Alcohol Usage Not Significantly Correlated with Memory Task Performance in Adults on The Islands

Belal Hajjaj, Aaron Jin, Justin Yu, Brandon Luong, Modhar Al Qasser

2024-03-27

Author Contributions

BH, AJ, JY, BL, and MA - Conceptualization and planning of research JY, AJ, BH - Design of methodologies BH, AJ, JY, BL, and MA - Data collection and recording JY - Annotation, management, and processing of data JY - Statistical analysis JY - Visualization and data presentation BH, AJ, JY, BL, and MA - Drafting report - Review and revision of final report

Introduction

Materials and Methods

Participants were sampled from the village of Arcadia on the island of Providence. Arcadia was chosen for its large population size of 4339, which encompasses a larger range of individuals than other villages, improving generalizability. Since there was no feasible method of obtaining a list of all individuals to take a simple random sample, a multistage sampling strategy was employed. Only adults at least 19 years of age were included to ensure participants were old enough to drink and complete study tasks.

Excel functions were used to generate a list of 100 house numbers out of the 1571 houses in Arcadia. From each selected house, the adults were numbered in order of appearance. R `runif()` was then used to generate a random number, and the corresponding resident was selected.

Participants who did not provide consent, empty houses, and duplicates were removed from the data, resulting in a final sample size of n = 90.

Results

To make comparisons between the quantitative cards test memory scores of different alcohol consumption categories, side-by-side boxplots were used (Figure 1). Means were also added to the plots as an additional factor for comparison. This graphical summary allows for the preliminary visualization of potential associations between alcohol consumption frequency and cards scores. Numerical summaries were then calculated to more precisely examine the centers and sample sizes of the groups (Table 1).

```
data %>%
  # create boxplots
  ggplot(aes(x = Alcohol_consumption, y = Cards)) +
  geom_boxplot() +
  stat_summary(fun.y = mean, geom = "point", shape = 18, size=3, color="red", fill="red") +
  theme_light() +
  # adjust text angle and position
  theme(axis.text.x = element_text(angle = 10, vjust = 0.5)) +
  labs(x = "Alcohol consumption frequency", y = "Cards Memory Score (/10)")
```

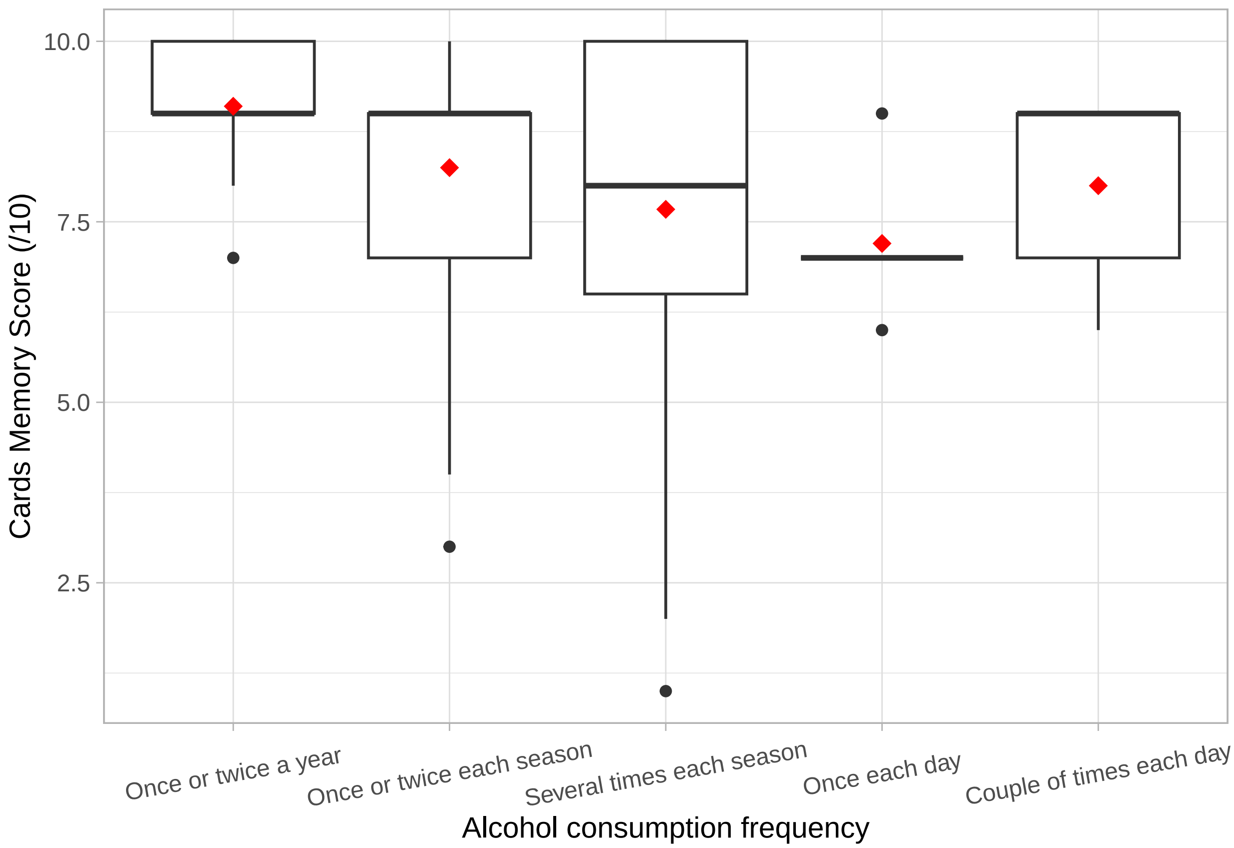

Figure 1: Cards Memory Score by Alcohol Consumption Frequency. Means are shown in red.

```
df <- data %>%
  group_by(Alcohol_consumption) %>%

  # calculate medians and means
  summarise("Number of participants" = n(),
            "Median Cards Score (/10)" = median(Cards),
            "Mean Cards Score (/10)" = mean(Cards)
  )

# use kable instead of directly displaying df for better appearance + caption
kable(df, caption = "Numerical summaries for Card Memory Score by Alcohol Consumption Frequency") %>%
  kable_styling()
```

