

# 2024 USRA Data Analysis

Benjamin Zubaly

2024-07-18

## Table of contents

<b>Introduction</b>	<b>2</b>
Hypotheses . . . . .	3
<b>Data Cleaning</b>	<b>4</b>
Variables in the Analysis . . . . .	4
<b>Internal Reliability</b>	<b>9</b>
Traditionalism . . . . .	9
Religiosity . . . . .	12
Pathogen Disgust . . . . .	13
Sociosexual Attitudes . . . . .	14
Preferences Subscale of SFGENE-7 . . . . .	15
<b>Descriptive Statistics</b>	<b>16</b>
Sample Size and Time Taken . . . . .	16
Categorical Variables . . . . .	17
Numeric Variables . . . . .	20
<b>Data Visualization</b>	<b>23</b>
Categorical Variables . . . . .	23
Numeric Variables . . . . .	27
<b>Testing Hypotheses</b>	<b>32</b>
Data Preparation . . . . .	33
Testing the Baseline Model . . . . .	35
Assumptions . . . . .	35
Summarizing the Model . . . . .	37
Hypothesis 1: Adherence to Traditional Practices . . . . .	39
Preliminary Simple Linear Regression Models . . . . .	39

Full Mediation Model . . . . .	47
Hypothesis 2: Out-Group Avoidance . . . . .	52
Preliminary Simple Linear Regression Models . . . . .	52
Full Mediation Model . . . . .	60
Hypothesis 3: Sexual Strategies . . . . .	65
Preliminary Simple Linear Regression Models . . . . .	65
Full Mediation Model . . . . .	73
<b>Exploratory Analyses</b>	<b>78</b>
Hypothesis 1: Controlling for Sex . . . . .	78
Preliminary Simple Linear Regression Models . . . . .	78
Full Mediation Model . . . . .	87
Hypothesis 2: Controlling for Sex . . . . .	91
Preliminary Simple Linear Regression Models . . . . .	91
Full Mediation Model . . . . .	99
Hypothesis 3: Controlling for Sex . . . . .	104
Preliminary Simple Linear Regression Models . . . . .	104
Full Mediation Model . . . . .	111
Hypothesis 1: Religious Importance . . . . .	115
Preliminary Simple Linear Regression Models . . . . .	115
Full Mediation Model . . . . .	124
Hypothesis 2: Religious Importance . . . . .	128
Preliminary Simple Linear Regression Models . . . . .	128
Full Mediation Model . . . . .	136
Hypothesis 3: Religious Importance . . . . .	141
Preliminary Simple Linear Regression Models . . . . .	141
Full Mediation Model . . . . .	149
Hypothesis 2: Removing GENE Preferences Item 1 . . . . .	154
Preliminary Simple Linear Regression Models . . . . .	154
Full Mediation Model . . . . .	163
<b>Saving Data</b>	<b>168</b>
<b>References</b>	<b>168</b>

## Introduction

The following tracks the data analysis for our 2024 Undergraduate Student Research Award (USRA) project, “What accounts for the relationship between disgust and religiosity?” This is the master analysis document, so all analyses are documented here—with any other code called within a chunk below. The goals of this analysis are to (1) clean the data, (2) check for internal reliability of our measures, (3) generate descriptive statistics, (4) visualize the

distribution of our variables, (5) test our hypotheses, and (6) conduct exploratory analyses as appropriate.

## Hypotheses

Individual differences in disgust sensitivity have repeatedly been shown to relate to religiosity, as predicted by the behavioral immune system model of disgust, leading to some scholars suggesting that religiosity is an evolved disease avoidance strategy. However, there are multiple mechanisms that aim to account for how religion contributes to disease avoidance, and these different mechanisms provide predictions for the psychological mediators of the relationship between disgust and religiosity. This study aims to test three such mediators, to better differentiate and understand why disgust sensitivity relates to religiosity. These mediators are represented by the following hypotheses.

1. The relationship between disgust sensitivity and religiosity is motivated by adherence to traditional practices.
  - Strong version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be fully mediated by conventionalism.
  - Weak version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be partially mediated by conventionalism.
2. The relationship between disgust sensitivity and religiosity is motivated by out-group avoidance.
  - Strong version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be fully mediated by ethnocentrism.
  - Weak version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be partially mediated by ethnocentrism.
3. The relationship between disgust sensitivity and religiosity is motivated by a monogamous mating strategy.
  - Strong version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be fully mediated by restricted sociosexual attitudes.
  - Weak version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be partially mediated by restricted sociosexual attitudes.

