How Sensory Modalities Affect Memory Recall

Experiment and Analysis

STAT 453 Design and Analysis of Experiments, Winter 2025

Submitted to: Dr. Michelle Miranda
Submitted by: Eric Huber, V00980396
Kush Manek, V00960352
Amritpal Singh, V00997628

Contents

Introduction

The ability to accurately recall information is fundamental to effective communication, learning, and decision-making across numerous contexts, including personal, academic and professional environments. As multimedia presentations become increasingly prevalent and influential in enhancing learning [1], refining and assessing which combination of stimuli optimizes effective learning and memory retention is vital. In the modern era, where podcasts [2] and YouTube [3] play a more significant role in research and exploration both within and outside the classroom, understanding which sensory modality- visual, auditory or a combined audiovisual- most effectively optimizes for memory retention could provide tremendous value in developing learning materials and improving everyday information retention. This research examines whether memory recall for a list of commonly found grocery items differs significantly based on whether the information is presented to the individual in these three manners, mainly visually, auditorily or through a combined audiovisual approach.

Experimental Design

In order to answer our research question, a Randomized Complete Block Design (RCBD) was selected where each participant is randomly assigned to all of the three treatment factors of:

1) **Visual** - Participants see a list of 10 grocery store items for exactly 20 seconds
2) **Auditory** - Participants hear a recording of the list of 10 grocery store items being read for exactly 20 seconds
3) **Audiovisual** - Participants simultaneously see and hear another list of 10 grocery store items for exactly 20 seconds

The independent variable here is the mode of presentation. The lists were of a fixed length of 10 items per list, and the presentation duration was exactly 20 seconds per list. Finally, three distinct lists were created to reduce the nuisance factor of list familiarity.

The response variable selected was the number of correctly recalled items from a list recorded immediately after the end of the exposure. This score was a minimum of 0 with a maximum of 10. It was selected based on Miller's Magic Number, which posits that the human short-term memory span typically ranges between 5 and 9 items [4].

Our main nuisance factor is gender, which we selected due to its influence on cognitive processing and memory retention as suggested by public perception [5] and supported by empirical research [6] [7]. To control for those differences, participants will be divided into two gender categories: male and female. Furthermore, equal sample sizes across gender groups ensure a blanched distribution across treatment groups; this avoids nuisance factors introduced by unequal sample sizes across gender groups.

Further participant-related nuisance factors could influence our results, such as individual memory ability variability, age variation and participant motivation or stress. To counter these factors, randomisation in the selection process was introduced to distribute the effects of these factors evenly across treatment groups. Some environmental factors were considered, such as ambient noise, time of day and location. These were largely more uncontrollable, and we attempted to limit their effects by standardising our instructions and procedures and splitting them into groups to record data at different times and locations.

The three basic principles of experimental design were implemented as follows:

1) **Randomisation:** After participants are blocked by gender, individuals within each gender group are randomly assigned to which list and receive which treatment to help mitigate any unforeseen biases related to personal cognitive abilities or other uncontrolled differences among individuals. Furthermore, the treatment order was randomised to address bias introduced by receiving treatments in the same order.

2) **Replication**: To address issues and ensure our results are robust and reliable, our design includes multiple participants within each treatment-gender combination, increasing our degrees of freedom. Replication will ensure a reliable estimation of variability and strengthen the conclusions drawn from our ANOVA tests.

3) **Blocking**: The reason we selected the Randomized Complete Block Design was to effectively control gender-related variability, which existing research suggests influences cognitive processing and memory retention. In our experiment, participants are grouped into blocks based on gender. This gender blocking controls for inherent differences in cognitive processing abilities between males and females, identified as a significant nuisance factor. By blocking in this manner, variability due to gender is isolated, enabling a more precise assessment of how the sensory mode of information presentation (auditory, visual, audiovisual) impacts memory retention.