Table 1: Numerical summaries for Card Memory Score by Alcohol Consumption Frequency

Alcohol_consumption	Number of participants	Median Cards Score (/10)	Mean Cards Score (/10)
Once or twice a year	10	9	9.100000
Once or twice each season	84	9	8.250000
Several times each season	55	8	7.672727
Once each day	5	7	7.200000
Couple of times each day	9	9	8.000000

From Figure 1, there appears to be somewhat of a negative correlation between alcohol consumption frequency and performance on cards memory tests. The medians appear to generally trend downwards with increasing consumption frequency, but the differences are small, ranging only between 7 and 9. The means behave similarly, with a slightly more clear trend. The *Couple of times each day* group does not appear to follow the trend, since it has both a higher median and mean compared to groups of less frequent alcohol consumption. The spread of data appears to vary significantly between groups, potentially due to differences in sample sizes. The boxplot for the *Once each day* group appears especially unusual due to the very small sample size (n = 5) and presence of outliers. The numerical summary in Table 1 supports the presence of a negative association, as both the median and mean cards scores generally appear to decrease with greater consumption. However, it can also be seen from Table 1 that a large majority of participants are in the *Once or twice each season* or *Several times each season* categories. Very few participants were in any of the other categories ($n \leq 10$), making these results less reliable.

Side-by-side boxplots (Figure 2) and numerical summaries (Table 2) were also used in a similar manner for the vocabulary memory tests.

```r
data %>%
  # create boxplot
  ggplot(aes(x = Alcohol_consumption, y = Vocab)) +
  geom_boxplot() +
  stat_summary(fun.y = mean, geom = "point", shape = 18, size=3, color="red", fill="red") +
  theme_light() +
  # adjust text angle and position
  theme(axis.text.x = element_text(angle = 10, vjust = 0.5)) +
  labs(x = "Alcohol consumption frequency", y = "Vocabulary Memory Score (/20)")
```

```r
df <- data %>%
  group_by(Alcohol_consumption) %>%

  # calculate medians and means
  summarise("Number of participants" = n(),
            "Median Vocab Score (/10)" = median(Vocab),
            "Mean Vocab Score (/10)" = mean(Vocab)
  )

# use kable instead of directly displaying df for better appearance + caption
kable(df, caption = "Numerical summaries for Vocabulary Memory Score by Alcohol Consumption Frequency") %>%
  kable_styling()
```

Table 2: Numerical summaries for Vocabulary Memory Score by Alcohol Consumption Frequency

Alcohol_consumption	Number of participants	Median Vocab Score (/10)	Mean Vocab Score (/10)
Once or twice a year	10	18.5	18.20000
Once or twice each season	84	16.0	15.69048
Several times each season	55	15.0	14.09091
Once each day	5	14.0	14.20000
Couple of times each day	9	13.0	14.00000

Similarly to the cards memory tests, the boxplots for the vocabulary memory test scores in Figure 2 seems to show a negative correlation with increasing alcohol consumption. It looks like there may be a stronger trend in this test, as the medians

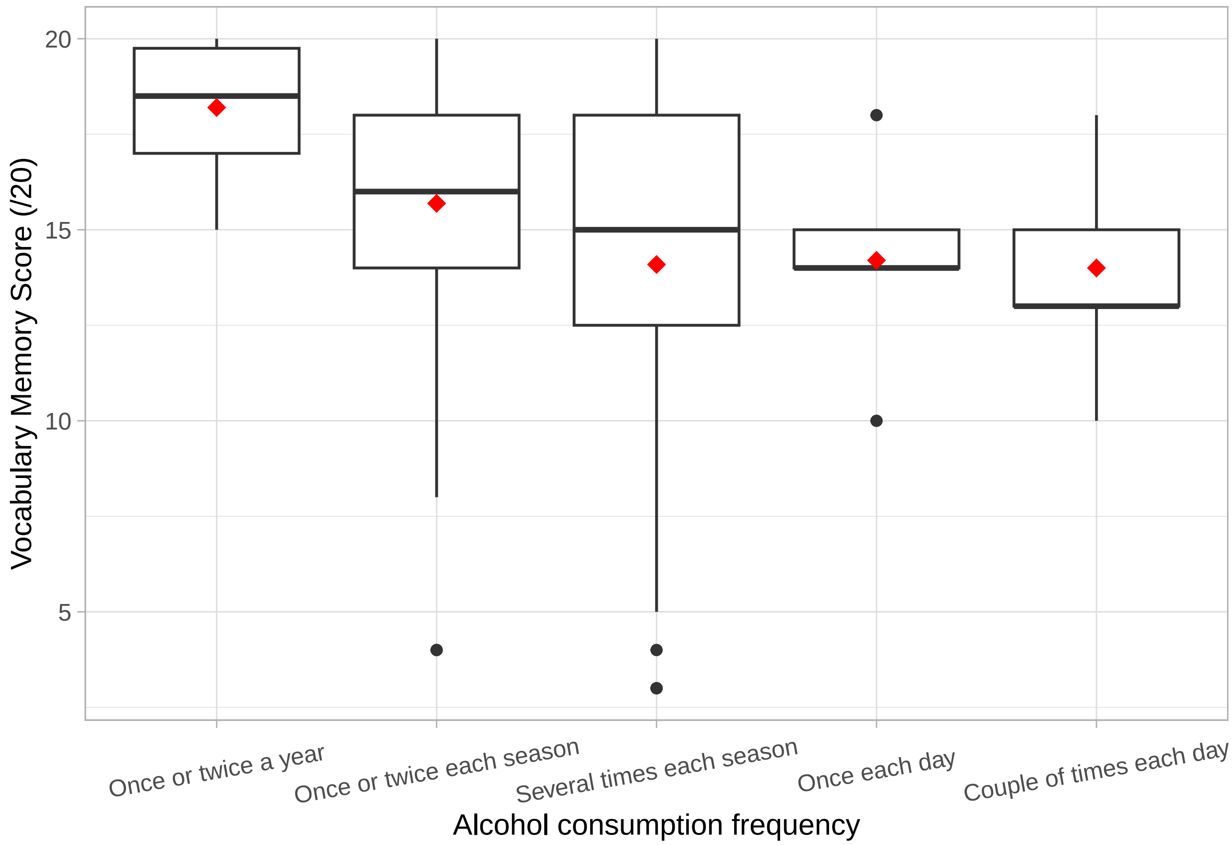

Figure 2: Side-by-side boxplots of Vocabulary Memory Score by Alcohol Consumption Frequency. Means are shown in red.