We will test these hypotheses through mediation analysis, the specifics of which will be reviewed in the [Testing Hypotheses](#) section.

## Data Cleaning

All data files are saved in `.csv` format in the `data` directory, and the steps that were taken to process these files are documented in `data/data_preparation_notes.txt`. To clean the data for analysis, we will call a pre-written script (`code/data_cleaning.R`) that returns a data frame called `data`. The script also saves this data frame as `data/data_clean.csv`.

```
# Call the data cleaning script
source("../code/data_cleaning.R")
```

Attached: 'Groundhog' (Version: 3.2.0)

Tips and troubleshooting: <https://groundhogR.com>

**Note:** The Three Domains of Disgust Scale (TDDS) is generally administered with seven response options (Olatunji et al., 2012), only anchored with 0 (Not disgusting at all) and 6 (Extremely disgusting). Due to a survey coding error, our sample completed the TDDS with only six response options and the same anchors. Because of this, each TDDS item has a range from 0 to 5 (rather than 6), and raw scores for the full scale and subscales will be underestimated accordingly.

## Variables in the Analysis

The `data` data frame contains the following variables (and `column names`).

- Participant ID (`ID`):
  - A unique randomly generated 10 digit string of numbers that identifies each participant.
- Start Time (`start_time`) and Finish Time (`finish_time`):
  - The date and time (down to the minute in PST) the participant started and finished the survey in POSIXct format (YYYY-MM-DD HH:MM:SS).
- Age (`age`):
  - The age of the participant in years.

- Sex (**sex**):
  - The sex of the participant as a factor variable (“Male” or “Female”).
- Gender (**gender**):
  - The gender of the participant. Responses were standardized into six categories formulated based on the genders reported. To see which original responses were put into these bins, see the appropriate section of the `code/data_cleaning.R` script. The categories are as follows:
    1. Male
    2. Female
    3. Non-binary
    4. Gender Queer
    5. Gender Fluid
    6. Agender
- Ethnicity (**ethnicity**):
  - The ethnicity of the participant. Participants were provided with a list of potential ethnicities to choose from as well as the options of writing in their own or “Rather Not Say”. Those who reported multiple ethnicities were categorized as “Mixed”. There are 11 unique ethnicities for this sample:
    1. South Asian
    2. East Asian
    3. South East Asian
    4. Latino or Hispanic
    5. White or Caucasian
    6. White or Sapharic Jew
    7. African
    8. Black or African American
    9. Black British
    10. Middle Eastern
    11. Mixed
- Nationality (**nationality**):
  - The nationality of the participant. Participants were provided with a list of potential ethnicities to choose from as well as the option of writing in their own. There were 41 unique nationalities reported by participants.
- Religious Affiliation (**religious\_affiliation**):

- The religion of participants. Participants were provided with a list of potential religions to choose from as well as the option of writing in their own. There were 12 religions reported, including “None”. Some participants wrote in “Catholic” or “Roman Catholic”, and were categorized as Christian (their Christian Affiliation was imputed as “Catholic”).
- Christian Affiliation (`christian_affiliation`):
  - The Christian sect or denomination reported by participants. Participants were provided with a list of potential groups to choose from as well as the option of writing in their own. Including “None” and “Non-religious Christian”, there were 16 affiliations reported.
- Religious Importance (`religious_importance`):
  - Participants were instructed to “Use the slider to indicate how important religion is in your life from 0 (not important at all) to 100 (extremely important).” This single-item measure will be used as a secondary indicator of religiosity, to examine whether results generalize beyond our primary DV of the Centrality of Religiosity Scale.
- The Centrality of Religiosity Scale (`CRS`):
  - The Centrality of Religiosity Scale (CRS-5) is a short measure of religiosity intended to triangulate the measurement of the multidimensional concept of religiosity by getting at five dimensions: intellect, ideology, public practice, private practice, and experience (Huber & Huber, 2012). Recent updates ensure that this measure is suitable for a cross-cultural sample. This measure shows strong concurrent validity with other religiosity measures and strong internal reliability.
  - Because the responses to items of the CRS-5 are qualitatively distinct, there are different response options, but the response format is the same, with a range of options that are scored from 1-5 (see the specific response option below their respective items). Where it is possible to get objective quantities for frequency (i.e., for religious service and prayer), the response options are the same. There are seven response options for these items, and they are collapsed into five scoring options (see below). Where objective quantities are potentially less frequent and are more difficult to capture (i.e., thinking about religious issues), the response options range from never to very often. Where objective quantities are not possible (i.e., degree of belief in God or the divine), response options are captured on a scale from not at all to very much so. Once responses are scored according to the 1-5 system, all items are summed and divided by the number of items (5) to provide a total score ranging from 1.0-5.0.
- The Three Domains of Disgust Scale (full, `TDDS_f`; pathogen, `TDDS_p`; sexual, `TDDS_s`, moral, `TDDS_m`):