The statistical model of our design is:

$$Y_{ij} = \mu + \tau_i + \beta_i + \epsilon_{ij}$$

where:

- Y_{ij}: Observed memory recall (response variable) for participant j receiving treatment i

- μ : Overall mean memory recall
- τ_i : Effect of the i^{th} treatment mode (visual, auditory, audiovisual)
- β_j : Effect of the j^{th} block (gender: male, female).
- ϵ_{ij} : Random error term (assumed independently and normally distributed)

Analysis and Diagnostics

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(ggplot2)
```

```r
library(dplyr)
library(ggplot2)
df <- read.csv("combined_data.csv")
head(df)
```

```
##   Person Score   Treatment Gender List
## 1      1     5 Audiovisual Female    1
## 2      1     5      Visual Female    2
## 3      1     4       Audio Female    3
## 4      2     4 Audiovisual   Male    2
## 5      2     4      Visual   Male    3
## 6      2     4       Audio   Male    1
```

```r
df$Treatment <- as.factor(df$Treatment)
df$Person <- as.factor(df$Person)
df$Score <- as.numeric(df$Score)
levels(df$Treatment)
```

```
## [1] "Audio"       "Audiovisual" "Visual"
```

```r
levels(df$Person)
```

```
##  [1] "1"  "2"  "3"  "4"  "5"  "6"  "7"  "8"  "9"  "10"
```

```r
df %>%
  group_by(Treatment) %>%
  summarise(
    count = n(),
    mean = mean(Score, na.rm = TRUE),
    sd = sd(Score, na.rm = TRUE)
  )
```

```
## # A tibble: 3 x 4
##   Treatment   count  mean    sd
##   <fct>       <int> <dbl> <dbl>
## 1 Audio          30  4.97  1.43
## 2 Audiovisual    30  5.63  1.54
## 3 Visual         30  5.87  1.66
```

```r
df %>%
  group_by(Person) %>%
  summarise(
    count = n(),
    mean = mean(Score, na.rm = TRUE),
    sd = sd(Score, na.rm = TRUE)
  )
```

```
## # A tibble: 10 x 4
##    Person count  mean    sd
##    <fct>  <int> <dbl> <dbl>
##  1 1          9  5.22 0.667
##  2 2          9  5.56 1.59
##  3 3          9  6    1.32
##  4 4          9  5.33 0.707
##  5 5          9  5.33 1.66
##  6 6          9  5.22 1.09
##  7 7          9  5.67 2.69
##  8 8          9  5.78 2.33
##  9 9          9  5.44 1.42
## 10 10         9  5.33 1.73
```

Hypothesis Testing Null Hypothesis (H_0): There is no significant difference between treatment means.

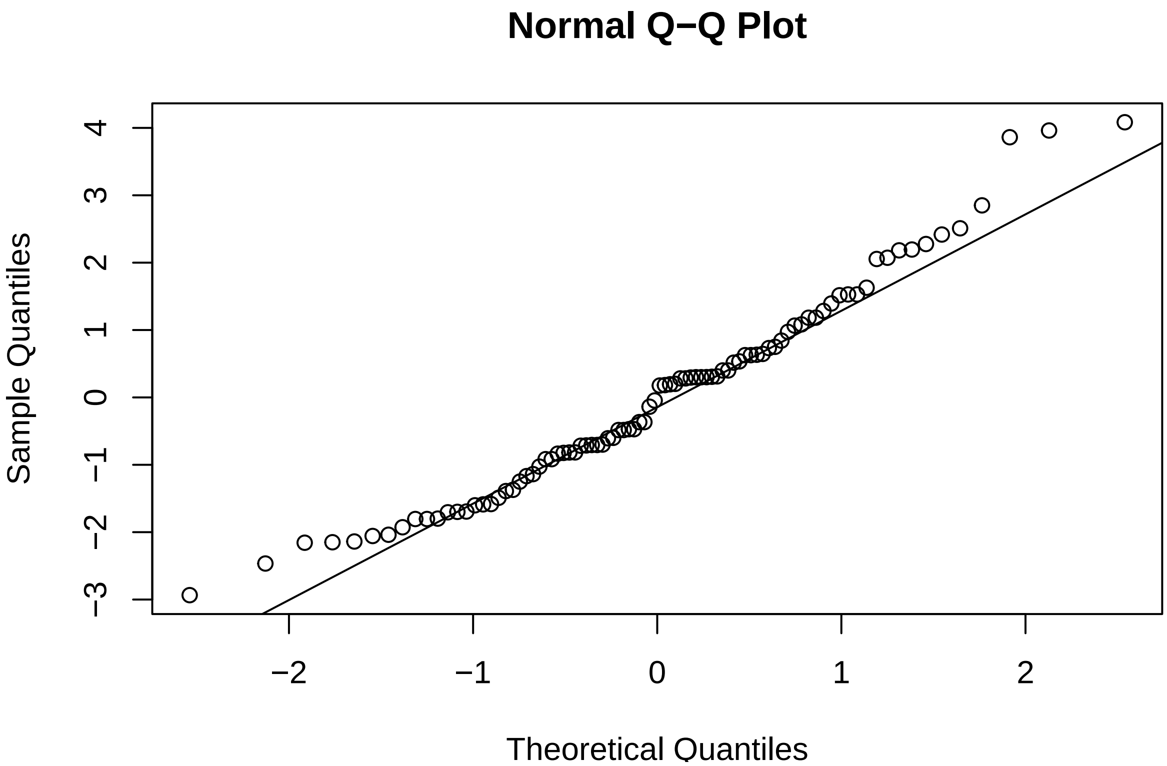
Alternative Hypothesis (H_1): At least one treatment has a different mean.

