MP2

Grace Rade, Carol Milton, Nikki Lin

2023-04-20

Introduction

It's usually pretty obvious what adults can do that children can't. Adults can drive, pay for what they want, work, drink alcohol, travel on their own, and so on. If a child does these things, then it is either illegal or socially frowned upon. Maybe less commonly discussed are what children can do that adults can't or rather shouldn't. Children can take off their clothes in public, speak at high volumes despite being indoors, pee in the swimming pool, have nap times in school, and many other things that are usually things that, if done by adults, would be inappropriate or seen as odd.

When taking another look at these actions, it's not that adults or children are strictly unable to do all of them (yes, children can drive if there are cars designed that way but we can't say that they can drive *well*), but rather that societal norms and regulations restrict people based on their ages. With increasing age comes better decision making ability and self control ("Goal B: Better Understand the Effects of Personal, Interpersonal, and Societal Factors on Aging, Including the Mechanisms Through Which These Factors Exert Their Effects" n.d.), so it makes sense that it is societally more acceptable for children to act immaturely or inappropriately compared to adults. Even so, just how much of an influence does a person's age and behavior have on others' opinions of that person? This project aims to study this by examining how peoples' opinion of a character in a vignette changes based on the stated age and the described behavior of that character.

Respondents read a vignette about a generic character, who performs the action of baking a cake while being described as a certain age and exhibiting either good or bad behavior. Two groups read vignettes where the character is stated to be a child, while two read vignettes where the character is an adult. Additionally, two groups read vignettes where the character exhibits bad behavior while two read vignettes where the character exhibits good behavior. A total of four respondent groups will read vignettes that contain one combination of age and behavior. All four combinations of age and behavior will be represented in the survey and are randomly assigned to respondents.

Hypotheses

Before conducting the analysis, the expectations for the results are below:

1. Respondents will exhibit more positive attitudes towards the vignette character of younger age.

a. Respondents who have viewed the vignettes featuring Sam as a 10 year old child will give higher levels of agreement to the statement "I like Sam".

2. Respondents will exhibit less positive attitudes towards the vignette character who displays negative behavior.

a. Respondents who have viewed the vignette featuring Sam leaving a mess in the kitchen and not sharing the cake will give lower levels of agreement to the statement "I like Sam".

3. Age will have a moderating effect on the attitudes of respondents based on behavior type.

a. Respondents will exhibit the most positive attitude towards the vignette character of younger age and positive behavior.

Methodology

We conducted a two-way randomized survey experiment to investigate our hypotheses. The survey was distributed to Smith College students, and we received a total of 58 complete, consenting responses. The two treatment factors, age and behavior, were randomized by a series of vignettes. Each factor has a similar number of respondents. 10 Good has 13 responses, 45 Bad has 15 responses, and 10 Bad and 45 Good each have 14 responses. The dependent variable is the opinion of the character, Sam, in the vignette, ascertained via a five-point likert scale agreement to the statement, "Sam seems likable". A score of 1 represents strong disagreement with the statement, and a score of 5 represents Strong Agreement with the statement.

Each respondent was shown one of four vignettes, which varied on two aspects: the age of the character in the vignette, and the type of behavior they displayed. All four vignettes featured the character Sam performing the action of baking a cake. Sam is stated to either be ten years old or forty-five years old. Good behavior is cleaning up the kitchen after baking and sharing the cake with friends. Bad behavior is leaving a mess in the kitchen and eating the cake alone. Four vignettes capture the total number of combinations of the two variables.

Results

Checking Anova Conditions

We begin with checking the conditions for Anova testing. We know that the response variable, how the much the participant likes Sam, is continuous while the factors, Sam's 'good' or 'bad' actions and Sam's age, are categorical, which satisfies the condition about the types of variables. We then check for constant variance.