- The TDDS was developed by Tybur et al. (2009) in response to a lack of theoretical and empirical grounding for prior measures of disgust sensitivity. Based on evolutionary theory, they predicted that items which people find disgusting should form three factors, one for pathogen, sexual, and moral disgust. Across multiple studies, they demonstrate that a three-factor solution is most parsimonious and fits the data well, that the full scale has good concurrent validity, and that the subscales have good convergent and divergent validity. Subsequent investigations have confirmed the three factor structure of the scale, although they find poorer validity evidence for the moral disgust subscale (Olatunji et al., 2012). All three subscales will be used in exploratory analyses and bivariate correlations, but the pathogen disgust scale will be used to test our central three hypotheses and six specific statistical hypotheses.
- Scores for the full scale and each subscale were calculated by summing item values.
- The Revised Sociosexual Orientation Inventory (full, `SOI_f`; behavior, `SOI_b`; attitudes, `SOI_a`; desire, `SOI_d`):
  - The original Sociosexual Orientation Inventory (SOI) was a 7-item global measure of sociosexual orientation (Simpson & Gangestad, 1991), but this measure included behavioral, attitudinal, and (one) desire items together, with different response formats. This is problematic because each of these three components are differentiable both conceptually and empirically, and the different response formats often lead to an improper weighting of each of the three components contributing to the total score, resulting in poor reliability. Penke & Asendorpf (2008) developed the SOI-R to address these, and other, problems. The SOI-R is a 9-item measure with three items each tailored towards behaviors, attitudes, and desires, which they differentiated based on exploratory and confirmatory factor analysis as well as convergent and divergent validation. Each of the subscales show adequate reliability, along with the global score. As we are interested in operationalizing a monogamous mating strategy, it is participants' *intentions* that we were most interested in measuring. In this case, one's desire and behavior may, or may not, correspond to their intentions, which would best be operationalized as sociosexual attitudes. Therefore, we will use the attitudes subscale of the SOI-R to operationalize a monogamous mating strategy. This will be used to test our mediational hypothesis involving monogamous sexual strategies.
  - Scores for the full scale and each subscale were calculated by taking the arithmetic mean of item values.
- The In-Group Preference Subscale of the Short Form of the Generalized Ethnocentrism Scale (SFGENE-7; full, `GENE_f`; preference, `GENE_p`; superiority, `GENE_s`):
  - The SFGENE-7 is simply a short form of the Generalized Ethnocentrism Scale (Neuliep, 2002; Neuliep & McCroskey, 1997). The GENE is composed of 22 items,

15 of which are scored, and out of all the measures that we have, the GENE would be by far the longest. By reducing the number of items to 7 with the SFGENE-7, we can allow for the recruitment of more participants. The SFGENE-7 has shown adequate reliability despite a loss of items, as well as evidence for convergent validity with constructs like tolerance and multicultural ideology. In addition, it has been broken down via factor analysis into two distinct subscales, the first three items comprising the in-group preference scale and the last four items comprising the in-group superiority scale. Each of these subscales also have strong internal reliabilities, and the in-group preference scale shows stronger negative correlations with tolerance and multicultural identity than the in-group superiority scale. We will use only the first three items to measure the in-group preference aspect of ethnocentrism in order to operationalize out-group avoidance, particularly because the items comprising the in-group superiority subscale are quite similar at face value to items in the conventionalism scale that we are using to measure the tendency towards adherence to tradition. The in-group preference subscale will be used to test the mediational hypothesis involving out-group avoidance.