```r
res.aov <- aov(Score ~ Treatment + Person + Gender, data = df)
summary(res.aov)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Treatment    2  13.09   6.544   2.494 0.0892 .
## Person       9   5.38   0.598   0.228 0.9895
## Gender       1   0.01   0.006   0.002 0.9614
## Residuals   77 202.02   2.624
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value > 0.05, we have no significant evidence against H_0 to reject H_0, meaning there is no significant difference between treatment means.

```
residuals <- residuals(res.aov)
qqnorm(residuals)
qqline(residuals)
```



It can be seen from the qq-plot that the residuals of the datapoints are not normally distributed towards the end. they are kind of skewed. Therefore dataset has possible skewness or heavy tails

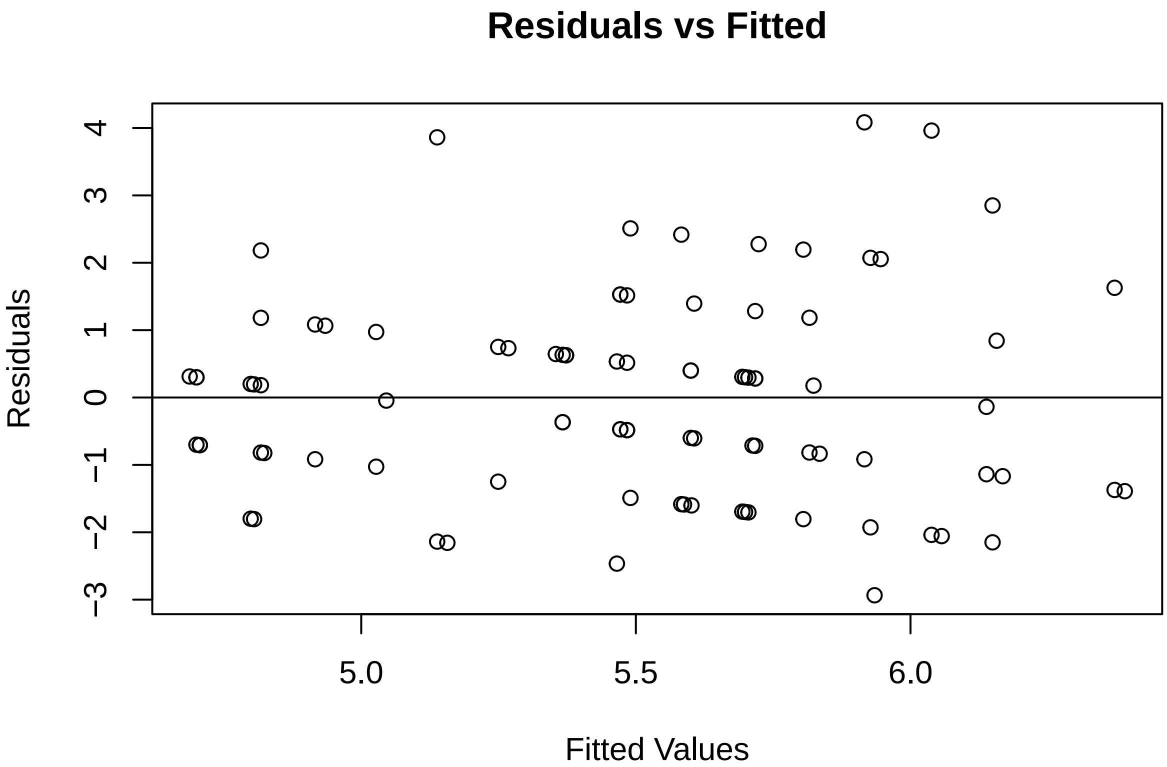
```
shapiro.test(residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.97084, p-value = 0.04054
```

To check the normality of the residues, we can do the shapiro test. After doing the test it can be interpreted with the help of the p-value of the test statistic that residuals are not normally distributed.