are more significantly different from each other, ranging from 13 to 18.5. The scoring scheme out of this test is out of 20 rather than 10, which may contribute to better sensitivity. The means follow a similar trend as the medians, but the means for the higher consumption groups are very close to each other. Interestingly, the *Couple of times each day* group does follow the trend for these scores, having the lowest median and mean. Once again, differences in the spread of data are observed between groups, possibly because of varying sample sizes. Table 2 supports the presence of a negative association between alcohol consumption frequency and vocabulary memory scores as well, with the medians and means generally lower for greater consumption.

Self-reported forgetfulness and alcohol consumption frequency are both categorical variables. A two-way table was first produced to examine the data (Table 3). Due to the varying sample sizes, many categories, and presence of cells with no observations, a mosaic plot would not be suitable this data. So, in order to see whether forgetfulness varies with alcohol consumption, a stacked bar chart of the conditional distributions of forgetfulness by alcohol consumption was created (Figure 3).

```r
# reorder forgetful levels
forgetfuldat <- data %>%
  mutate(Forgetful = factor(Forgetful, levels = c("Not at all", "A little", "Moderately")))

# create two-way table
forgetfultab <- table(forgetfuldat$Alcohol_consumption, forgetfuldat$Forgetful)

# use kable instead of directly displaying for better appearance + caption
kable(forgetfultab, caption = "Two-way table of Forgetfulness by Alcohol Consumption Frequency") %>%
  kable_styling()
```

Table 3: Two-way table of Forgetfulness by Alcohol Consumption Frequency

	Not at all	A little	Moderately
Once or twice a year	3	7	0
Once or twice each season	18	53	13

Several times each season	14	33	8
Once each day	0	2	3
Couple of times each day	1	7	1

```r
data %>%
  # order levels
  mutate(Forgetful = factor(Forgetful, levels = c("Moderately", "A little", "Not at all"))) %>%
  ggplot(aes(x = Alcohol_consumption, fill = Forgetful)) +
  # use position = "fill" for conditional distributions
  geom_bar(position = "fill") +
  theme_light() +
  labs(x = "Alcohol consumption frequency", y = "Proportion") +
  theme(axis.text.x = element_text(angle = 10, vjust = 0.5))
```

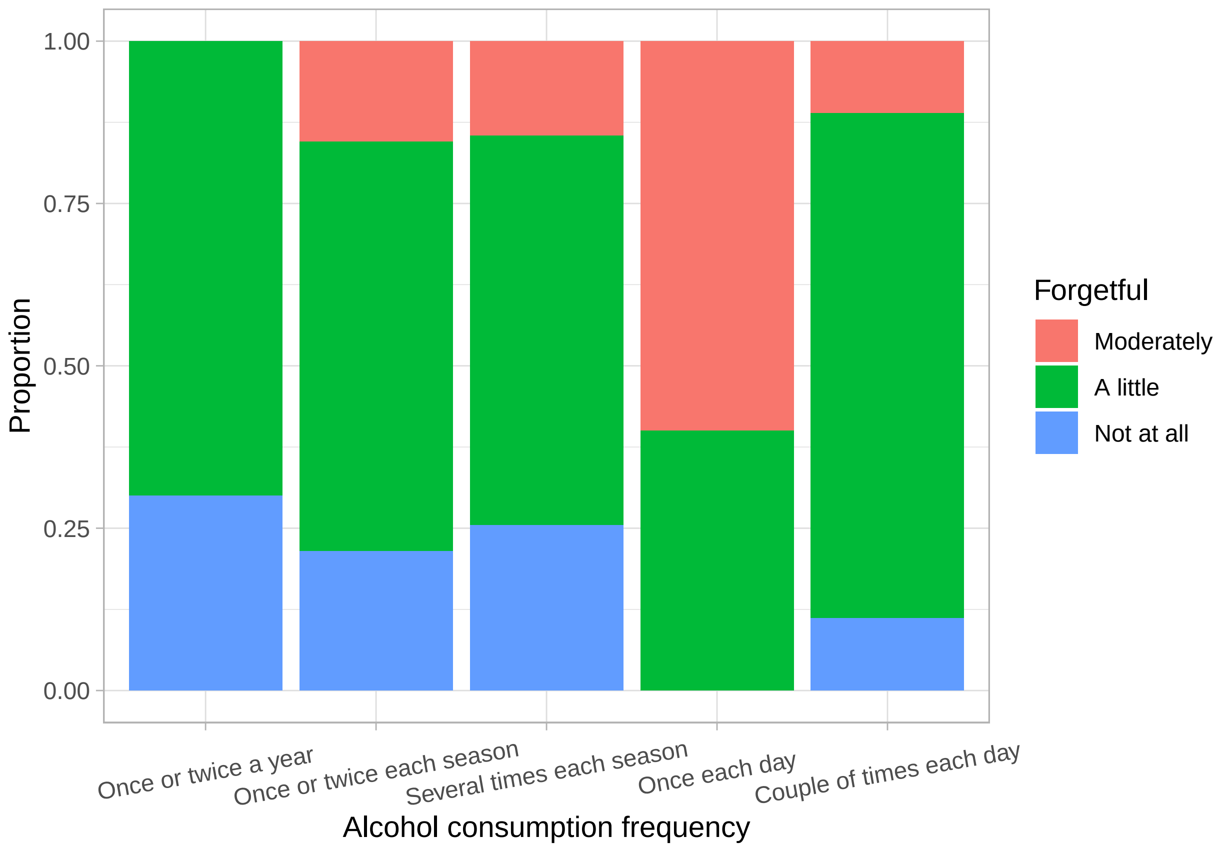

Figure 3: Conditional distributions of Forgetfulness by Alcohol Consumption Frequency