- Items were summed to provide the total, preference, and superiority scores.
- Conventionalism Subscale of Aggression-Submission-Conventionalism Scale of Authoritarianism (ASC-C; CONV) and additional Traditionalism Items (TRAD; together, CONV\_f):
  - Traditionalism is often operationalized as the traditionalism subscale of the Right Wing Authoritarianism (RWA) scale (Duckitt et al., 2010), but there are problems with this operationalization for our purposes. The RWA scale, and the traditionalism scale in particular, uses explicitly religious, sexual, and, some have argued, prejudicial content. Dunwoody & Funke (2016) address these limitations of the RWA scale by developing the Aggression-Submission-Conventionalism scale of Right Wing Authoritarianism that omits any politically charged, evaluative, prejudicial, or explicitly religious or sexual language. This removes the problems of tautology when predicting, in our case, religiosity, while retaining the same factor structure and similar reliability and validity as previous RWA scales. The conventionalism subscale, in particular, is meant to measure the same thing as the traditionalism subscale of the RWA scale but without religious language. Despite removing religious language, it retains a positive relationship with religiosity, and it seems to have better face validity than the traditionalism subscale for measuring adherence to traditional norms/practices. It is also quite short, allowing for quick administration without sacrificing psychometric properties.
  - In an attempt to improve our measure of traditionalism, we added some items which hopefully will get at our construct more specifically. These items are the TRAD items in the data frame. We will calculate Chronbach’s alpha and only accept these items in the measure if we keep an alpha above .8. If the alpha is below .8, we will remove items until it performs at this level.
  - Items were summed to provide each of the total scores.



## Internal Reliability

To quickly assess the internal reliability of our scales, we will calculate Chronbach's alpha for each of our scales. We will start with the CONV\_f scale, because we may remove items we added to it.

```
# Defining the packages to use to do the alpha analysis
pkg <- c("tidyverse", "psych")
# tidyverse if not attached from data cleaning script (for pipe and select function)
# psych for alpha (and later descriptives)

# Loading the groundhog package to install packages from a certain date
library(groundhog)

# Reading in the packages with the groundhog package for reproducibility
# Message suppressed in order to allow for rendering of quarto document
suppressMessages(groundhog.library(pkg = pkg, date = "2024-06-01"))

# Remove the pkg vector
rm(pkg)
```

## Traditionalism

```
# Chronbach's Alpha for CONV_f (which includes all CONV and TRAD items)
alpha(x = data %>% select(starts_with("CONV."), starts_with("TRAD.")))
```

Reliability analysis

Call: alpha(x = data %>% select(starts\_with("CONV."), starts\_with("TRAD.")))

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	sd	median_r
	0.83	0.83	0.87	0.29	4.9	0.015	3	0.86	0.28

95% confidence boundaries

	lower	alpha	upper
Feldt	0.8	0.83	0.86
Duhachek	0.8	0.83	0.86

Reliability if an item is dropped:

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha	se	var.r	med.r
--	-----------	-----------	---------	-----------	-----	-------	----	-------	-------

CONV.1	0.82	0.82	0.86	0.30	4.6	0.015	0.034	0.29
CONV.2	0.80	0.80	0.84	0.27	4.0	0.017	0.031	0.25
CONV.3	0.82	0.82	0.85	0.29	4.6	0.015	0.029	0.30
CONV.4	0.81	0.81	0.85	0.28	4.4	0.016	0.034	0.28
CONV.5	0.82	0.81	0.85	0.28	4.4	0.016	0.034	0.28
CONV.6	0.82	0.82	0.85	0.29	4.5	0.016	0.028	0.29
TRAD.1	0.84	0.84	0.87	0.33	5.4	0.014	0.023	0.31
TRAD.2	0.82	0.82	0.85	0.29	4.5	0.016	0.030	0.28
TRAD.3	0.82	0.82	0.86	0.29	4.5	0.015	0.034	0.28
TRAD.4	0.82	0.81	0.85	0.28	4.4	0.016	0.031	0.29
TRAD.5	0.81	0.80	0.84	0.27	4.1	0.017	0.028	0.26
TRAD.6	0.81	0.81	0.85	0.28	4.3	0.016	0.030	0.28

