

Soda Taste Test Factorial Design with Blocking

Kian Jadbabaei

2025-03-16

Introduction

Drinks are essential for survival, providing the hydration and nutrients needed for daily life. Not only just essential, they are woven into culture, social gatherings, and personal enjoyment. A morning coffee is the start of the day for many, while an ice-cold soda on a hot afternoon offers refreshment for others. Despite their constant presence, little thought is given to how small details—like the way a drink is served—affect the experience. Taste is not just about flavor; it is about perception. The way a beverage is poured, the container it is served in, and the method of drinking can subtly alter how it tastes and how much it is enjoyed. Something as simple as the presence of bubbles, the temperature of the drink, or even the texture of the container can create a noticeably different experience for the consumer. Despite the widespread consumption of carbonated beverages, few consider how serving methods might impact their experience. Could pouring soda down the side of a glass instead of straight in the middle make it more enjoyable? Does drinking from a straw versus directly from the can change the way it feels and tastes? Research suggests that carbonation affects sweetness perception by reducing the distinction between sugar and artificial sweeteners, meaning the way a soda is poured and consumed may influence how it is perceived (Di Salle et al., 2013). For instance, a soda that has lost some of its carbonation might taste flat and overly sweet, whereas a freshly poured, highly carbonated soda can create a sharper, more refreshing sensation. So does a different pouring method alter the carbonation which alters the taste? Additionally, external factors like lighting, background music, and even the weight of a glass can influence taste perception, showing that serving methods may have a greater impact than commonly assumed (Spence, Harrar, & Piqueras-Fiszman, 2012). The method of consumption of a drink is also important because plastic straws, a very common vessel, significantly contribute to marine and coastal waste, harming the world around us (Neto et al., 2021). As concerns about environmental impact grow, many companies have shifted towards alternative materials such as paper or metal straws, yet these changes also affect the drinking experience. A flimsy paper straw can alter the texture and even add an unintended aftertaste to the beverage, whereas a metal straw can make the drink feel colder. This experiment aims to investigate how pouring method and consumption method affect the taste of soda. By testing three pouring styles: down the side of the glass, straight into the center, and leaving it in the can, and two consumption methods: drinking with a straw or directly from the container, this experiment investigates how each factor influences enjoyment individually and whether the impact of pouring method depends on how the soda is consumed. Understanding how these factors shape taste perception is important beyond just personal preference—it affects external factors like marketing strategies and product design.

Methods

The experiment was carried out by having participants taste soda under 6 different conditions and rate the taste on a 0-10 scale where 0 was the worst and 10 was the best. The two factors were pouring method and consumption method. Pouring method had 3 levels: down the side of the glass, straight in the center, and left in can and the consumption method had 2 levels: straw or no straw. Each participant was a block since they served as their own control. I had a total of 9 different participants that participated in the experiment, but I could not get everyone at one time, so each person tried all 6 conditions in a row individually.

I randomized the order of conditions for each participant using a simple sample function with R code. The

participant then drank each condition and gave a rating on the taste from 0-10 and I recorded every rating. Each participant was given water to reset the taste in their mouth between each sample. The statistical approach for this study was designed to assess how pouring method and consumption method impact taste ratings. To analyze these effects, a linear model was initially used to estimate the influence of both factors and their interaction on taste perception. An ANOVA model was also fit to compare variance across groups. These models assume that residuals follow a normal distribution, variances are equal across groups, and residuals are independent. To verify these assumptions, normality was tested using a Shapiro-Wilk test and visualized with QQ plots, homoscedasticity was assessed through residual plots, and independence was checked by plotting residuals against the order of data collection. If these assumptions hold, lm and ANOVA provide valid estimates of significance, but if they are violated, alternative methods may be necessary. To account for the possibility that parametric methods might not be suitable, a permutation test was also planned. This test works by randomly shuffling taste ratings within each participant to generate an empirical null distribution, allowing comparison of observed test statistics against a non-parametric baseline. The permutation test was performed both for the overall model and separately for pouring method, consumption method, and their interaction. Since this approach does not assume normality or equal variance, it provides a more flexible alternative if assumption violations are detected. A power analysis was also incorporated to determine whether the sample size was sufficient to detect meaningful effects. A few technical challenges arose during data collection. Some participants noted that carbonation levels seemed to change over time as some sodas could have been more flat if they were opened sooner, potentially influencing taste ratings. Additionally, ensuring a consistent order of taste testing was difficult, as palate fatigue may have affected later responses. These factors were controlled for as much as possible, but they represent potential errors.