There doesn't appear to be a substantial association from the conditional distribution bar chart in Figure 3. The *Once each day* group has a much greater proportion of moderate forgetfulness, but this may be due to chance or sampling error. Table 3 shows that there were only 5 total responses which reported drinking *Once each day*, and of those, only 3 reported *Moderate* forgetfulness. From Figure 3, it can also be seen that the proportion of *Moderate* responses actually decreases with increasing alcohol consumption over several other groups. However, the lowest alcohol consumption group, *Once or twice a year*, does have the highest proportion of *Not at all* forgetful responses and no *Moderate* forgetfulness responses. Additionally, the two highest consumption groups have the lowest proportions of *Not at all* responses.

Although it looks like there may be an association between memory game scores and alcohol consumption frequency based on the graphical and numerical summaries, it is possible that these results are due to chance and this correlation isn't present in the population. To determine the extent to which these data actually provide evidence of an association, a one-way ANOVA was used to test for significant differences between mean cards and vocabulary memory test scores across alcohol consumption groups. A 5% significance level was used, as there is no clear need to set a different significance level to minimize a specific type of error. If no significant differences are detected, the data does not provide evidence against the absence of a correlation; it would not be possible to conclude that there is a real correlation between alcohol consumption and memory test scores. If one or more significant differences are detected, the data indicates that there may be some correlation between alcohol consumption and memory scores. Further tests could then be used to determine which groups significantly differ from each other.

A simple random sample could not be achieved, so participants were randomly recruited through a multistage sample from all houses in Arcadia, then all adults from each house. Although there may not be an equal chance of each individual being selected due to differences in the number of residents in each house, there should be no dependency between individuals since the sample randomly selected both houses and residents. Since this was an observational study and participants were not assigned to groups beforehand, the groups should be independent from each other. To check if the population σs might be the same for the groups, the sample standard deviations (Tables 4-5) and strip plots of residuals (Figures 4-5) were examined for both memory tests.

```r
# calculate n and stdev for vocab
sd_cards <- data %>%
  group_by(Alcohol_consumption) %>%
  summarise(n = n(), sd = sd(Cards))

# display via kable
kable(sd_cards, caption = "Standard deviations of Cards Memory Scores by Alcohol Consumption Frequency") %>%
  kable_styling()
```

Table 4: Standard deviations of Cards Memory Scores by Alcohol Consumption Frequency

Alcohol_consumption	n	sd
Once or twice a year	10	0.9944289
Once or twice each season	84	1.6122647
Several times each season	55	2.4271195
Once each day	5	1.0954451
Couple of times each day	9	1.3228757

```r
# calculate n and stdev for vocab
sd_vocab <- data %>%
  group_by(Alcohol_consumption) %>%
  summarise(n = n(), sd = sd(Vocab))

# display via kable
kable(sd_vocab, caption = "Standard deviations of Vocab Memory Scores by Alcohol Consumption Frequency") %>%
  kable_styling()
```

Table 5: Standard deviations of Vocab Memory Scores by Alcohol Consumption Frequency

Alcohol_consumption	n	sd
Once or twice a year	10	1.686548
Once or twice each season	84	3.400837
Several times each season	55	4.819992
Once each day	5	2.863564
Couple of times each day	9	2.500000

```r
# run ANOVA first to get residuals
anova_cards <- aov(data$Cards ~ data$Alcohol_consumption)

# create dataframe for strip plot
strip <- data.frame(Alcohol_consumption = data$Alcohol_consumption, Residuals = anova_cards$residuals)

# create strip plot
strip %>%
  # color code
```

```r
ggplot(aes(Alcohol_consumption, Residuals, colour = Alcohol_consumption)) +
# hide color legend, ensure 0 vertical jitter, reduce horizontal jitter
geom_jitter(show.legend = F, height = 0, width = 0.2) +
theme_light() +
# rotate labels
theme(axis.text.x = element_text(angle = 10, vjust = 0.5)) +
labs(x = "Alcohol consumption frequency")
```