Item statistics

	n	raw.r	std.r	r.cor	r.drop	mean	sd
CONV.1	289	0.53	0.53	0.47	0.42	2.3	1.4
CONV.2	289	0.76	0.75	0.74	0.68	3.4	1.5
CONV.3	289	0.54	0.55	0.52	0.43	3.4	1.5
CONV.4	289	0.64	0.63	0.58	0.53	2.7	1.5
CONV.5	289	0.63	0.63	0.59	0.54	2.2	1.5
CONV.6	289	0.57	0.58	0.55	0.47	3.9	1.3
TRAD.1	289	0.27	0.27	0.17	0.14	1.5	1.4
TRAD.2	289	0.57	0.57	0.53	0.47	3.9	1.4
TRAD.3	289	0.57	0.57	0.51	0.46	2.0	1.5
TRAD.4	289	0.63	0.63	0.60	0.53	3.0	1.5
TRAD.5	289	0.73	0.73	0.73	0.66	3.7	1.4
TRAD.6	289	0.63	0.64	0.61	0.54	3.7	1.4

Non missing response frequency for each item

	0	1	2	3	4	5	6	miss
CONV.1	0.13	0.17	0.27	0.25	0.12	0.03	0.02	0
CONV.2	0.04	0.07	0.14	0.28	0.24	0.13	0.10	0
CONV.3	0.04	0.06	0.12	0.29	0.25	0.15	0.09	0
CONV.4	0.07	0.13	0.27	0.26	0.13	0.08	0.06	0
CONV.5	0.13	0.18	0.31	0.21	0.10	0.04	0.03	0
CONV.6	0.01	0.03	0.09	0.25	0.30	0.19	0.12	0
TRAD.1	0.31	0.25	0.22	0.15	0.03	0.02	0.01	0
TRAD.2	0.02	0.04	0.10	0.20	0.31	0.17	0.15	0
TRAD.3	0.19	0.23	0.25	0.16	0.11	0.04	0.01	0
TRAD.4	0.06	0.09	0.23	0.25	0.23	0.09	0.06	0
TRAD.5	0.03	0.05	0.09	0.26	0.32	0.13	0.12	0
TRAD.6	0.02	0.07	0.08	0.25	0.30	0.20	0.08	0

- The raw alpha value (.831; standardized .830) is acceptable here, so we will use our full CONV\_f variable to operationalize traditionalism.
- To compare these values to those for the original conventionalism measure, I will calculate its' alpha as well.

```
# Chronbach's Alpha for CONV (only CONV items)
alpha(x = data %>% select(starts_with("CONV.")))
```

Reliability analysis

Call: alpha(x = data %>% select(starts\_with("CONV.")))

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	sd	median_r
0.74	0.74	0.76	0.33	2.9	0.024	3	0.97	0.38

95% confidence boundaries

	lower	alpha	upper
Feldt	0.7	0.74	0.79
Duhachek	0.7	0.74	0.79

Reliability if an item is dropped:

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha	se	var.r	med.r
CONV.1	0.73	0.73	0.75	0.36	2.8	0.025	0.028	0.40	
CONV.2	0.67	0.67	0.70	0.29	2.1	0.031	0.038	0.21	
CONV.3	0.72	0.72	0.70	0.34	2.6	0.025	0.018	0.39	
CONV.4	0.70	0.70	0.71	0.32	2.4	0.028	0.031	0.36	
CONV.5	0.69	0.69	0.70	0.31	2.3	0.029	0.030	0.35	
CONV.6	0.72	0.71	0.70	0.33	2.5	0.026	0.020	0.39	

Item statistics

	n	raw.r	std.r	r.cor	r.drop	mean	sd
CONV.1	289	0.59	0.59	0.45	0.39	2.3	1.4
CONV.2	289	0.75	0.75	0.67	0.59	3.4	1.5
CONV.3	289	0.62	0.63	0.56	0.43	3.4	1.5
CONV.4	289	0.68	0.67	0.58	0.50	2.7	1.5
CONV.5	289	0.70	0.70	0.62	0.53	2.2	1.5
CONV.6	289	0.63	0.64	0.58	0.45	3.9	1.3

Non missing response frequency for each item

	0	1	2	3	4	5	6	miss
CONV.1	0.13	0.17	0.27	0.25	0.12	0.03	0.02	0

CONV.2	0.04	0.07	0.14	0.28	0.24	0.13	0.10	0
CONV.3	0.04	0.06	0.12	0.29	0.25	0.15	0.09	0
CONV.4	0.07	0.13	0.27	0.26	0.13	0.08	0.06	0
CONV.5	0.13	0.18	0.31	0.21	0.10	0.04	0.03	0
CONV.6	0.01	0.03	0.09	0.25	0.30	0.19	0.12	0

- As you can see, the CONV\_f variable has stronger internal reliability than the CONV items alone.