Results

```
library(readxl)
soda_data <- read_excel("soda_data.xlsx")
colnames(soda_data) <- c("Participant", "Pouring_Method", "Consumption_Method", "Taste_Rating")

library(knitr)
library(ggplot2)
library(dplyr)

source("power_factorial_32.R")

lm_model <- lm(Taste_Rating ~ Pouring_Method * Consumption_Method, data = soda_data)
coef_summary <- summary(lm_model)$coefficients

sample_sizes <- seq(5, 50, by=5)
beta_se_values <- c(0.75, 1, 1.5)

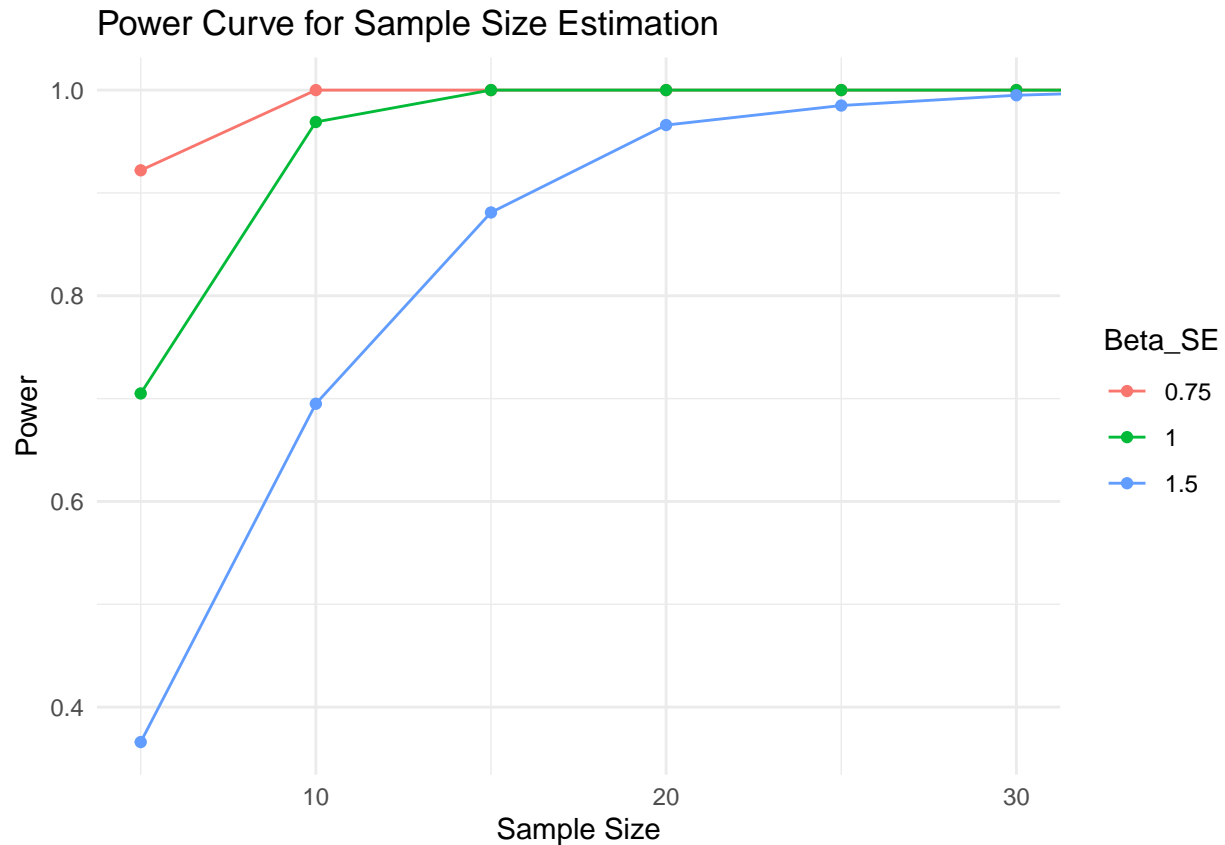
results <- data.frame()

for (beta_se in beta_se_values) {
  for (n in sample_sizes) {
    beta_mean <- c(coef_summary[1,1], coef_summary[2,1], coef_summary[3,1], coef_summary[4,1], coef_summary[5,1])
    power <- power_factorial_32(beta_mean, rep(beta_se, length(beta_mean)), n)
    results <- rbind(results, data.frame(Sample_Size=n, Power=power, beta_se=beta_se))
  }
}

results$beta_se <- factor(results$beta_se, levels = c(0.75, 1, 1.5))

ggplot(results, aes(x=Sample_Size, y=Power, color=beta_se, group=beta_se)) +
```

```
geom_line() +
geom_point() +
labs(
  x = "Sample Size",
  y = "Power",
  title = "Power Curve for Sample Size Estimation",
  color = "Beta_SE"
) +
coord_cartesian(xlim = c(5, 30)) +
theme_minimal()
```