Now we will check for homoscedasticity.

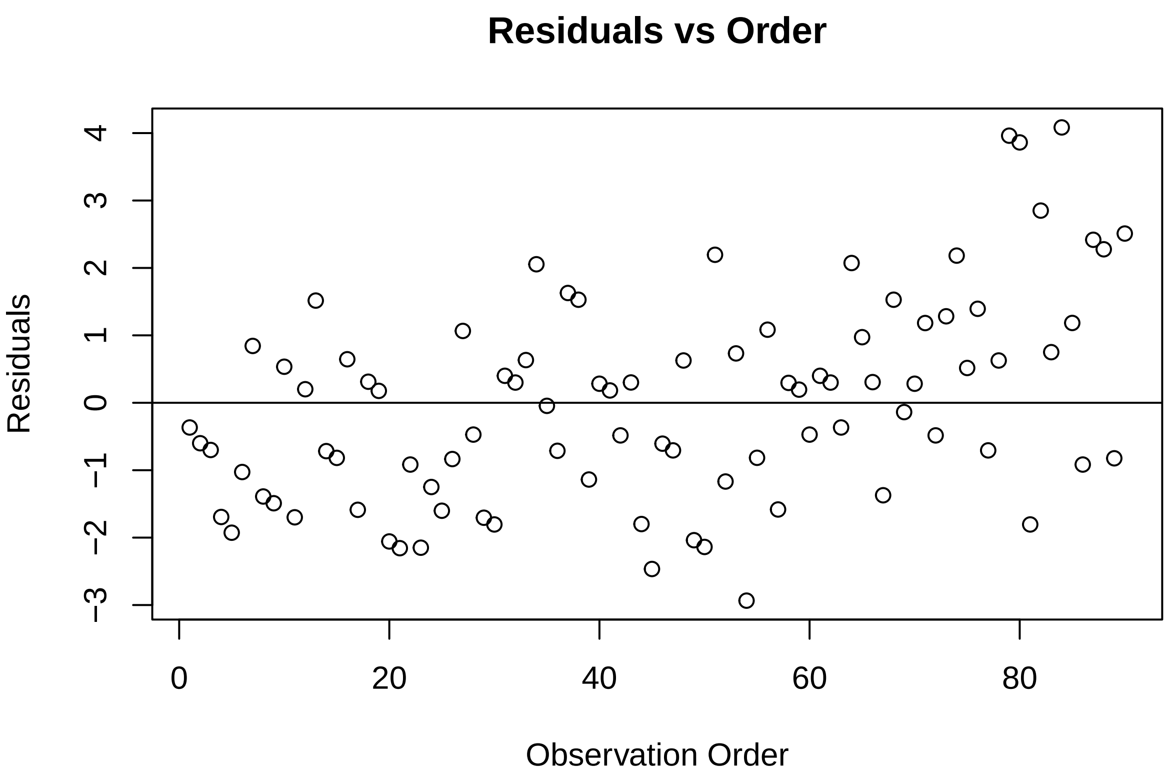
```
fitted_values <- fitted(res.aov)
plot(fitted_values, residuals,
     ylab="Residuals", xlab="Fitted Values", main="Residuals vs Fitted")
abline(h=0)
```



From the graph it can be interpreted that residuals are randomly scattered around the horizontal line with constant spread. No clear pattern is visible on the graph. therefore it can be said that the data has homoscedasticity.

Now we will check for autocorrelation

```
plot(1:length(fitted_values), residuals,
     ylab="Residuals", xlab="Observation Order", main="Residuals vs Order")
abline(h=0)
```



It can be seen that residuals appear randomly scattered around zero, with no visible patterns or trends. No clustering, cycles, or trends over time are visible in the graph. Therefore it can be said that there is no autocorrelation present in the data set.

```
TukeyHSD(res.aov, "Treatment", conf.level=0.95)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Score ~ Treatment + Person + Gender, data = df)
##
## $Treatment
##                           diff         lwr      upr     p adj
## Audiovisual-Audio  0.6666667 -0.33281491 1.666148 0.2543323
## Visual-Audio       0.9000000 -0.09948158 1.899482 0.0862168
## Visual-Audiovisual 0.2333333 -0.76614824 1.232815 0.8427193
```

From the Tukey test, it can be interpreted that the difference between the three treatments is not significant as the p-values are all greater 0.05. To confirm our hypothesis, we can see that confidence interval also includes Zero in it which means that the difference among the treatment is not significant.