## Religiosity

```
# Chronbach's Alpha for Centrality of Religiosity Items
alpha(x = data %>% select(starts_with("CRS.")))
```

Reliability analysis

Call: alpha(x = data %>% select(starts\_with("CRS.")))

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	sd	median_r
0.92	0.92	0.91	0.69	11	0.0069	2.6	1.3	0.72

95% confidence boundaries

	lower	alpha	upper
Feldt	0.90	0.92	0.93
Duhachek	0.91	0.92	0.93

Reliability if an item is dropped:

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha	se	var.r	med.r
CRS.1	0.93	0.93	0.92	0.78	13.8	0.0064	0.002	0.80	
CRS.2	0.89	0.89	0.88	0.68	8.4	0.0094	0.012	0.66	
CRS.3	0.90	0.90	0.88	0.69	8.8	0.0090	0.016	0.70	
CRS.4	0.89	0.89	0.86	0.66	7.7	0.0104	0.011	0.65	
CRS.5	0.89	0.89	0.87	0.67	8.1	0.0096	0.015	0.65	

Item statistics

	n	raw.r	std.r	r.cor	r.drop	mean	sd
CRS.1	289	0.73	0.76	0.65	0.63	2.6	1.1
CRS.2	289	0.90	0.89	0.87	0.83	3.0	1.5
CRS.3	289	0.88	0.88	0.84	0.81	2.3	1.4
CRS.4	289	0.93	0.92	0.91	0.87	2.5	1.7
CRS.5	289	0.90	0.90	0.88	0.85	2.5	1.4

Non missing response frequency for each item

	1	2	3	4	5	miss
CRS.1	0.14	0.37	0.28	0.13	0.08	0
CRS.2	0.18	0.29	0.15	0.11	0.27	0
CRS.3	0.42	0.22	0.13	0.09	0.14	0
CRS.4	0.46	0.15	0.08	0.03	0.28	0
CRS.5	0.33	0.26	0.16	0.11	0.14	0

- The CRS shows strong internal consistency here ( $\alpha = .919$ ; standardized = .919).

## Pathogen Disgust

```
# Chronbach's Alpha for Pathogen Disgust Items
```

```
alpha(x = data %>% select(., c("TDDS.12", "TDDS.15", "TDDS.9", "TDDS.3", "TDDS.21", "TDDS.18"
```

Reliability analysis

```
Call: alpha(x = data %>% select(., c("TDDS.12", "TDDS.15", "TDDS.9",
  "TDDS.3", "TDDS.21", "TDDS.18", "TDDS.6")))
```

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	sd	median_r
0.77	0.77	0.76	0.33	3.4	0.021	3	0.91	0.34

95% confidence boundaries

	lower	alpha	upper
Feldt	0.73	0.77	0.81
Duhachek	0.73	0.77	0.81

Reliability if an item is dropped:

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha	se	var.r	med.r
TDDS.12	0.74	0.74	0.72	0.33	2.9	0.023	0.0094	0.34	
TDDS.15	0.72	0.72	0.70	0.30	2.6	0.026	0.0063	0.28	
TDDS.9	0.72	0.73	0.70	0.31	2.7	0.025	0.0064	0.31	
TDDS.3	0.75	0.75	0.74	0.34	3.1	0.023	0.0091	0.34	
TDDS.21	0.75	0.75	0.74	0.33	3.0	0.023	0.0114	0.35	
TDDS.18	0.75	0.75	0.74	0.34	3.1	0.022	0.0095	0.35	
TDDS.6	0.75	0.76	0.74	0.34	3.1	0.023	0.0076	0.34	

Item statistics

	n	raw.r	std.r	r.cor	r.drop	mean	sd
--	---	-------	-------	-------	--------	------	----