The power curve for sample size estimation demonstrates how statistical power changes with sample size for different levels of variability in effect estimates (Beta_SE). With 9 participants, the study achieves over 80% power for Beta_SE values of 0.75 and 1, indicating that the sample size is sufficient to detect effects under these conditions. However, for Beta_SE = 1.5, power remains lower at smaller sample sizes, requiring a larger sample to reach the 80% threshold. This suggests that while the current sample size is adequate for detecting effects with moderate variability, greater uncertainty in effect estimates would necessitate a larger sample to maintain statistical confidence.

```
summary_table <- soda_data %>%
  group_by(Pouring_Method, Consumption_Method) %>%
  summarise(
    Mean_Taste = round(mean(Taste_Rating), 2),
    SD_Taste = round(sd(Taste_Rating), 2),
    Sample_Size = n()
  )
```

```
kable(summary_table, caption = "Summary Statistics of Taste Ratings",
      col.names = c("Pouring Method", "Consumption Method", "Mean Taste Rating", "SD", "Sample Size"))
```

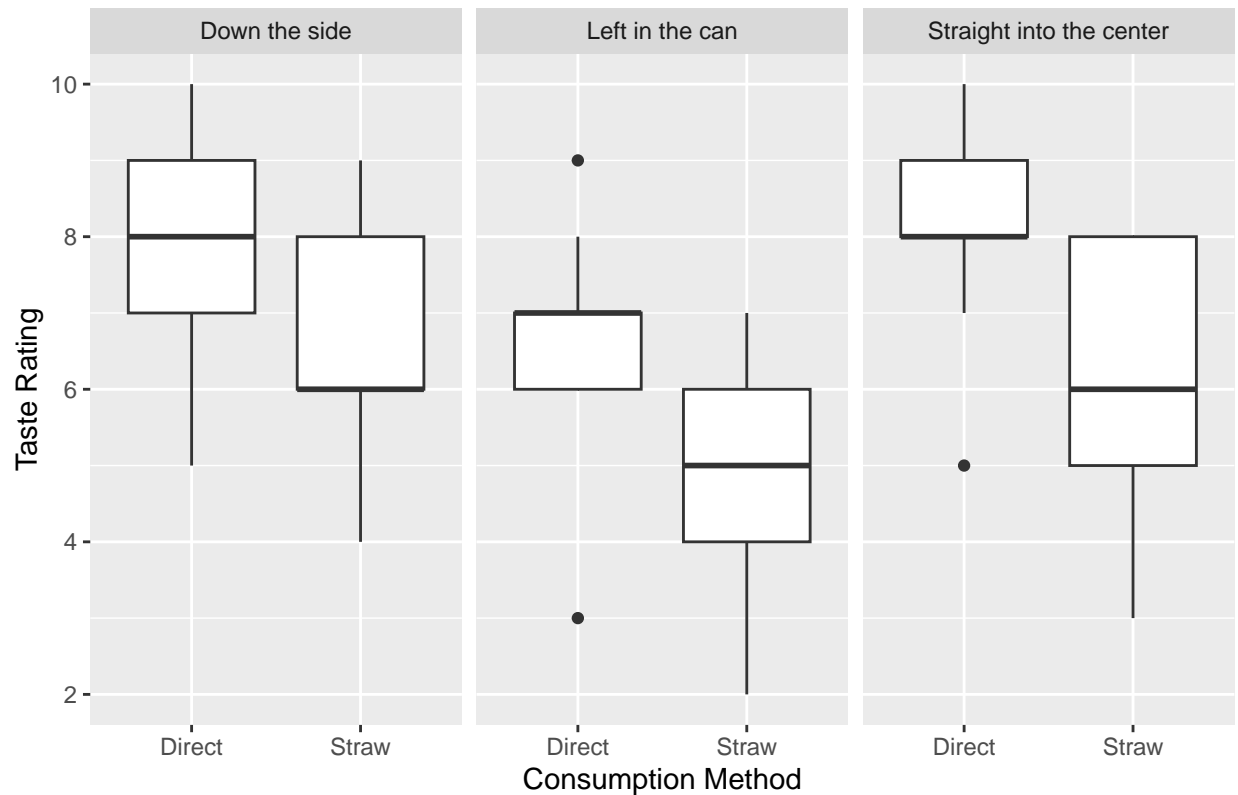
Table 1: Summary Statistics of Taste Ratings

