Mini Project 2

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 socially 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 (10 for child VS 45 for adult) and the described behavior (good VS bad) of that character.

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", regardless of Sam's described behavior.

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

a. Respondents who have viewed the vignette featuring Sam exhibiting bad behavior will give lower levels of agreement to the statement "I like Sam", regardless of Sam's age.

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 56 complete, consenting responses. Our participants consisted primarily of white, CIS females between the ages 18-24. The first treatment factor, age, is a categorical variable with two levels (10 as a representative age for children, and 45 as a representative for adults). The second treatment variable, is a categorical behavior with two levels (relatively "good" and "bad"). The character of the vignette is given a gender neutral name, "Sam", and is not addressed by any pronouns in order to avoid gender bias. The dependent variable, participant's opinion of the vignette character, is a continuous numeric variable 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. We also measure another variable, participant's perception of the vignette's behavior as good or bad, which is a continuous numeric variable measured on a five-point Likert scale regarding "Sam's behavior is good". Likewse, 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 based on two experimental factors: 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 age (either 10 or 45 years old) is stated in the vignette and their behavior is described. 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. The four resulting combination of vignettes is therefore: Sam is 10 years old and exhibits good behavior (10 Good); Sam is 45 years old and exhibits bad behavior (45 Bad); Sam is 10 years old and exhibits bad behavior (10 Bad); and Sam is 45 years old and exhibits good behavior (45 Good). Each combination received 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 four vignettes capture the total number of combinations of the two variables. To test our hypotheses, we conducted both additive and interactive Anova models as well as a Kruskal-Wallis Test.

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 (Fig. 1) shows overall similar variability in terms of the length of the 'boxes', however the 'whiskers' are of varying lengths. The distribution for the category in which Sam is 10 years old and conducts good behavior appears to be non-normal since the median is the lower boundary 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. We will also check the standard deviations and means for each treatment group.