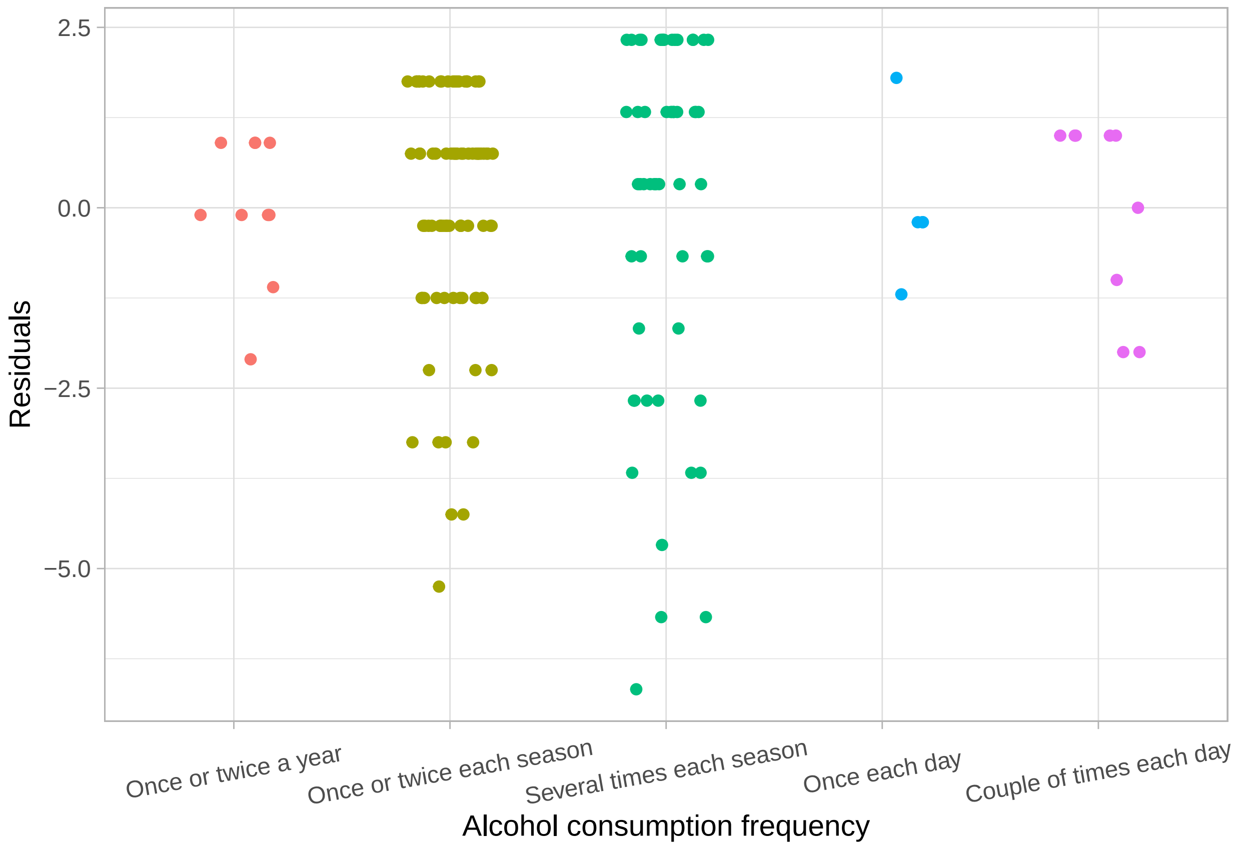

Figure 4: Strip plot of residuals for Cards memory test.

```r
# run ANOVA first to get residuals
anova_vocab <- aov(data$Vocab ~ data$Alcohol_consumption)

# create dataframe for strip plot
strip <- data.frame(Alcohol_consumption = data$Alcohol_consumption, Residuals = anova_vocab$residuals)

# create strip plot
strip %>%
  # color code
  ggplot(aes(Alcohol_consumption, Residuals, colour = Alcohol_consumption)) +
  # hide color legend, ensure 0 vertical jitter, reduce horizontal jitter
  geom_jitter(show.legend = F, height = 0, width = 0.2) +
  theme_light() +
  # rotate labels
  theme(axis.text.x = element_text(angle = 10, vjust = 0.5)) +
  labs(x = "Alcohol consumption frequency")
```

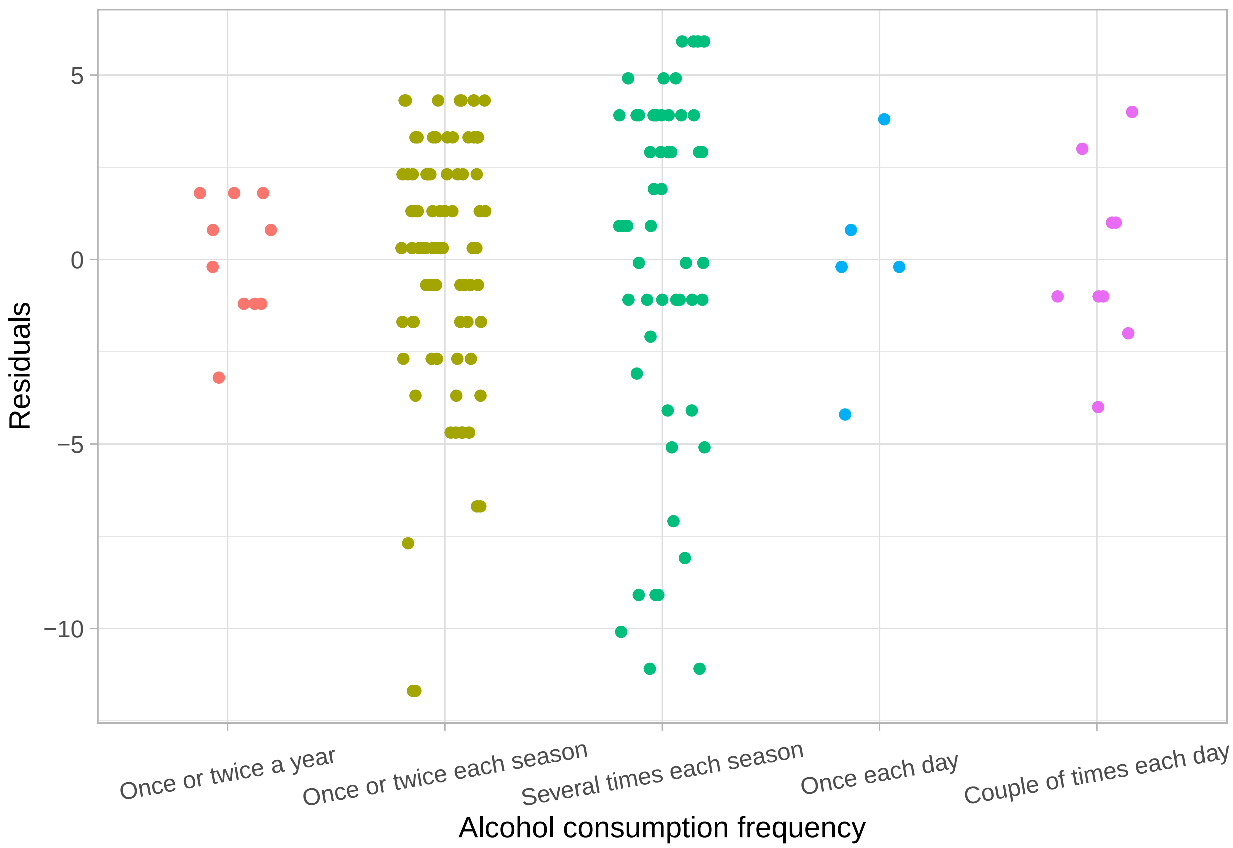

Figure 5: Strip plot of residuals for Vocabulary memory test.