Pouring Method	Consumption Method	Mean Taste Rating	SD	Sample Size
Down the side	Direct	7.78	1.56	9
Down the side	Straw	6.56	1.74	9
Left in the can	Direct	6.67	1.66	9
Left in the can	Straw	5.00	1.58	9
Straight into the center	Direct	8.00	1.41	9
Straight into the center	Straw	6.11	1.69	9

The summary statistics indicate that direct consumption consistently results in higher mean taste ratings across all pouring methods compared to using a straw. Among pouring methods, soda poured straight into the center and consumed directly had the highest average rating (8.00), while soda left in the can and consumed with a straw had the lowest (5.00). The standard deviations suggest moderate variability in ratings, with the highest variability observed for sodas consumed through a straw, indicating greater inconsistency in taste perception when using a straw.

```
ggplot(data = soda_data, aes(x = Consumption_Method, y = Taste_Rating)) +
  geom_boxplot() +
  facet_grid(cols = vars(Pouring_Method)) +
  labs(
    title = "Taste Ratings Based on Pouring and Consumption Methods",
    x = "Consumption Method",
    y = "Taste Rating"
  )
```

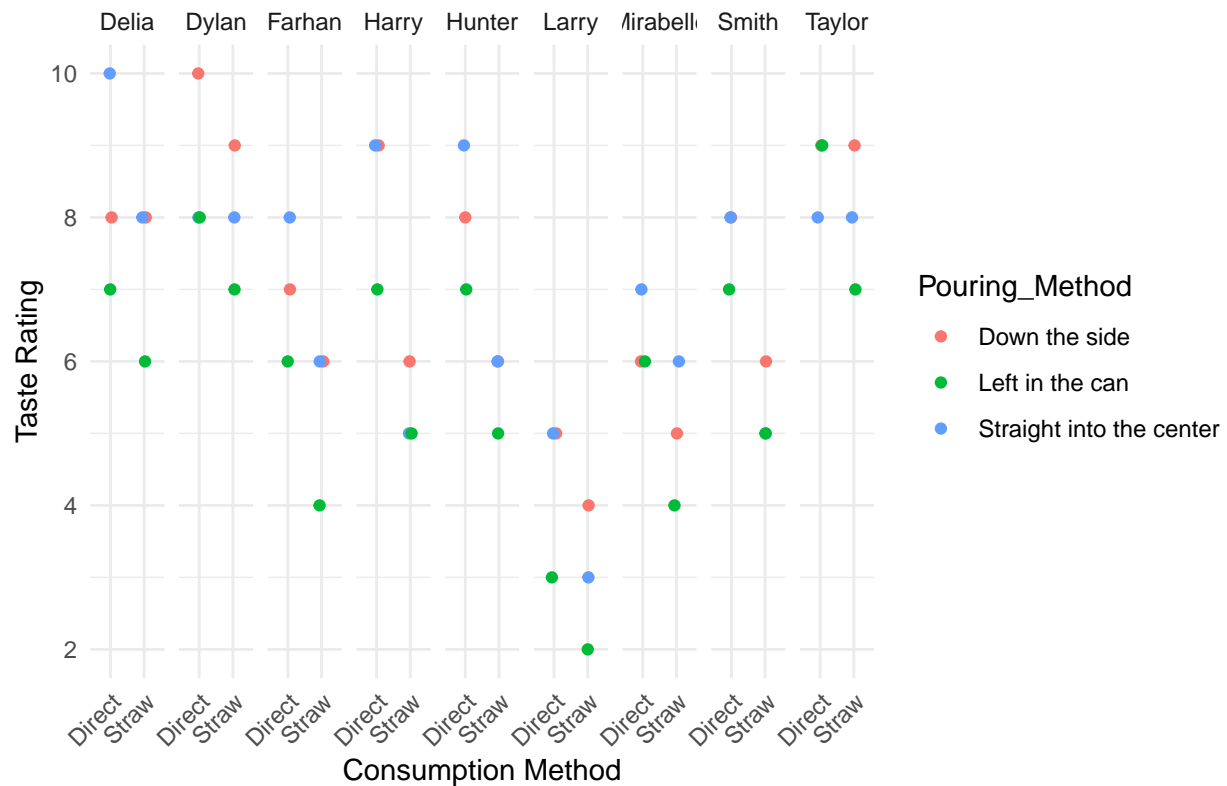
Taste Ratings Based on Pouring and Consumption Methods