Conclusion

This study investigated the effect of different sensory modalities—visual, auditory, and audiovisual—on memory recall of grocery items. We used a Randomized Complete Block Design (RCBD) to control for gender differences and ensure randomization. The data collected were analyzed using ANOVA, and the results indicated no statistically significant differences in recall performance across the three modalities. The p-value obtained was greater than 0.05, suggesting that the mode of presentation did not substantially impact short-term memory recall in this experiment.

One of the primary strengths of this study was its methodological rigor. The RCBD approach allowed us to account for potential confounding factors, and the use of ANOVA provided a robust statistical framework for analyzing variance between groups. However, despite these methodological strengths, our findings did not reveal a significant relationship between sensory modality and memory recall. This suggests that, at least within the scope of this study, individuals may not inherently recall information more effectively when it is presented in a specific sensory format.

We also ran additional tests to check the accuracy of our results. The Shapiro-Wilk test for normality indicated potential skewness in the data, suggesting that the assumption of normality was not fully met. However, homoscedasticity assessments confirmed that variance was relatively stable across groups. Additionally, the Tukey post-hoc analysis further corroborated that there were no significant pairwise differences among the treatment conditions. These findings reinforce the conclusion that no single sensory modality had a markedly superior effect on memory recall.

While the results of this study do not support a strong effect of presentation modality on short-term memory recall, they do provide valuable insights for future research. Several factors could have influenced the outcomes, including individual differences in cognitive processing, familiarity with the test materials, and variations in attention levels. Future studies may benefit from expanding the sample size and including a more diverse group of participants. They could also explore how sensory modalities affect long-term memory instead of just short-term recall. Additionally, investigating the role of contextual cues, emotional engagement, and multi-sensory integration strategies may offer further understanding of how information is best retained.

Even though we didn’t find a strong effect, this study adds to the conversation about how people learn and remember things. These findings highlight the complexity of memory retention and suggest that the effectiveness of different presentation formats may depend on additional cognitive and contextual factors. Future research can build upon these findings to explore alternative methodologies and deeper cognitive mechanisms that influence learning and recall.

References

[1] Abdulrahaman, M., Faruk, N., Oloyede, A., Surajudeen-Bakinde, N., Olawoyin, L., Mejabi, O., Imam-Fulani, Y., Fahm, A., & Azeez, A. (2020). Multimedia tools in the teaching and learning processes: A systematic review. *Heliyon*, *6*(11), e05312. https://doi.org/10.1016/j.heliyon.2020.e05312

[2] Lonn, S., & Teasley, S. D. (2009). Podcasting in higher education: What are the implications for teaching and learning? *The Internet and Higher Education*, *12*(2), 88–92. https://doi.org/10.1016/j.iheduc.2009.06.002

[3] The Guardian. (2024, September 29). *The Guardian view on the other influencers: A golden era for science education.* https://www.theguardian.com/commentisfree/2024/sep/29/the-guardian-view-on-the-other-influencers-a-golden-era-for-science-education

[4] Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158

[5] Loftus, E. F., Banaji, M. R., Schooler, J. W., & Foster, R. A. (1987). Who remembers what?: Gender differences in memory. *Michigan Quarterly Review*. Retrieved from https://projects.iq.harvard.edu/files/banaji/files/1987_loftus_mqr.pdf

[6] McDougall, G. J., Pituch, K. A., Stanton, M. P., & Chang, W. (2014). Memory performance and affect: Are there gender differences in community-residing older adults? *Issues in Mental Health Nursing, 35*(3), 212–219. https://doi.org/10.3109/01612840.2013.853332

[7] Li, R. (2024). Why women see differently from the way men see? A review of sex differences in cognition and sports. *Neuroscience & Biobehavioral Reviews, 157*, 104794. https://doi.org/10.1016/j.neubiorev.2023.104794