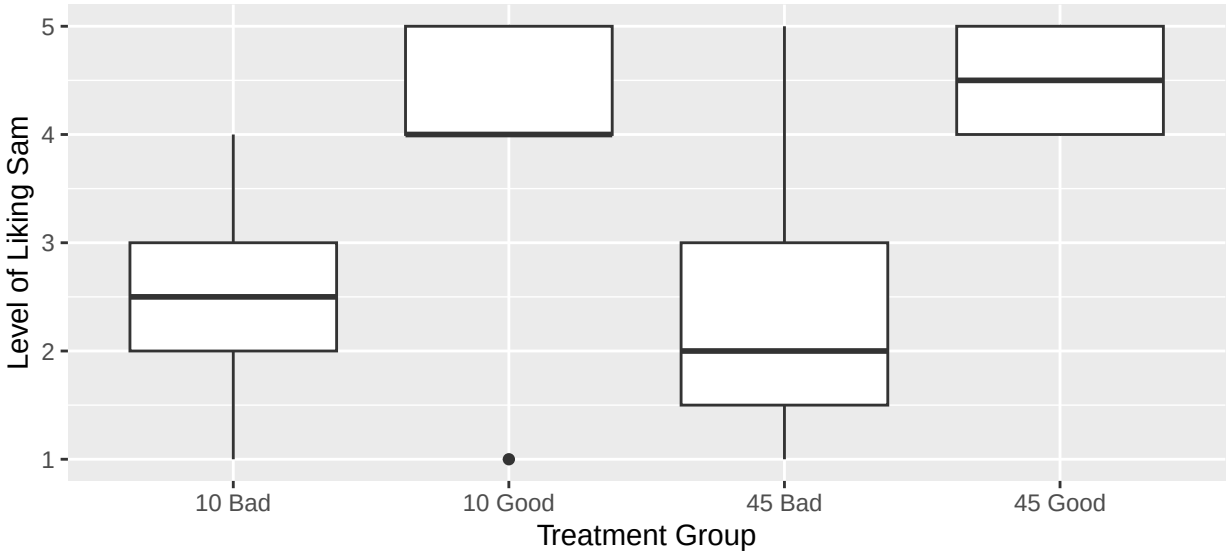

Figure 1: Mean Opinion of Sam by Treatment Group

```
## # 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 descriptive statistics table showing the standard deviations reveal 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 treatment groups (Fig. 2). The "Good" treatment groups (10 Good and 45 Good), both have long left tails and appear display bimodal distributions. The "Bad" treatment groups (10 Bad and 45 Bad) have asymmetrical distributions, and the 10 Bad group is multimodal.

Transformation We now check to see if a transformation of the responses can help satisfy the condition of normality. First, we plot the logs of the mean and standard deviations for the response variable, likability of Sam, across the treatment groups (Fig. 3).

Using the means and standard deviations, we create a linear regression model and look at the slope. 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 normalize the spread of the data.