The boxplot visualization reinforces the trends observed in the summary statistics, showing that direct consumption consistently results in higher taste ratings across all pouring methods.

```
ggplot(data = soda_data, aes(x = Consumption_Method, y = Taste_Rating, color = Pouring_Method)) +
  geom_jitter(width=0.08, height=0) +
  facet_grid(cols = vars(Participant)) +
  labs(title = "Taste Ratings Based on Pouring and Consumption Methods",
        x = "Consumption Method",
        y = "Taste Rating") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

Taste Ratings Based on Pouring and Consumption Methods

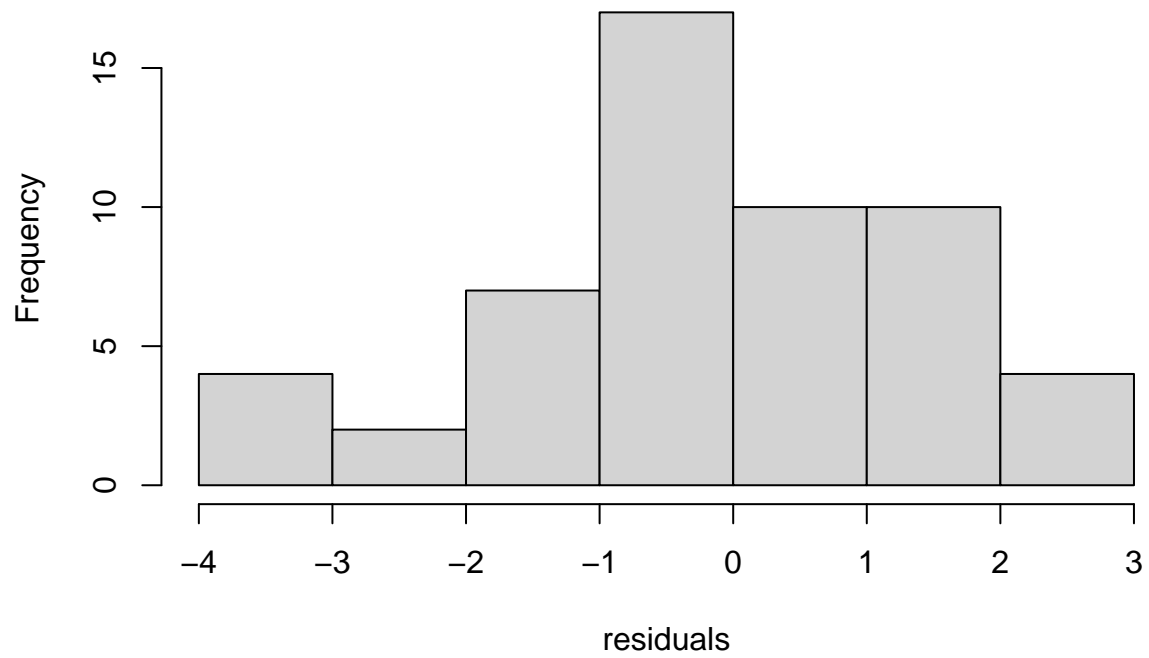


This scatterplot provides a participant-level breakdown of taste ratings, allowing for a closer look at individual variability. The pattern aligns with previous findings—direct consumption tends to yield higher ratings, while ratings for soda consumed with a straw show more spread. Differences in pouring methods are also visible, with soda poured straight into the center showing more high ratings across participants. This visualization helps confirm that the observed effects are not driven by just a few individuals but are present across multiple participants.

Model Checking

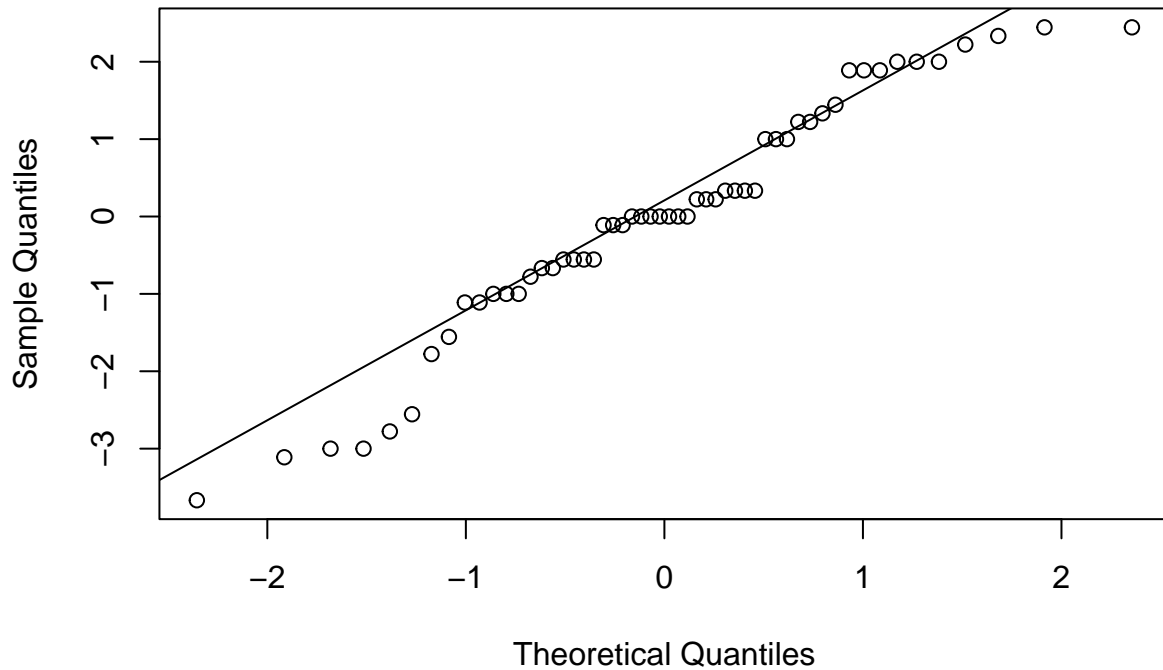
```
model_norm <- lm(Taste_Rating ~ Pouring_Method + Consumption_Method + Pouring_Method*Consumption_Method)
hist(model_norm$residuals, xlab="residuals", main="Soda Data Residuals")
```

Soda Data Residuals