From Tables 4-5 and Figures 4-5, the groups look like they may belong to populations with different standard deviations. As a result, it may not be appropriate to assume the observed ANOVA F statistic will follow an F distribution, and the ANOVA p-value may not be reliable.

To check for normality in the data, normal Q-Q plots and histograms of the residuals for each memory test were used (Figure 6-9).

```r
qqnorm(anova_cards$residuals, main = NULL)
qqline(anova_cards$residuals)
```

```r
hist(anova_cards$residuals, breaks = 10, main = NULL, xlab = "Residuals")
```

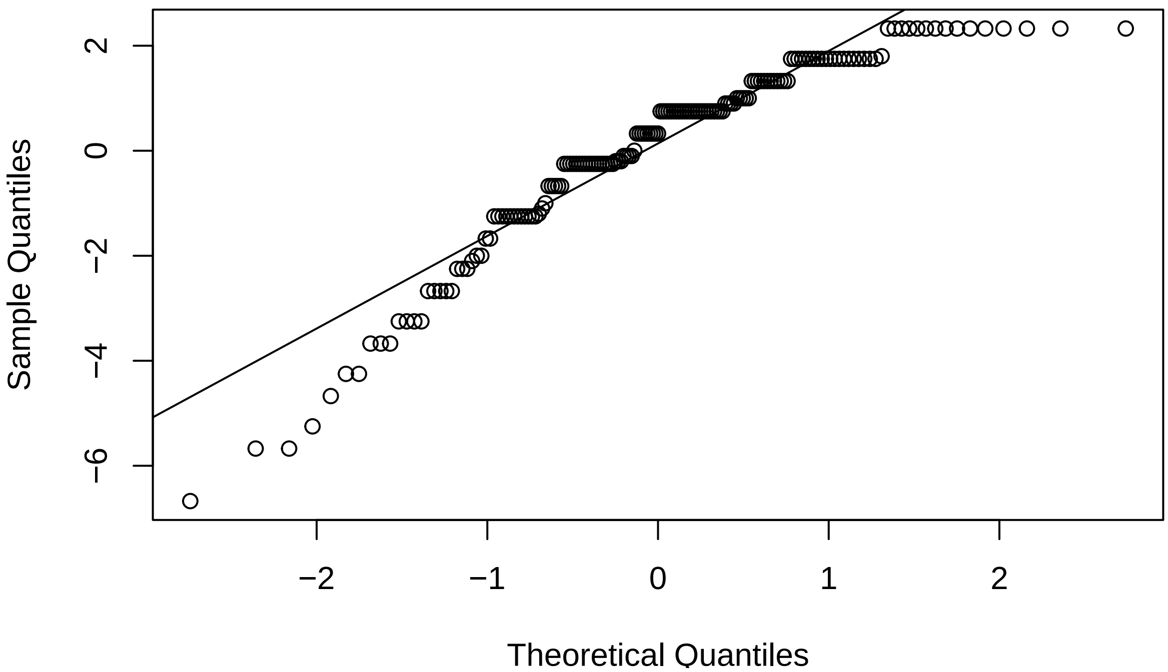

Figure 6: Normal Q-Q Plot of residuals for Cards memory test.

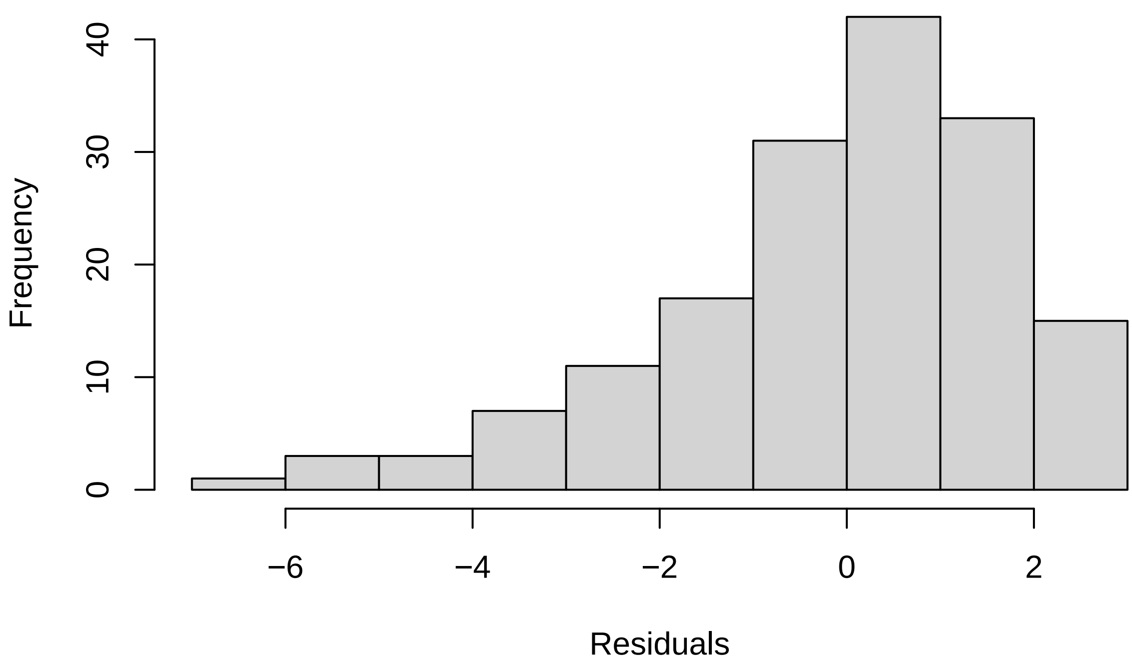

Figure 7: Histogram of residuals for Cards memory test.

```r
qqnorm(anova_vocab$residuals, main = NULL)
qqline(anova_vocab$residuals)
```

