleading effect of symbolic information on the perceived healthiness of food. **Appetite**, 95, 252-261.
<https://doi.org/10.1016/j.appet.2015.07.011>