```
qqnorm(model_norm$residuals)
qqline(model_norm$residuals)
```

Normal Q-Q Plot



```
shapiro_result <- shapiro.test(model_norm$residuals)

shapiro_df <- data.frame(
  "Statistic" = round(shapiro_result$statistic, 3),
  "P-Value" = round(shapiro_result$p.value, 3)
)

kable(shapiro_df, col.names = c("Shapiro-Wilk Statistic", "P-Value"), caption = "Shapiro-Wilk Test for Normality of Residuals")
```

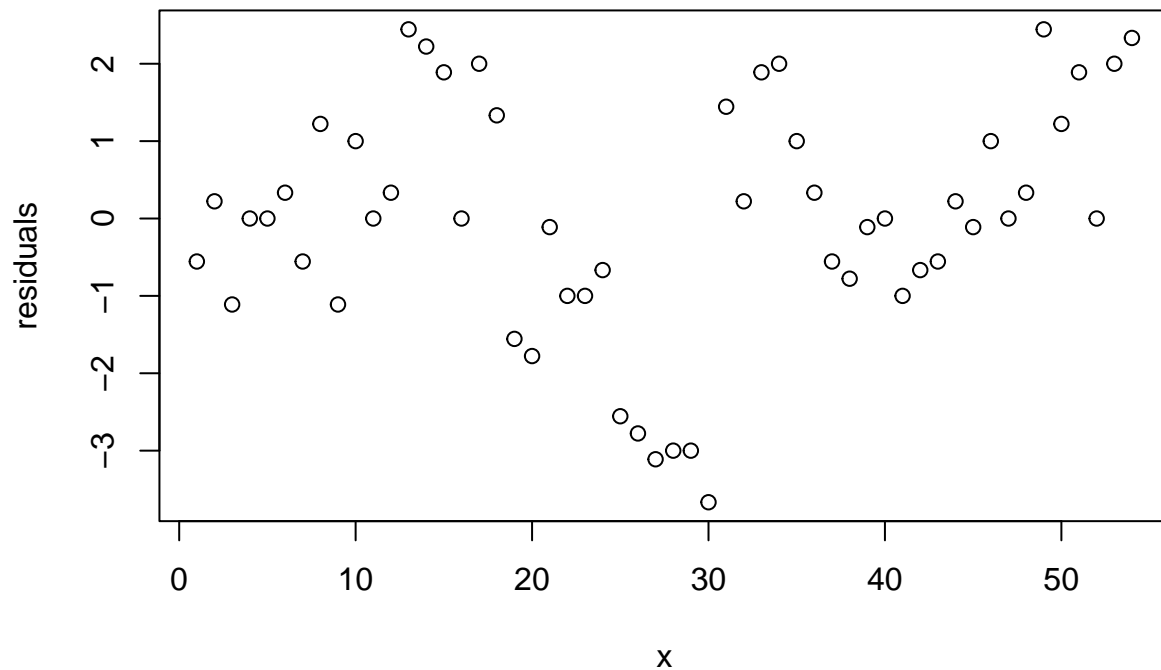
Table 2: Shapiro-Wilk Test for Normality of Residuals

	Shapiro-Wilk Statistic	P-Value
W	0.954	0.038

At $\alpha = 0.05$ we conclude that there is statistically significant evidence that the residuals do not follow a normal distribution seen through the skewness in the histogram, deviations at the tails of the Q-Q plot as well as the p-value of the Shapiro-Wilk test being 0.03767 which is less than 0.05, so we reject the null hypothesis that the residuals are normally distributed, proving that we have at least one assumption violated.