Constant Variance
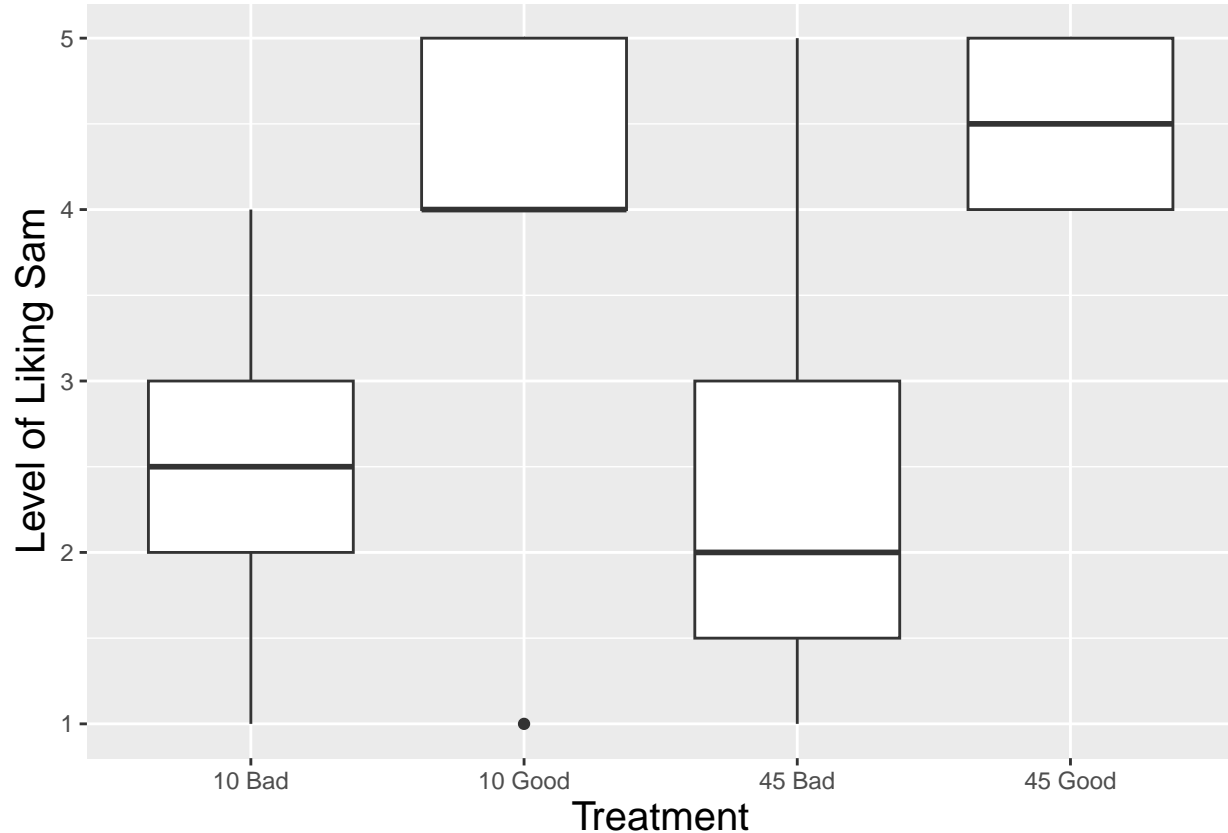
The box plot shows overall similar variability in terms of the 'boxes', however the 'whiskers' are of varying lengths, The distribution for the category in which Sam is 10 years old and does good behavior appears to be non-normal since the mean is at the end of the 'box' and there is also an outlier. The condition for constant variance must be further verified.

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  3  0.9618 0.4178
##       52
```

The Levene's Test for Homogeneity shows that the p-value is greater than alpha (0.05), so we fail to reject the null hypothesis that there is no constant variance.