The distributions are still not normal so a reciprocals transformation did not ressolve the normality issues with our data according to the density plot (Fig.4). Given that the continued issue of normality in the data, we can use a Tukey Non-additivity plot to assess the need for interaction in our model. (Fig. 5)

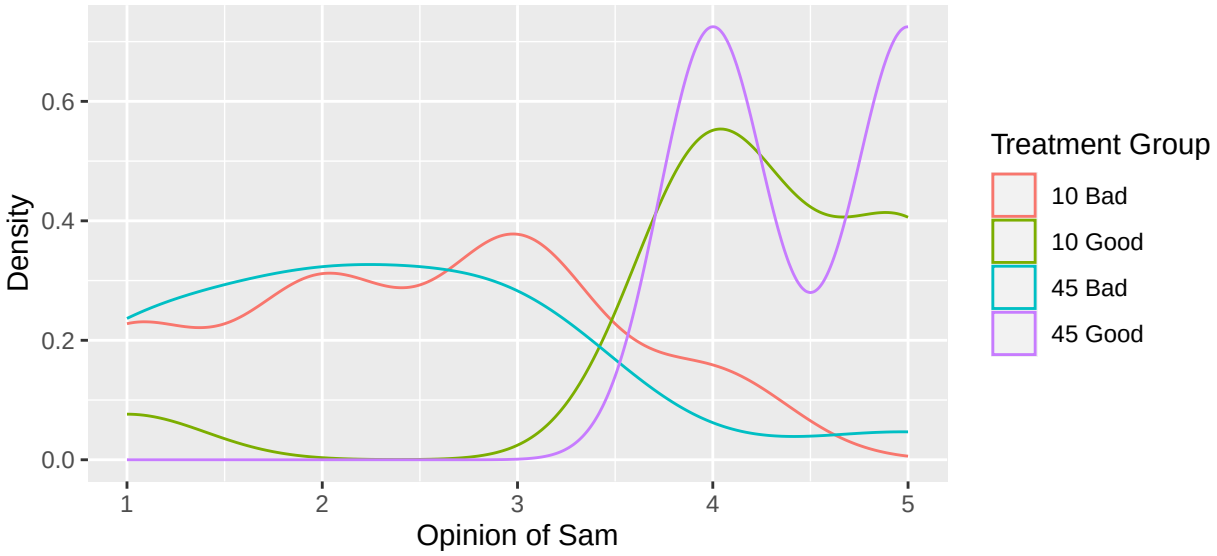

Figure 2: Density Plot for Opinion of Sam

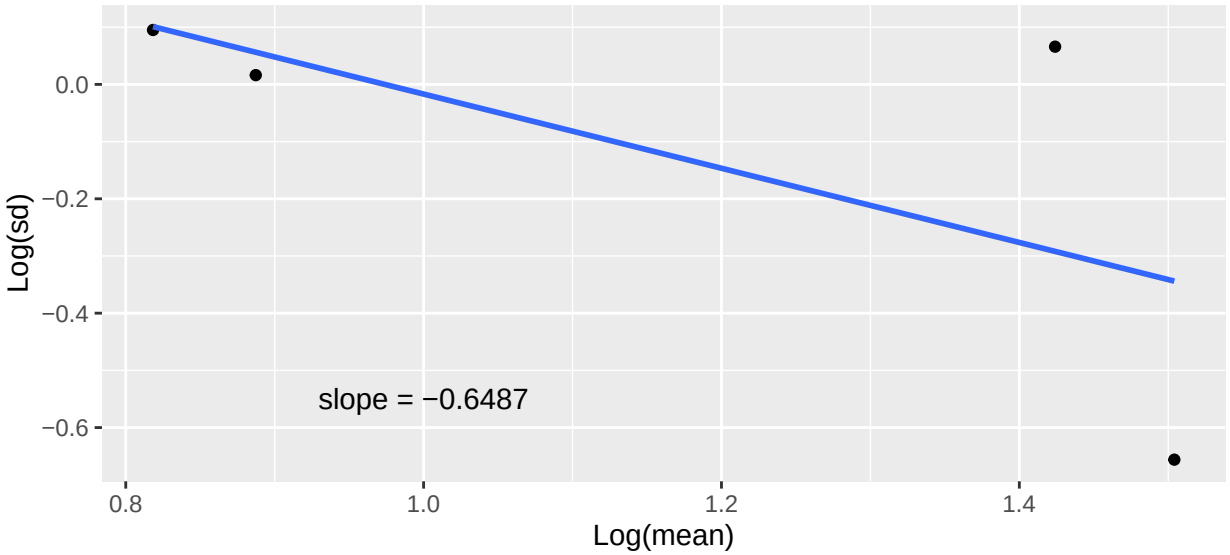

Figure 3: Log(mean) vs. Log(sd) Plot to Assess Possible Data Transformations

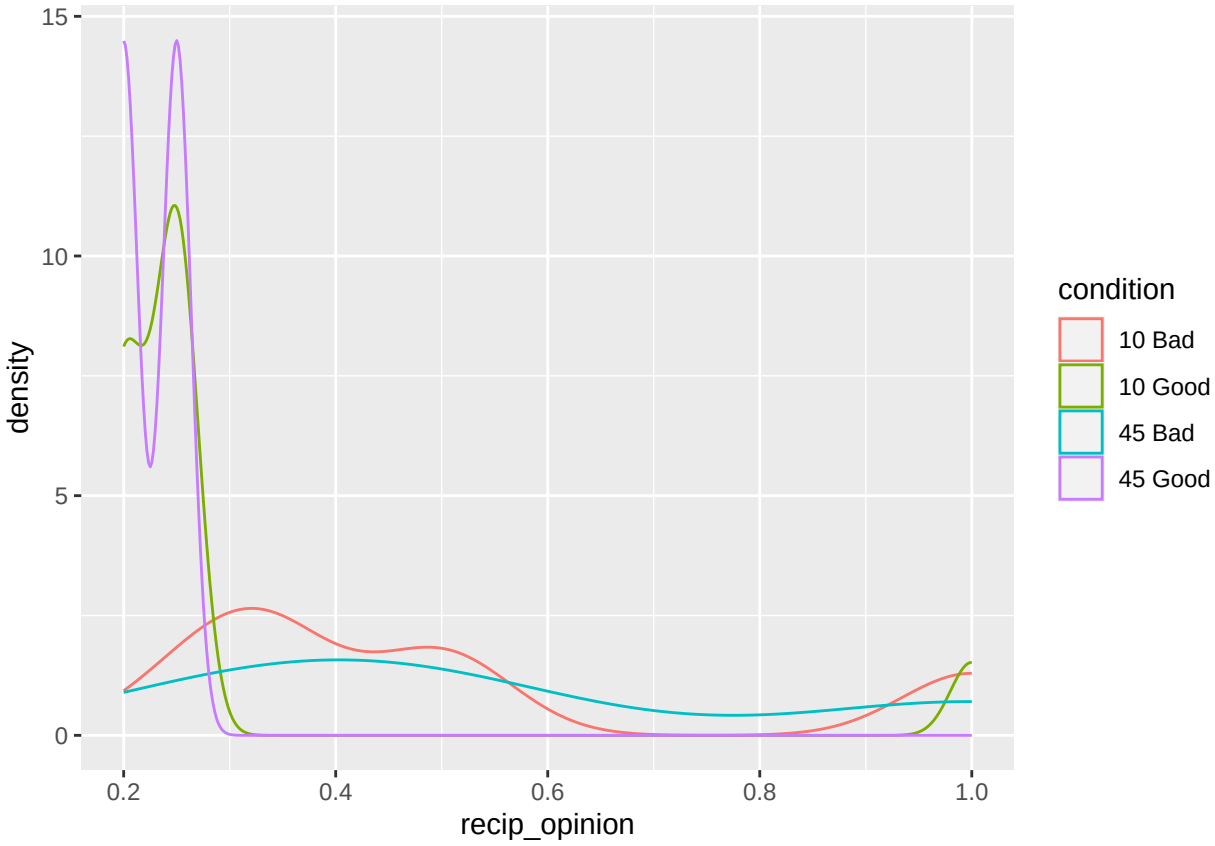

Figure 4: Density Plot of Transformed DV

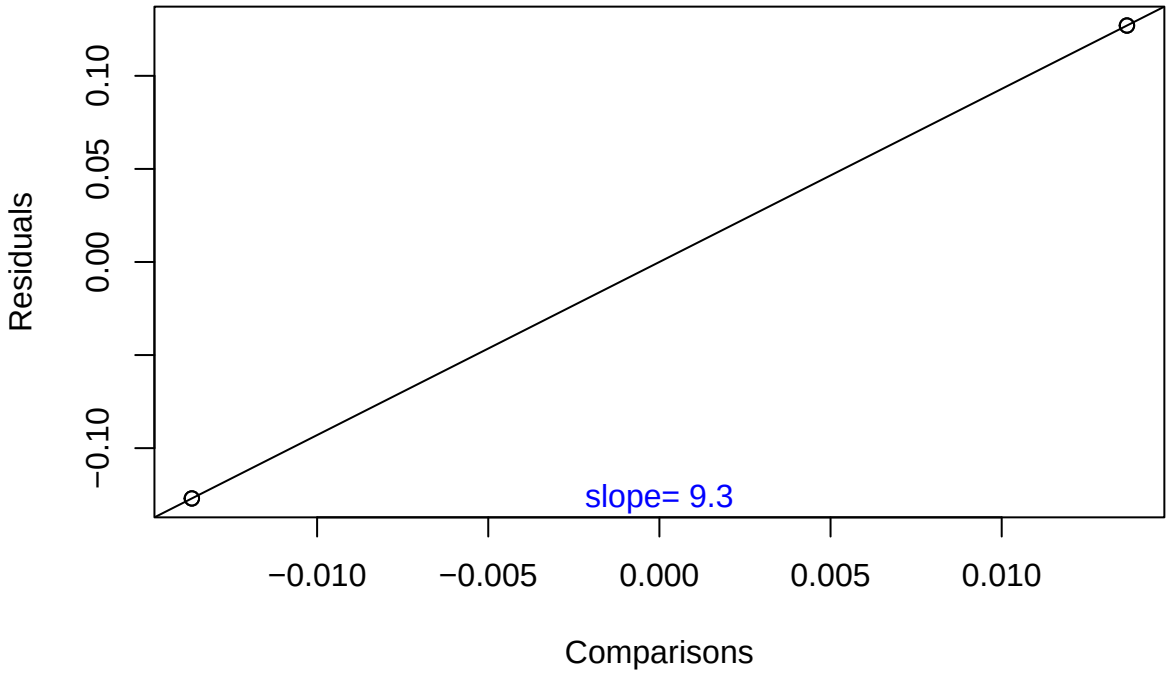

Figure 5: Tukey Non-Addtivity Plot for Opinion by Vignette Age and Behavior

The Tukey Non-additivity plot for our un-transformed dependent variable shows a non-zero slope, which indicates either a transformation of the