```
x <- 1:length(model_norm$residuals)
plot(model_norm$residuals ~ x, ylab="residuals",
     main="Residuals vs order of data collection")
```


Residuals vs order of data collection



There appears to be a slight pattern in the residuals so there is some concern. There is also concern about inequality of variances from the graph of residuals vs fitted values.

Permutation Test

```
library(broom)

# permutation test for overall model
perm_f <- numeric(10000)
reps <- 10000
blocks <- unique(soda_data$Participant)

for(i in 1:reps){
  soda_perm <- soda_data

  for(curr_block in blocks){
    ind <- which(soda_data$Participant == curr_block)

    soda_perm$Taste_Rating[ind] <- sample(soda_perm$Taste_Rating[ind])
  }

  perm_f[i] <- summary(aov(Taste_Rating ~ Pouring_Method * Consumption_Method, data=soda_perm))[[1]][1,4]
}

f_actual <- summary(aov(Taste_Rating ~ Pouring_Method * Consumption_Method, data=soda_data))[[1]][1,4]

perm_p <- sum(perm_f >= f_actual) / reps
```

```
lm_model <- lm(Taste_Rating ~ Pouring_Method * Consumption_Method, data=soda_data)
p_lm <- glance(lm_model)$p.value

p_values_overall <- data.frame(
  "Test" = c("Overall Permutation Test", "Traditional lm model"),
  "P-Value" = c(perm_p, p_lm)
)

kable(p_values_overall, col.names = c("Test", "P-Value"), caption = "Comparison of P-Values: Permutation Test vs lm Model")
```

Table 3: Comparison of P-Values: Permutation Test vs lm Model

	Test	P-Value
	Overall Permutation Test	0.0002000
value	Traditional lm model	0.0028696

The permutation test ($p = 0.0005$) confirms that Pouring Method and/or Consumption Method significantly affect taste ratings. Since permutation tests don't rely on normality or equal variances, they provide a stronger and more reliable measure of significance. `lm()` also showed significance ($p = 0.00287$), but it assumes conditions that don't fully hold based on earlier model checks. With non-normal residuals and unequal variances, the permutation test is the better indicator of real effects in this study.

```
# permutation test for Pouring Method, Consumption Method, and Interaction
reps <- 10000
blocks <- unique(soda_data$Participant)

perm_f_pouring <- numeric(reps)
perm_f_consumption <- numeric(reps)
perm_f_interaction <- numeric(reps)

for (i in 1:reps) {
  soda_perm <- soda_data

  for (curr_block in blocks) {
    ind <- which(soda_data$Participant == curr_block)
    soda_perm$Taste_Rating[ind] <- sample(soda_perm$Taste_Rating[ind])
  }

  aov_perm <- aov(Taste_Rating ~ Pouring_Method * Consumption_Method, data = soda_perm)
  aov_summary <- summary(aov_perm)

  perm_f_pouring[i] <- aov_summary[[1]][1, 4] # Pouring Method
  perm_f_consumption[i] <- aov_summary[[1]][2, 4] # Consumption Method
  perm_f_interaction[i] <- aov_summary[[1]][3, 4] # Interaction
}

aov_actual <- aov(Taste_Rating ~ Pouring_Method * Consumption_Method, data = soda_data)
aov_actual_summary <- summary(aov_actual)

f_pouring_actual <- aov_actual_summary[[1]][1, 4]
f_consumption_actual <- aov_actual_summary[[1]][2, 4]
f_interaction_actual <- aov_actual_summary[[1]][3, 4]
```