```
## # A tibble: 4 x 4
##   condition Group_mean Sample_size Standard_dev
##   <fct>          <dbl>       <int>        <dbl>
## 1 10 Bad          2.43          14         1.02
## 2 10 Good         4.15          13         1.07
## 3 45 Bad          2.27          15         1.10
## 4 45 Good         4.5           14         0.519
```

The box plot and the table with the standard deviations show that there is a slight violation of the constant variance assumption, especially for the category in which Sam is 10 years old and does good behavior. However, the Levene's Test indicates that the variance across groups is not statistically different, so we will proceed.
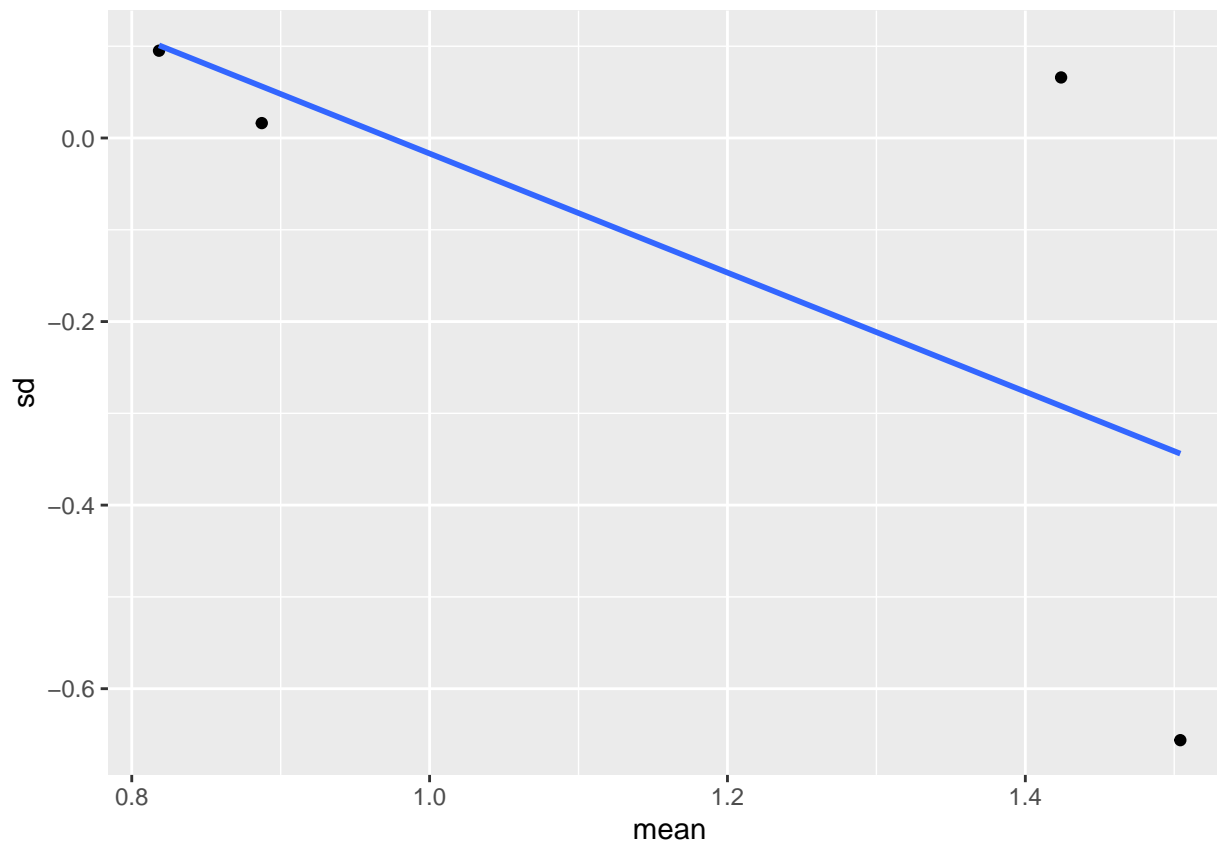
Normality
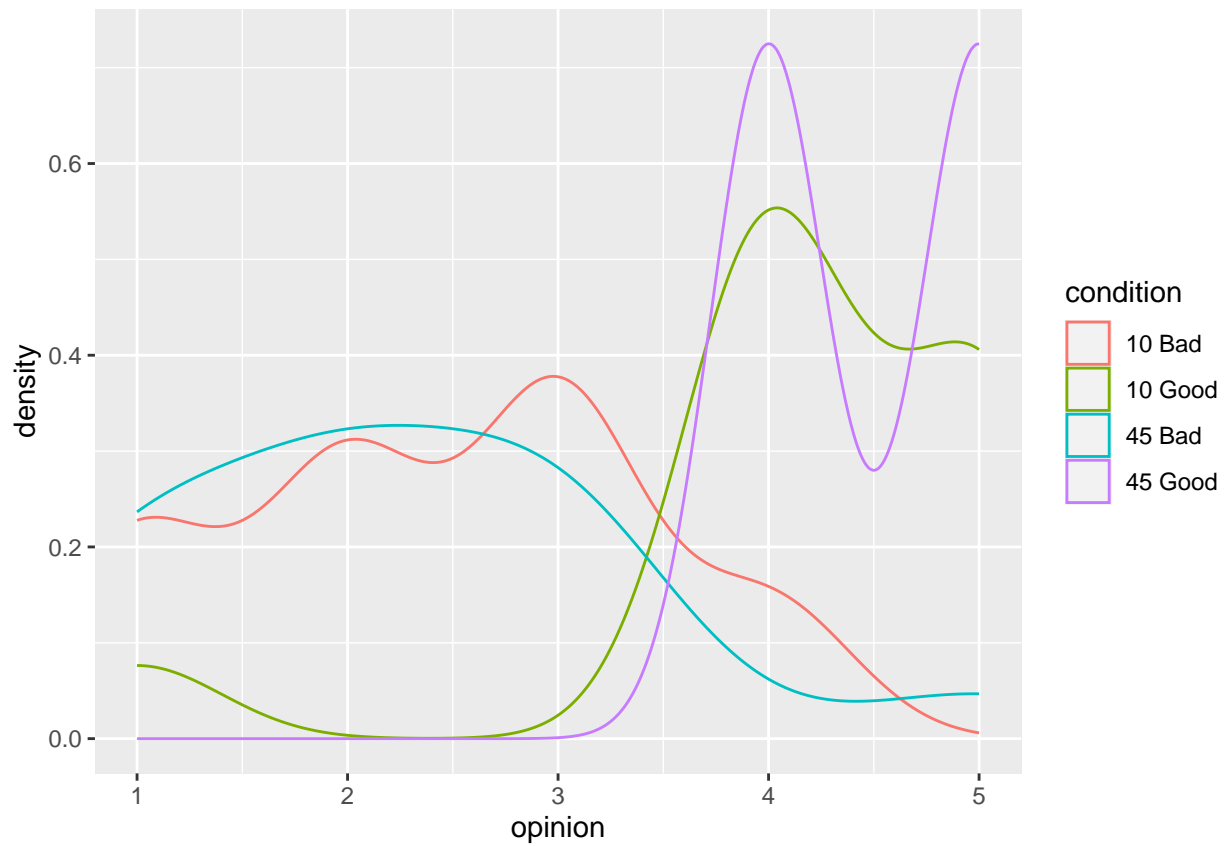
There are clear violations of normality for all combinations of factors (age and behavior).
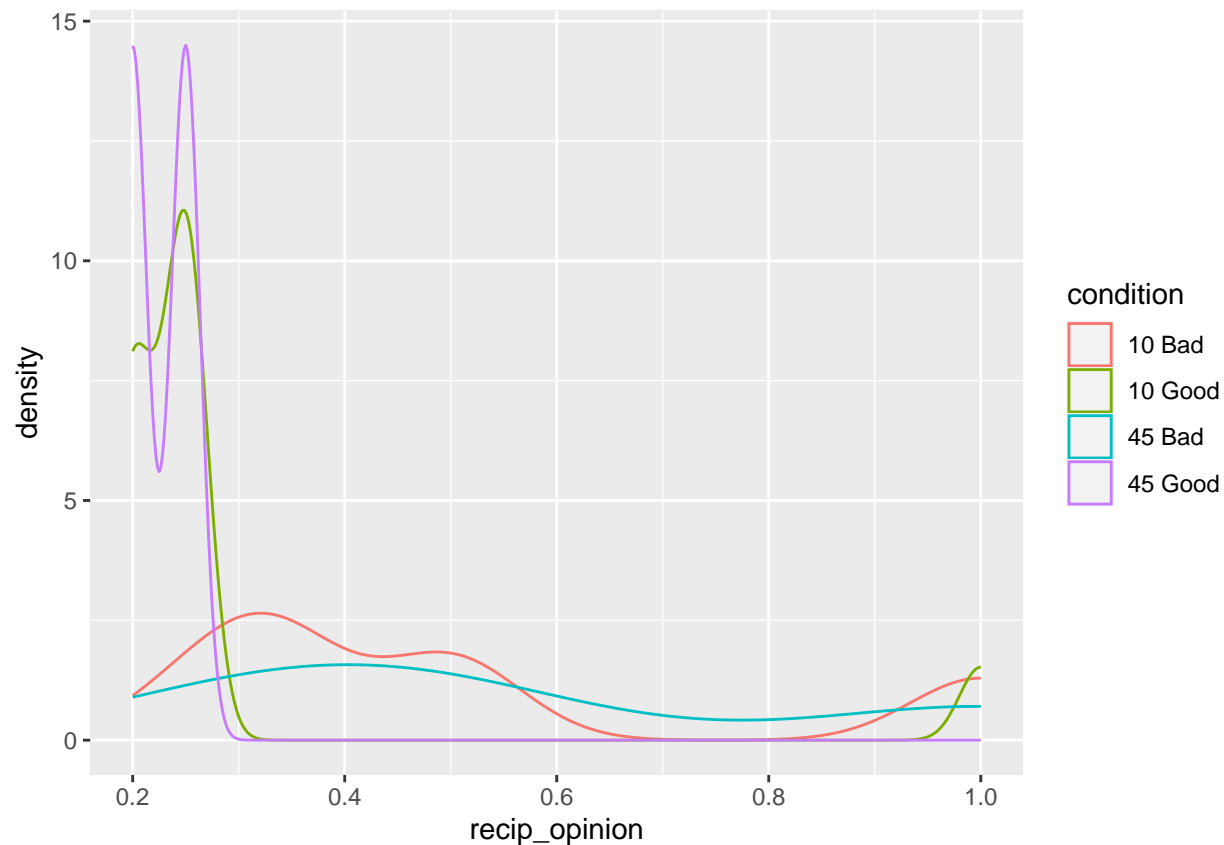
Transformation

We now check to see if a transformation of the responses that can help satisfy the condition of normality. First, we plot the mean and standard deviations for the response variable, liking of Sam, across the treatment groups.

Using the means and standard deviations, we create a linear model and look at the slope.