data or interaction. The Tukey non-additivity plot for the transformed data also has a non-zero slope that indicates the need for interaction in the model. (Fig. 6)

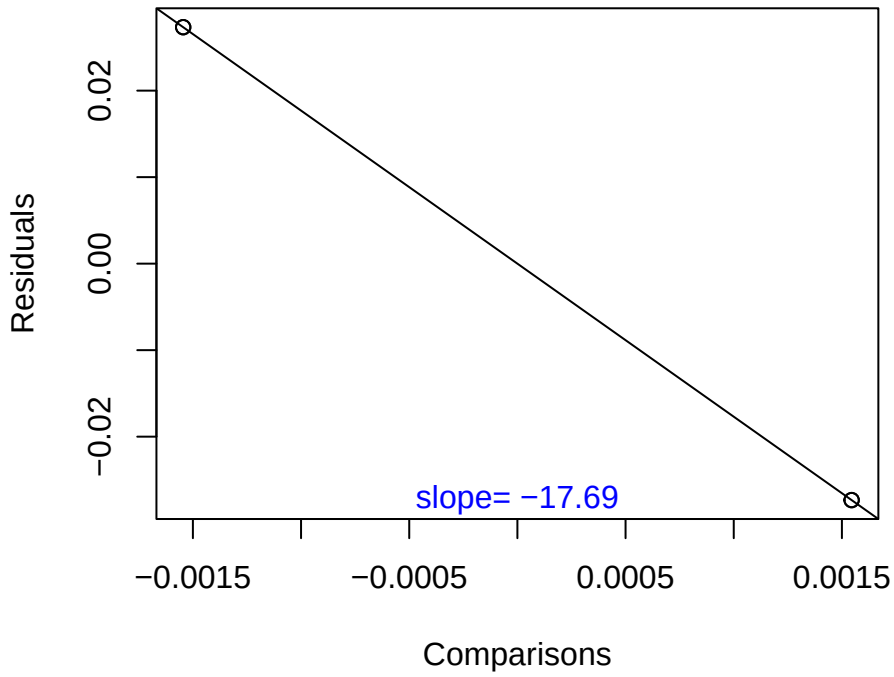

Figure 6: Tukey Non-Addtivity Plot for Reciprical Opinion by Vignette Age and Behavior

The condition for normality is clearly violated for both the transformed and un=transformed data. 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 using an interaction plot (Fig. 7).

The above Tukey plot has a nonzero slope (Fig. 5), which indicates that there is evidence of non-additivity. However, the interaction plot shows that the lines for Good and Bad behavior are reasonably parallel, but do have slopes of opposite signs. This is mixed evidence for the presence of interaction in our data.

We will fit both an additive and interaction model to accommodate this uncertainty and allow us to test our third hypotheses.