```

perm_p_pouring <- mean(perm_f_pouring >= f_pouring_actual)
perm_p_consumption <- mean(perm_f_consumption >= f_consumption_actual)
perm_p_interaction <- mean(perm_f_interaction >= f_interaction_actual)

p_values <- data.frame(
  "Factor" = c("Pouring Method", "Consumption Method", "Interaction"),
  "P-Value" = c(perm_p_pouring, perm_p_consumption, perm_p_interaction)
)

kable(p_values, col.names = c("Factor", "P-Value"), caption = "Permutation Test P-Values")

```

Table 4: Permutation Test P-Values

Factor	P-Value
Pouring Method	0.0002
Consumption Method	0.0000
Interaction	0.6550

The permutation test results show that Pouring Method ($p = 0.0007$) and Consumption Method ($p = 0.000$) have statistically significant effects on Taste Rating at the $\alpha = 0.05$ level since they are both less than 0.05, meaning that both factors independently influence soda enjoyment. However, the Interaction effect ($p = 0.6608$) is not significant, indicating that the effect of Pouring Method does not depend on Consumption Method—each factor impacts taste perception independently rather than in combination. Since these results are derived from a permutation test, they are more reliable than traditional ANOVA due to assumption violations established earlier. This suggests that how soda is poured and how it is consumed both significantly impact enjoyment, but their effects do not interact, meaning neither factor amplifies or diminishes the impact of the other.

Discussion

The results of this study provide strong evidence that both pouring method and consumption method significantly impact taste, while their interaction does not. The permutation tests revealed that both main effects had statistically significant p-values, confirming their influence on taste ratings. The interaction term, however, was not statistically significant, indicating that the effect of one factor did not depend on the other. This suggests that differences in taste scores were primarily driven by individual factors rather than a combined effect. The sample size analysis confirmed that the chosen number of participants was sufficient under certain conditions. At Beta_SE values of 0.75 and 1, the power exceeded 80%, meaning that the study had a strong likelihood of detecting real effects under these levels of variability. However, for higher levels of variability, such as Beta_SE = 1.5, power was noticeably lower, suggesting that a larger sample size would be required to maintain better conclusions. This is a limitation of the study, while the current sample size was sufficient for moderate variability, results may not generalize well under higher variance conditions. Additionally, model diagnostics indicated that an ANOVA approach or linear model would not be appropriate due to the non-normality in residuals. This justified the use of permutation tests, which do not rely on these assumptions, ensuring that p-values were valid. The use of blocking by participant helped control for individual differences in taste perception, but there were still uncontrolled external factors that could have influenced results. Research has shown that carbonation influences the perception of sweetness and overall taste experience, which aligns with the finding that different pouring and drinking methods affect enjoyment (Di Salle et al., 2013).

Beyond statistical considerations, the findings have useful implications for both consumer behavior and industry applications. The fact that pouring method affects taste means that serving techniques could be optimized for a better sensory experience. For example, pouring down the side of the glass may improve

taste ratings by reducing carbonation loss. The significant effect of consumption method suggests that drinking directly may lead to a more favorable taste experience than using a straw, possibly due to the way carbonation interacts with the palate. These insights could be useful for beverage companies in designing packaging and marketing strategies, as well as for consumers seeking to maximize enjoyment. Despite these contributions, there are still limitations. The study only taste right after taking a sip and did not account for taste in the longer term. Additionally, while the factorial design allowed for testing interactions, future research could explore additional factors such as temperature or type of glass to better understand sensory perception. Increasing the sample size and finding more diverse participants would further strengthen these results, every participant in this study was in the same age and educational range. Overall, this study demonstrates that simple changes in pouring and drinking methods can significantly alter taste perception. The statistical approaches used ensured valid conclusions despite assumption violations, and the sample size analysis confirmed that under moderate variability, the experiment was well-powered. These findings contribute to a deeper understanding of how sensory experiences are shaped by external factors, with potential applications in both consumer behavior and product optimization. Future research could expand on these results to explore further sensory influences and find better strategies for enhancing enjoyment of drinking beverages.

References

- Di Salle, Francesco, et al. "Effect of carbonation on brain processing of sweet stimuli in humans." *Gastroenterology* 145.3 (2013): 537-539.
- Neto, Arnaldo Mailes, et al. "An overview of plastic straw policies in the Americas." *Marine Pollution Bulletin* 172 (2021): 112813.
- Spence, Charles, Vanessa Harrar, and Betina Piqueras-Fiszman. "Assessing the impact of the tableware and other contextual variables on multisensory flavour perception." *Flavour* 1 (2012): 1-12.