```
##
## Call:
## lm(formula = sd ~ mean, data = summary_stats)
##
## Residuals:
##         1         2         3         4
## -0.039971  0.357967 -0.005833 -0.312164
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.6318     0.6561   0.963    0.437
## mean         -0.6487     0.5474  -1.185    0.358
##
## Residual standard error: 0.3371 on 2 degrees of freedom
## Multiple R-squared:  0.4126, Adjusted R-squared:  0.1189
## F-statistic: 1.405 on 1 and 2 DF,  p-value: 0.3577
```
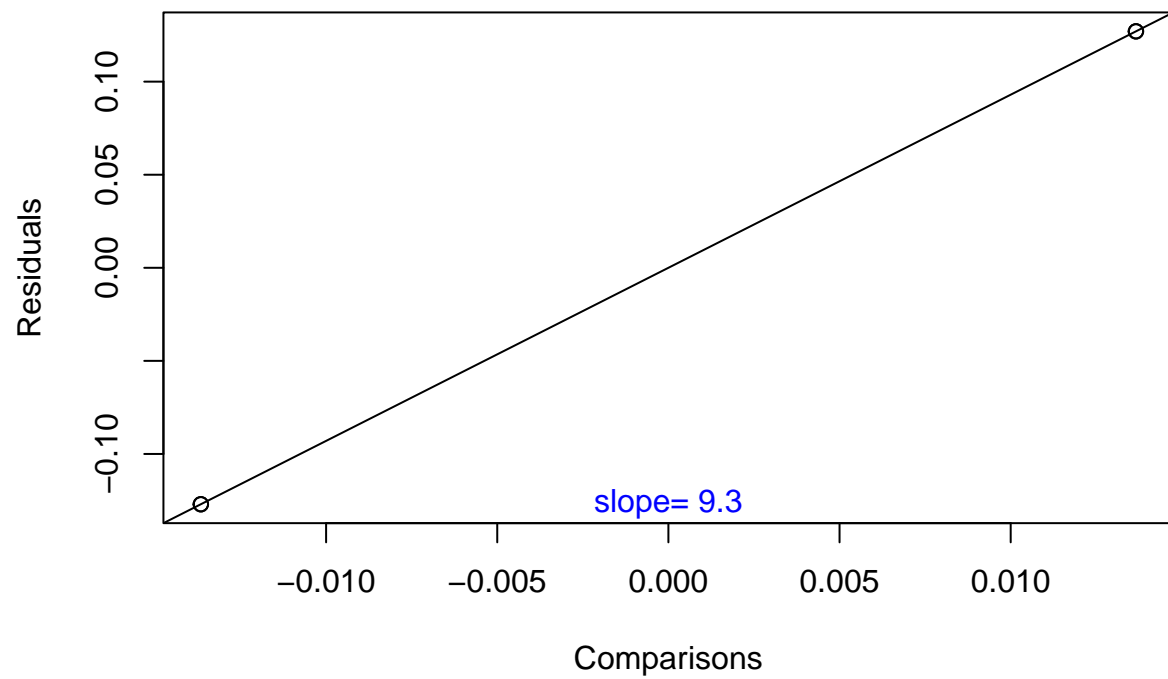
The slope is about -0.65 and so 1-slope is 1.65, thus, we can try to see if a reciprocals transformation can help.
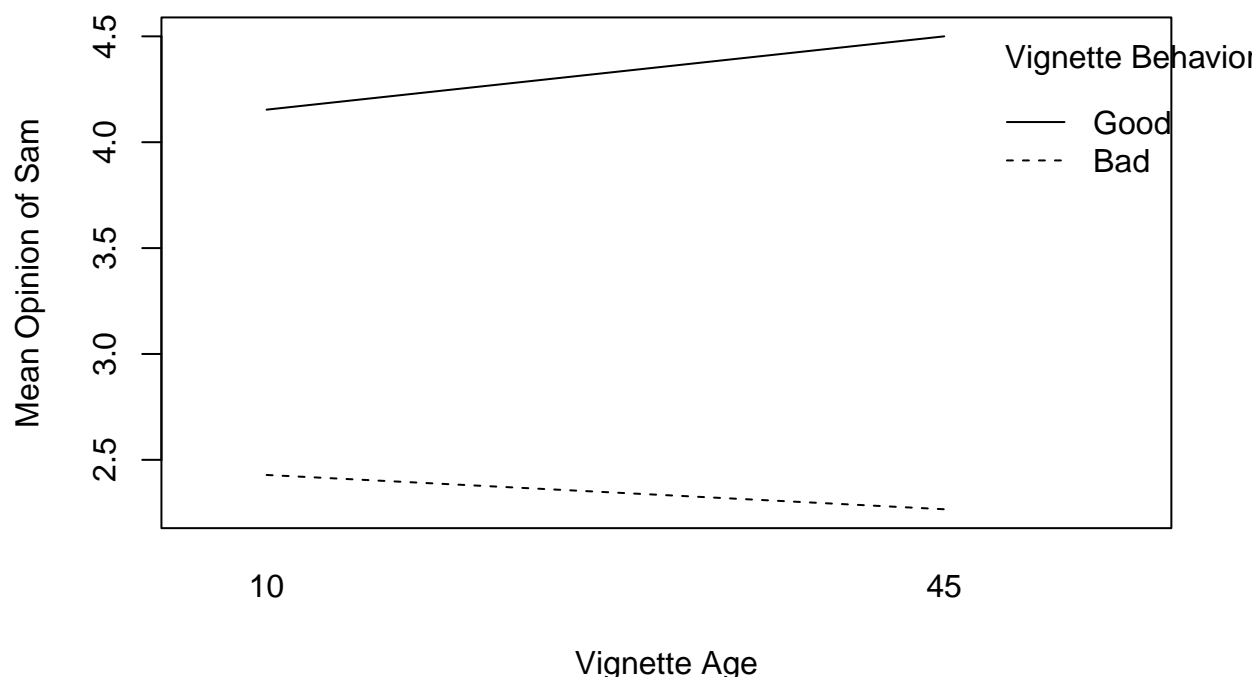
The distributions are still not normal so a reciprocals transformation did not help. Therefore, the condition for normality is violated. However, we will continue with Anova testing, despite the condition for normality being violated, and then check the robustness of the model with the non-parametric test, the Kruskal Wallis test.

Anova Testing

First, we will assess the data for the appearance of interaction between our dependent variables.

slope= 9.3

Residuals

Comparisons

Mean Opinion of Sam

4.5
4.0
3.5
3.0
2.5

Vignette Behavior

—— Good
- - - - Bad

10
45

Vignette Age

The Tukey plot has a nonzero slope, which indicates that there is evidence of non-additivity. However, the interaction plot shows that the lines for Good and Bad behavior do not intersect but have slopes of opposite signs. We will fit both an additive and interaction model.

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## v_age        1   0.10    0.10   0.112    0.739
## behavior     1  55.28   55.28  60.466 2.55e-10 ***
## Residuals   53  48.46    0.91
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


##               Df Sum Sq Mean Sq F value   Pr(>F)
## v_age          1   0.10    0.10   0.112    0.739
## behavior       1  55.28   55.28  60.450 2.89e-10 ***
## v_age:behavior 1   0.90    0.90   0.985    0.325
## Residuals     52  47.55    0.91
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In both the additive and interactive models, we see that the mean opinions of the vignette character are different across the levels of the behavior variable (Good/Bad). The F-value is large and the p-value is smaller than 0.05, which provides sufficient evidence to reject the null hypothesis. From this, we can conclude, that the behavior described in the vignette influences the opinion of the vignette character of Smith College students. We do not have sufficient evidence to conclude that the described age of the vignette character has any affect on Smithie's mean opinion of Sam. There is also not sufficient evidence that the vignette character age variable interacts with the behavior variable.

8

Kruskal Wallis Test

Given the violation of the normality assumption, we will verify our results with a non-parametric test. Since we have more than two groups in our analysis, we will use a Kruskal-Wallis test. To do so, we will conduct separate tests for each of our explanatory variables.