```
##             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
```

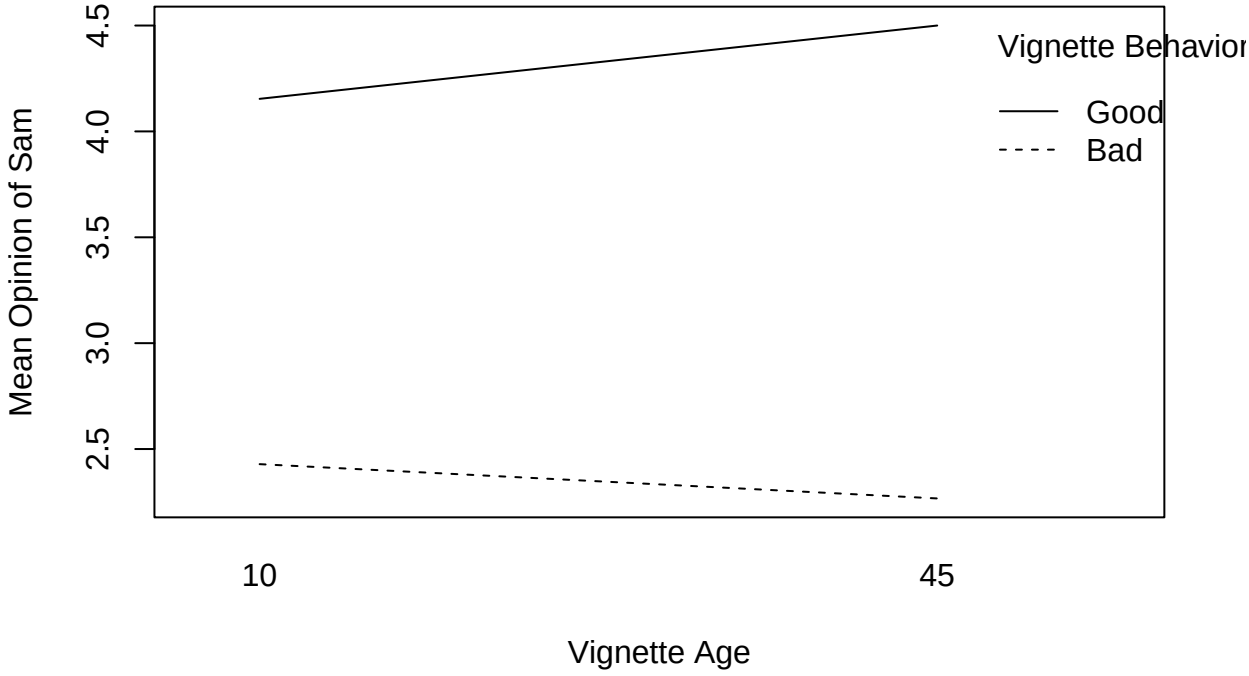

Figure 7: Interaction Plot for Mean Opinion of Sam by Vignette Age and Behavior