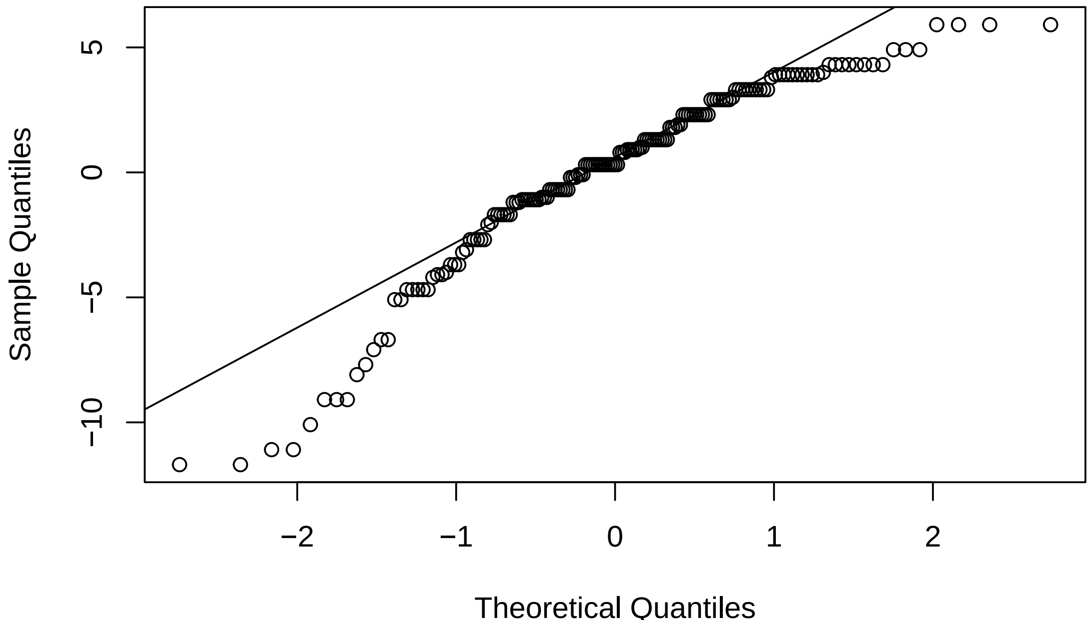

Figure 8: Normal Q-Q Plot of residuals for Vocabulary memory test.

```r
hist(anova_vocab$residuals, breaks = 10, main = NULL, xlab = "Residuals")
```

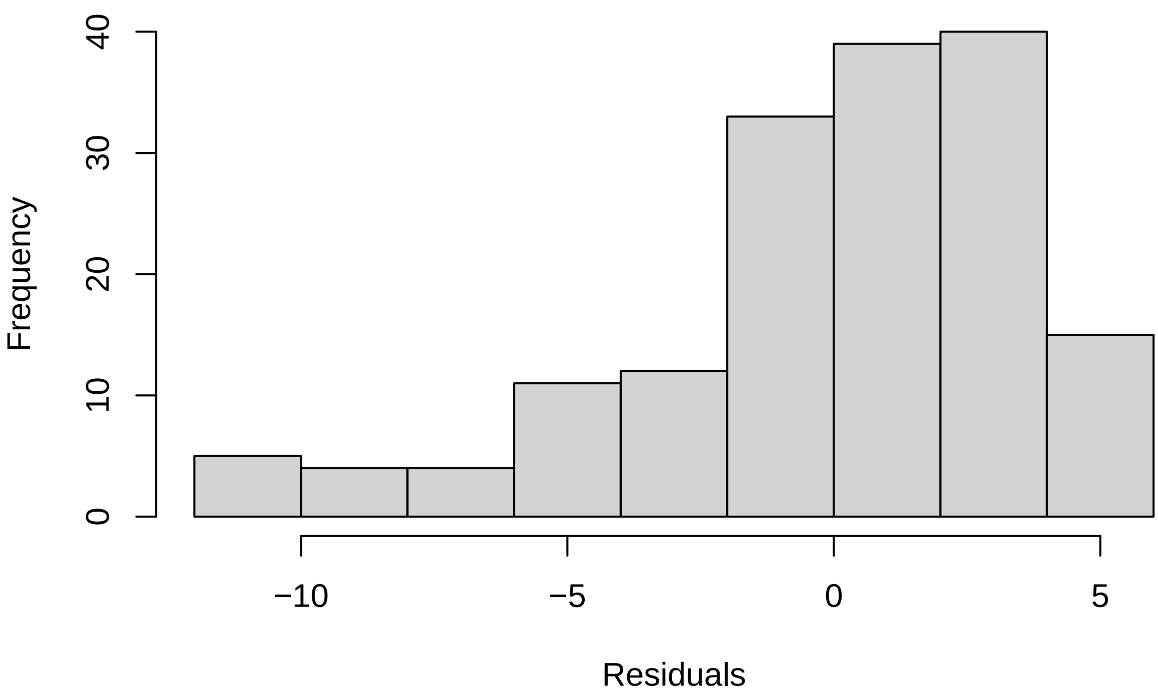

Figure 9: Histogram of residuals for Vocabulary memory test

From Figures 6-9, the data for both tests appears to be significantly left skewed and not normally distributed. This is another factor contributing to the potential unreliability of a one-way ANOVA test on this data.