```
##
##  Kruskal-Wallis rank sum test
##
## data:  opinion by behavior
## Kruskal-Wallis chi-squared = 31.14, df = 1, p-value = 2.401e-08


##
##  Kruskal-Wallis rank sum test
##
## data:  opinion by v_age
## Kruskal-Wallis chi-squared = 0.08652, df = 1, p-value = 0.7686
```

We see that the behavior variable once again is significant while the vignette age variable is not. Our results from the ANOVA are confirmed using non-parametric methods.

Pairwise Comparisons

To test our pairwise comparisons, we will use both a TukeyHSD test and a Bonferonni Adjustment. The Tukey test will conduct pairwise comparisons for both of our explanatory variables, while the Bonferroni test will conduct pairwise comparisons for each of the four groups using the condition variable, which indicates which combination of age and behavior was described in the vignette for each respondent.

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = opinion ~ v_age + behavior, data = df)
##
## $v_age
##            diff        lwr       upr     p adj
## 45-10 0.08556833 -0.427319 0.5984556 0.7392237
##
## $behavior
##             diff       lwr      upr p adj
## Good-Bad 1.988396 1.475509 2.501284     0


##
## 95% Bonferroni confidence intervals
##
##                        Diff    Lower    Upper  Decision Adj. p-value
## mu10 Bad-mu10 Good  -1.72527 -2.73559 -0.71496 Reject H0     0.000123
## mu10 Bad-mu45 Bad     0.1619 -0.81286  1.13667    FTR H0            1
## mu10 Good-mu45 Bad   1.88718  0.89321  2.88115 Reject H0        2e-05
## mu10 Bad-mu45 Good  -2.07143 -3.06286    -1.08 Reject H0        3e-06
## mu10 Good-mu45 Good -0.34615 -1.35647  0.66416    FTR H0            1
## mu45 Bad-mu45 Good  -2.23333  -3.2081 -1.25857 Reject H0            0
```

The Tukey HSD test shows that the behavior groups are different from one another, while the age groups are not. This supports our earlier conclusions that behavior is a significant predictor of mean opinion of the vignette character, but not stated character age.

The Bonferroni Adjustment also supports this conclusion. We see that for each of the four significant pairwise comparisons, the behavior is different (i.e. Good vs. Bad comparison). For two of the significant pairwise comparisons, the age of the character varies, but does not for the other two. This also supports the conclusion that the described type of behavior in the vignette influences the mean opinion of Sam.

Post-Modeling Condition Checks

We want to check the residuals to further evaluate our model.