```
## ---
## 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 respondents' mean opinion of the vignette character. We do not have sufficient evidence to conclude that the described age of the vignette character has any affect on the respondents' mean opinion of Sam. There is also not sufficient evidence that the age factor of the vignette character interacts with the behavior factor on participant opinions.

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, as well as a third test for the interaction between the two 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
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  opinion by interaction(v_age, behavior)
## Kruskal-Wallis chi-squared = 31.636, df = 3, p-value = 6.245e-07
```

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. However we do find that the interaction between vignette age and behavior is significant according to the Kruskall-Wallis test. From this we conclude that interaction actually *is* present among our variables, as a non-parametric test is more appropriate method of analysis for our data given the significant violations of the normality condition for all treatment groups.

Pairwise Comparisons

To test our pairwise comparisons, we will use both a TukeyHSD test to compare both the levels of the explanatory vairables in isolation and all of the treatment group pairs.

```
##   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 -0.4276201 0.5987568 0.7392836
##
## $behavior
##              diff      lwr      upr p adj
## Good-Bad 1.988396 1.475208 2.501585     0
##
## $`v_age:behavior`
##                         diff        lwr       upr     p adj
## 45:Bad-10:Bad   -0.1619048 -1.1050949 0.7812854 0.9682425
## 10:Good-10:Bad   1.7252747  0.7476862 2.7028633 0.0001185
## 45:Good-10:Bad   2.0714286  1.1121143 3.0307428 0.0000030
## 10:Good-45:Bad   1.8871795  0.9254086 2.8489504 0.0000193
## 45:Good-45:Bad   2.2333333  1.2901432 3.1765235 0.0000004
## 45:Good-10:Good  0.3461538 -0.6314347 1.3237424 0.7836331
```

The Tukey HSD test shows that opinions of 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. Pairwise comparisons of the treatment groups show that the comparisons when behavior differs (Good vs. Bad) have a meaningful difference in means, where those that only vary based on vignette age do not.

The treatment group pairwise comparison results do not support our conclusion from the Kruskall-Wallis test that there is interaction in the data. However,

Post-Modeling Condition Checks

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

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. (Fig. 8)

The density plot shows more evidence of non-normal distribution in the residuals. (Fig. 9)

Effect Sizes and R Squared