Since the normality and equal variance conditions for the one-way ANOVAs were violated, a non-parametric test may be a better alternative. The Kruskal-Wallis rank sum test is a non-parametric analogue to one-way ANOVA which does not assume normality and equal variance, and only requires independence (Kruskal and Wallis 1952).

```
# ANOVA F tests
anova_cards <- aov(data$Cards ~ data$Alcohol_consumption)
anova_cards_sum <- summary(anova_cards)

anova_vocab <- aov(data$Vocab ~ data$Alcohol_consumption)
anova_vocab_sum <- summary(anova_vocab)
```

```
# display results - process shown in full code
kable(anova_results, caption = "One-way ANOVA F test results") %>% kable_styling()
```

Table 6: One-way ANOVA F test results

Test Type	F statistic	P value
Cards one-way ANOVA	1.816697	0.1282062
Vocab one-way ANOVA	3.326610	0.0120024

```
# kruskal tests
krus_cards <- kruskal.test(Cards ~ Alcohol_consumption, data = data)
krus_vocab <- kruskal.test(Vocab ~ Alcohol_consumption, data = data)
```

```
# display results - process shown in full code
kable(krus_results, caption = "Kruskal-Wallis rank sum test results") %>% kable_styling()
```

Table 7: Kruskal-Wallis rank sum test results

Test Type	Kruskal-Wallis chi-squared	P-value
Cards Kruskal-Wallis	6.90348	0.1410776
Vocab Kruskal-Wallis	14.12826	0.0068968

Since a significant difference was observed for the mean vocabulary memory test scores in the one-way ANOVA and Kruskal-Wallis tests, a pairwise t-test using the Bonferroni correction was used to examine which groups have significantly different mean scores.

```
# run pairwise t test
# use as.data.frame to store p-values in data frame for presentation
pairwise <- as.data.frame(pairwise.t.test(data$Vocab, data$Alcohol_consumption, p.adj = "bonf")$p.value)

# replace NAs with - for clarity
pairwise[is.na(pairwise)] <- "-"

# display results
kable(pairwise, caption = "Pairwise comparisons for Vocab scores by Alcohol Consumption using t tests and Bonf
```

Table 8: Pairwise comparisons for Vocab scores by Alcohol Consumption using t tests and Bonferroni correction

	Once or twice a year	Once or twice each season	Several times each season	Once each day
Once or twice each season	0.5216892	-	-	-
Several times each season	0.0216865	0.173250224354024	-	-
Once each day	0.5864402	1	1	-
Couple of times each day	0.1831194	1	1	1

Considering the two categorical variables of alcohol consumption and self-reported forgetfulness, a Chi-squared test would be appropriate to test whether the two variables appear to be independent or dependent. A small p-value from this test provides evidence that the two variables are dependent, suggesting that there is some correlation between them.

The Chi-squared test requires independent observations, which was established when evaluating the one-way ANOVA conditions above. The counts can sucessfully be organized into a two-way table, as seen in Table 3. Also from Table 3, it is apparent that there are several cells which have less than 5 observations and even some cells with 0 observations. So, the Chi-squared test results may not be accurate.

Since the sample size conditions of the Chi-squared test were not met, a Monte Carlo simulation Chi-squared test may be more appropriate. This test uses randomization, and does not require >5 observations in each cell (Hope 1968).

```r
# normal chi-squared
chis <- chisq.test(data$Alcohol_consumption, data$Forgetful)

# monte carlo
# set seed for reproducibility
set.seed(2024)
monte <- chisq.test(data$Alcohol_consumption, data$Forgetful, sim = T)
```

```r
# display results - process shown in full code
kable(chis_results, caption = "Self-reported Forgetfulness Chi-squared results") %>% kable_styling()
```

Table 9: Self-reported Forgetfulness Chi-squared results

Test Type	Chi-squared statistic	P-value
Chi-squared	11.23233	0.1888796
Monte Carlo Simulation	11.23233	0.1814093

Conclusions

Move some of the interpretation of summaries from results section here. I'm not sure which

Since the p-value for the one-way ANOVA on mean Cards memory scores (0.128) is greater than the significance level of 0.05, there is no evidence against the mean Cards score for all alcohol consumption frequencies being the same. So, this data does not provide evidence for a correlation between alcohol consumption frequency and Cards memory test scores.

As with the previous one-way ANOVA, the p-value for the Kruskal-Wallis rank sum test on Cards memory scores (0.1411) is greater than the significance level of 0.05. This result does not provide evidence against there being no correlation between alcohol consumption frequency and Cards memory test scores. Note that the Kruskal-Wallis rank sum test should not be interpreted as detecting differences between means when the distributions of groups are significantly different, but rather detecting differences between the groups in general (Kruskal and Wallis 1952).

The p-value for the one-way ANOVA on mean vocabulary memory scores (0.012) is smaller than the significance level of 0.05, but it is not extremely small. Therefore, this data provides moderate evidence against the mean vocabulary score for all alcohol consumption frequencies being the same. As a result, this data does provide evidence for some association between alcohol consumption frequency and vocabulary memory test scores.

The p-value for the Krukal-Wallis rank sum test on vocabulary memory scores (0.006897) is substantially smaller than the p-value from the corresponding one-way ANOVA (0.012). This suggests the data may actually provide strong evidence for a significant difference in vocabulary memory test performance for different alcohol consumption frequencies.

Discussion

References

Hope, Adery C. A. 1968. "A Simplified Monte Carlo Significance Test Procedure." *Journal of the Royal Statistical Society: Series B (Methodological)* 30 (3): 582–98. https://doi.org/10.1111/j.2517-6161.1968.tb00759.x.

Kruskal, William H., and W. Allen Wallis. 1952. "Use of Ranks in One-Criterion Variance Analysis." *Journal of the American Statistical Association* 47 (260): 583–621. https://doi.org/10.1080/01621459.1952.10483441.

Appendix

Dataset

Variable Name	Description
House	House number of participant. Corresponds to house numbers in Arcadia on The Islands. Randomly generated using Excel.
Name	Name of participant.
Alcohol_consumption	Alcohol consumption frequency. As reported by the participant after being asked "How often do you drink alcohol?". Responses fall into one of the following categories: ("Rarely/Once or twice a year", "Once or twice each season", "Several times each season", "Once each day", "Couple of times each day"). Categorical ordinal variable.
Cards	Cards memory test score. Number of cards recalled in one minute, out of 10 drawn from a 52 card deck. Numerical discrete variable.
Vocab	Vocabulary memory test score. Number of words recalled in 30 seconds, out of a list of 20 words seen for 1 minute. Numerical discrete variable.
Forgetful	Subjective forgetfulness. As reported by the participant after being asked "How forgetful do you feel right now?". Responses fall into one of the following categories: ("Not at all", "A little", "Moderately"). Categorical ordinal variable.

The dataset can be found here.