Residuals vs Fitted

Normal Q–Q

Theoretical Quantiles
aov(opinion ~ v_age + behavior)
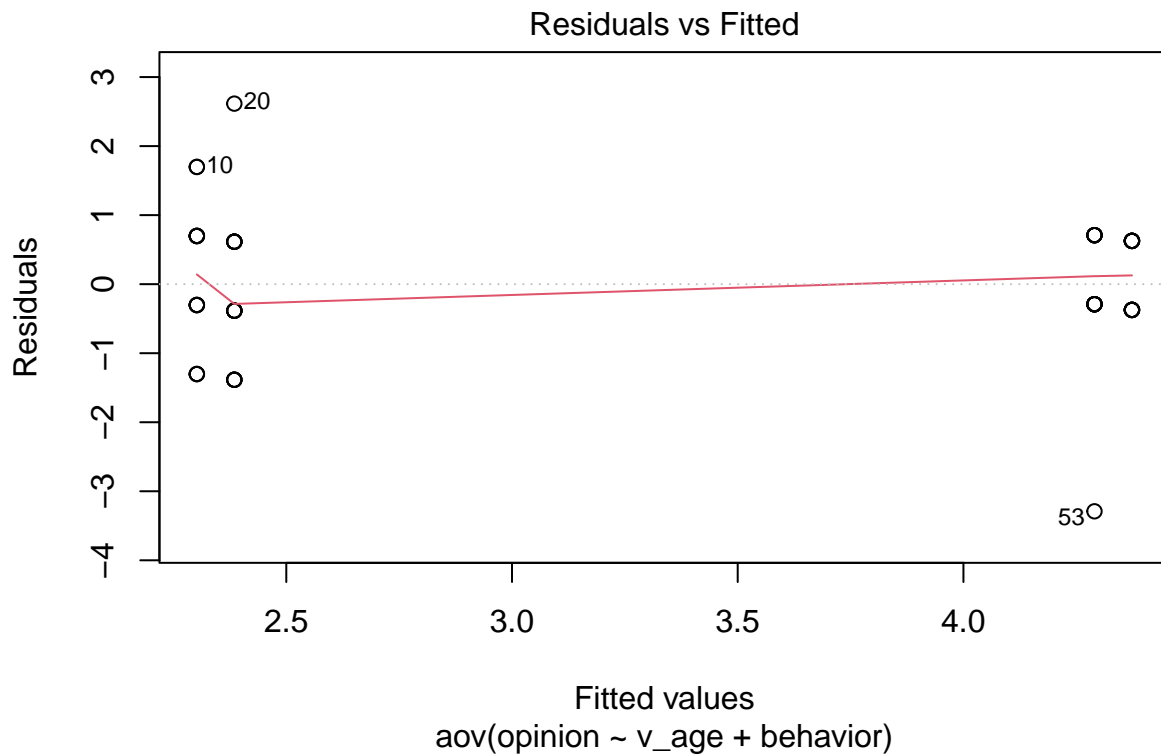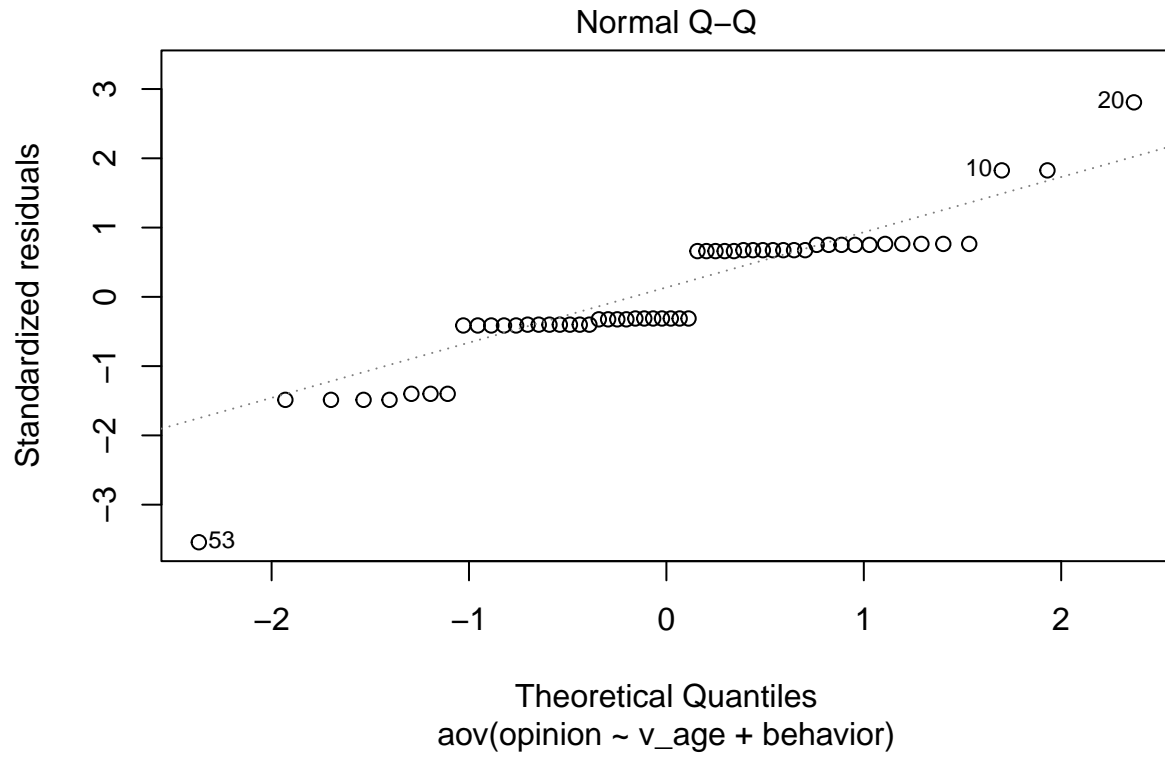
The Fitted vs. Residuals plot shows that the residuals for the additive ANOVA model are reasonably symmetrically distributed. However, the normality condition is very clearly violated as the residuals do not fall along a line. We also see that the residuals are not symmetrical from the density plot.
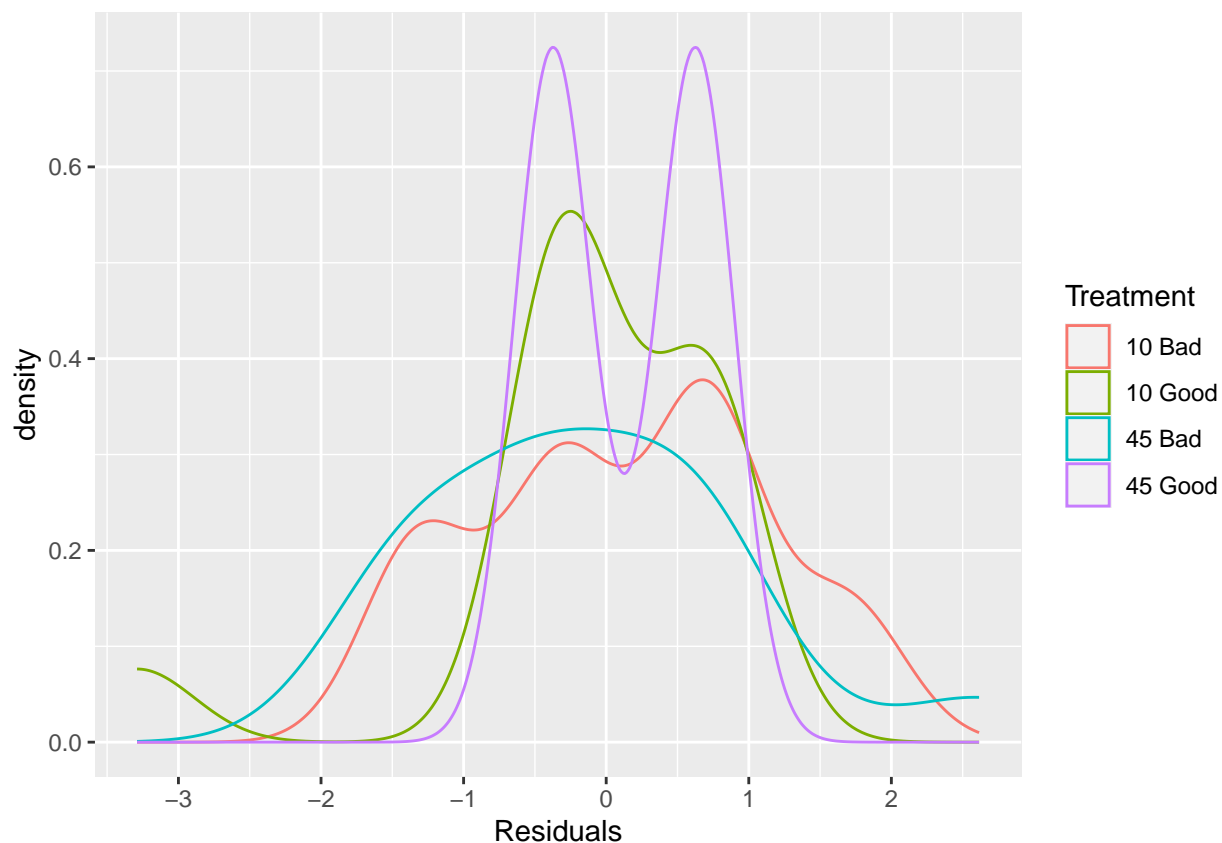
Effect Sizes

```
## $'10 Bad vs. 10 Good'
## [1] -1.895901
##
## $'10 Bad vs. 45 Bad'
## [1] 0.1779121
##
## $'10 Good vs. 45 Bad '
## [1] 2.073824
##
## $'10 Bad vs. 45 Good'
## [1] -2.276297
##
## $'10 Good vs. 45 Good'
## [1] -0.3803846
##
## $'45 Bad vs. 45 Good'
## [1] -2.454209
```

The groups that have large effects are the groups that vary by behavior type. Additionally, some groups with large effects also vary by age, but not all. The groups that do not vary by behavior type do not have large effects. 10 Bad vs. 45 Bad has a trivial effect and 10 Good vs. 45 Good has a small effect.

Conclusion

References

"Goal B: Better Understand the Effects of Personal, Interpersonal, and Societal Factors on Aging, Including the Mechanisms Through Which These Factors Exert Their Effects." n.d. *National Institute on Aging.* Accessed May 1, 2023. https://www.nia.nih.gov/about/aging-strategic-directions-research/goal-behavioral-psychological-factors.