```
## $`Good vs. Bad`
## [1] 2.185051
##
## $`10 vs. 45`
## [1] 0.09403168
##
```

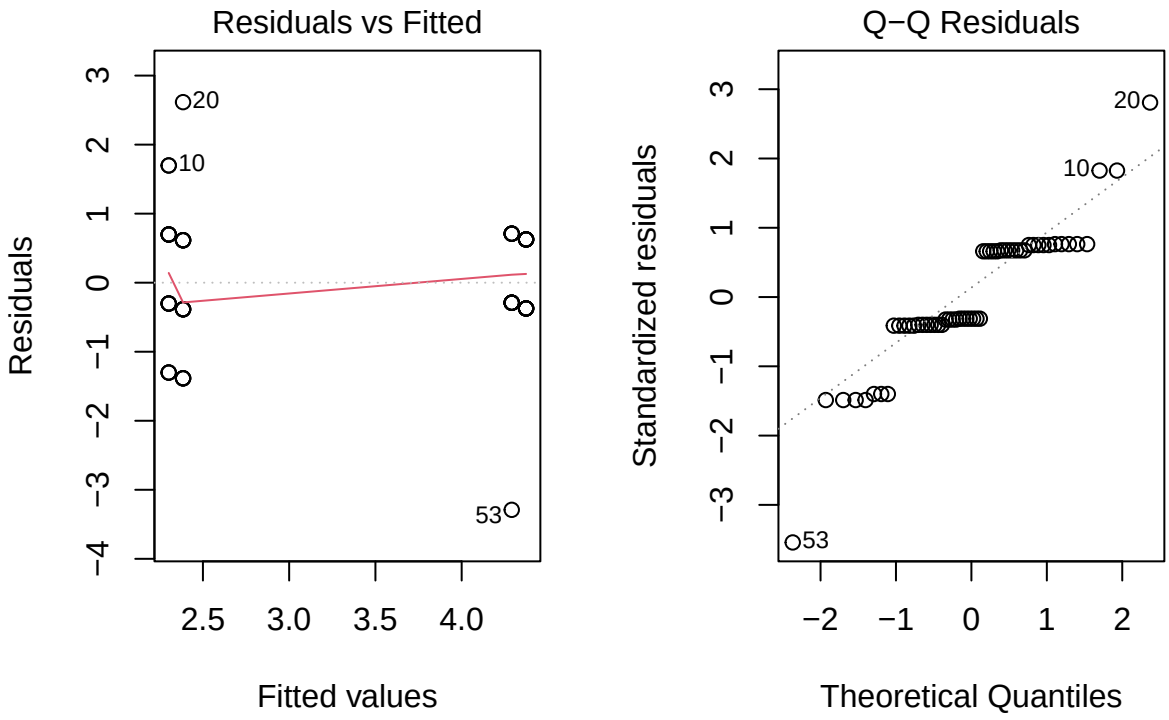

Figure 8: Fitted vs. Residuals Plot and QQ Plot for the Addtive ANOVA Model

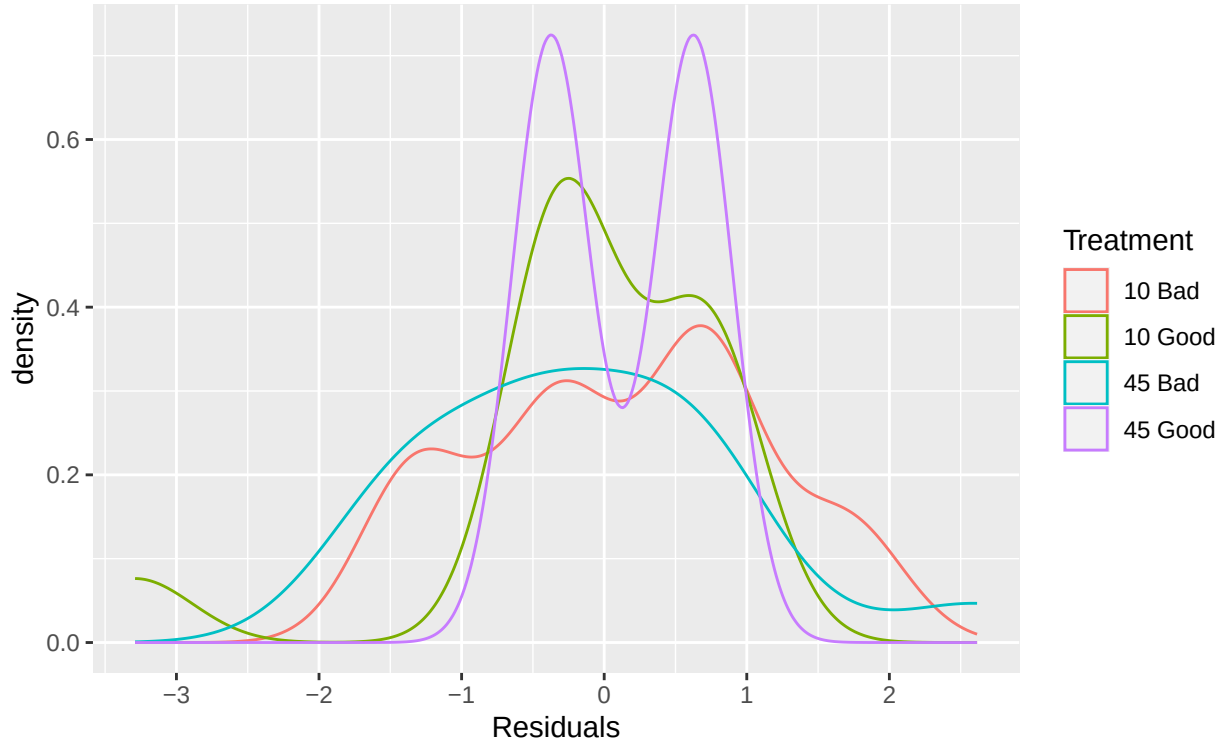

Figure 9: Desnity Plot of Model Residuals

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

Effect sizes show that changing the type of behavior in the vignette has a large effect on the mean opinion of Sam, where the vignette age only has a trivbial effect. Pairwise effect size analysis shows that 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.

```
## $`R squared: Additive`
## [1] 0.5332691
```

```
##
## $`R squared: Interaction`
## [1] 0.5420399
```

The R squared values show that the additive model explains about 53% of the variation in our data, while the interactive model explains about 54% of the variation in the model. The small difference in the R squared values reinforces the conclusion that the character age and behavior in the vignettes do not interactively influence the opinion of Sam in the parametric model.

Conclusion

This study investigated how individual's opinions toward others differed depending on the other person's age and type of behavior. There is not sufficient evidence to suggest that the participants' opinion of the vignette character differed depending on the vignette character's age, so we failed to prove the first hypothesis suggesting that respondents will exhibit more positive attitudes towards the vignette character of younger age. However, there are significant differences in the participants' opinions depending on the type of behavior the vignette character displayed. In concurrence with our second hypothesis, participants were more likely to exhibit less positive attitudes towards the vignette character who displays negative behavior, and more positive attitudes towards the vignette character with positive behavior. Based on our interactive Anova model, we failed to discover any moderating effect of age on the attitudes of respondents across different behavior types, so there is not sufficient evidence to prove our third hypothesis. However, this conclusion is contradicted by the results of the Kruskall-Wallis test on the interaction between the vignette character's age and behavior.

There were several limitations to this study. First, the normality condition is violated in our Anova models, so the significant results indicating that a person's opinion of another differed depending on the other person's behavior may not be reliable. However, our Kruskal Wallis Test results later yielded significant results indicating a difference in opinion across varying behavior types and for interaction. Given the significant normality violations, we are believe that the results of the Kruskall-Wallis tests are more accurate to the true nature of the data, as parametric analysis is not appropriate.

Secondly, our participants were not chosen through random sampling and do not represent the entire Smith student population. Therefore, the results of this study cannot be generalized to the entire Smith student population. Additionally, a greater sample size can possibly serve as a solution to our normality violation.

Third, most of our respondents are CIS white females. As studies have shown, women tend to be more empathetic than men (Hoffman 1977), so this may influence our results.

Future research can use other questions to measure participant opinion towards the vignette character, using statements such as such as "I am willing to be friends with Sam". In our study, we chose a gender-neutral name for our vignette character. In future studies, researchers can investigate how a vignette character's gender may influence participant attitudes towards their behavior, especially when such behavior is related to traditional gender norms, as cooking and cleaning is often considered a female responsibility. Other areas of exploration can include changing the vignette character's racial identity. Future researchers can also investigate if age had a moderating effect on respondents attitudes toward varying degrees of negative behavior, especially ones that may put others in danger. For example, we may compare milder scenarios where both younger and older Sam simply made a mess in the kitchen with more extreme scenarios such as both versions of Sam causes an accidental fire in the kitchen. This way, we can investigate how perceptions of responsibility, tolerance, and forgiveness are related to age and degrees of negative behaviors.

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.

Hoffman, M. L. 1977. "Sex Differences in Empathy and Related Behaviors." *Psychological Bulletin*. https://doi.org/10.1037/0033-2909.84.4.712.