TDDS.12	289	0.65	0.65	0.57	0.49	3.4	1.4
TDDS.15	289	0.73	0.74	0.70	0.61	3.3	1.3
TDDS.9	289	0.71	0.71	0.67	0.57	2.5	1.4
TDDS.3	289	0.58	0.61	0.51	0.44	4.0	1.1
TDDS.21	289	0.65	0.63	0.52	0.47	3.0	1.5
TDDS.18	289	0.63	0.61	0.51	0.44	3.1	1.5
TDDS.6	289	0.61	0.60	0.50	0.44	1.9	1.4

Non missing response frequency for each item

	0	1	2	3	4	5	miss
TDDS.12	0.02	0.09	0.15	0.21	0.25	0.28	0
TDDS.15	0.01	0.11	0.15	0.26	0.24	0.24	0
TDDS.9	0.07	0.20	0.28	0.21	0.15	0.10	0
TDDS.3	0.00	0.03	0.09	0.17	0.23	0.48	0
TDDS.21	0.06	0.15	0.18	0.22	0.15	0.24	0
TDDS.18	0.04	0.14	0.18	0.19	0.17	0.27	0
TDDS.6	0.17	0.24	0.26	0.18	0.08	0.06	0

- The TDDS Pathogen Disgust (TDDS\_p) shows adequate internal reliability here (alpha = .769; standardized = .772).

## Sociosexual Attitudes

```
# Chronbach's Alpha for the Attitudes Subscale of the SOI-R
alpha(x = data %>% select(., c("SOI.4", "SOI.5", "SOI.6")))
```

Reliability analysis

Call: alpha(x = data %>% select(., c("SOI.4", "SOI.5", "SOI.6")))

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	sd	median_r
0.81	0.81	0.74	0.58	4.2	0.02	5.2	2.4	0.6

95% confidence boundaries

	lower	alpha	upper
Feldt	0.76	0.81	0.84
Duhachek	0.77	0.81	0.84

Reliability if an item is dropped:

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha	se	var.r	med.r
SOI.4	0.76	0.76	0.62	0.62	3.2	0.028	NA	0.62	

SOI.5	0.69	0.69	0.52	0.52	2.2	0.037	NA	0.52
SOI.6	0.75	0.75	0.60	0.60	3.0	0.029	NA	0.60

#### Item statistics

	n	raw.r	std.r	r.cor	r.drop	mean	sd
SOI.4	289	0.83	0.83	0.70	0.63	5.9	2.7
SOI.5	289	0.88	0.87	0.78	0.70	4.3	2.9
SOI.6	289	0.84	0.84	0.71	0.64	5.3	2.8

#### Non missing response frequency for each item

	1	2	3	4	5	6	7	8	9	miss
SOI.4	0.11	0.04	0.08	0.06	0.11	0.09	0.15	0.11	0.25	0
SOI.5	0.29	0.11	0.10	0.04	0.10	0.09	0.09	0.06	0.13	0
SOI.6	0.17	0.04	0.11	0.05	0.14	0.09	0.08	0.12	0.19	0

- Despite only comprising three items, the attitudes subscale of the SOI-R (SOI\_a) seems to have good internal consistency here ( $\alpha = .806$ ; standardized = .806).

## Preferences Subscale of SFGENE-7

```
# Chronbach's Alpha for the Preferences Subscale of the GENE
alpha(x = data %>% select(., c("GENE.1", "GENE.2", "GENE.3")))
```

#### Reliability analysis

```
Call: alpha(x = data %>% select(., c("GENE.1", "GENE.2", "GENE.3")))
```

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	sd	median_r
0.52	0.53	0.55	0.27	1.1	0.05	2.3	0.68	0.095

#### 95% confidence boundaries

	lower	alpha	upper
Feldt	0.42	0.52	0.61
Duhachek	0.43	0.52	0.62

#### Reliability if an item is dropped:

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha	se	var.r	med.r
GENE.1	0.801	0.804	0.672	0.672	4.09	0.023	NA	0.672	
GENE.2	0.093	0.093	0.049	0.049	0.10	0.107	NA	0.049	
GENE.3	0.174	0.174	0.095	0.095	0.21	0.097	NA	0.095	

```

Item statistics
      n raw.r std.r r.cor r.drop mean  sd
GENE.1 289  0.54  0.53 0.096  0.078  3.2 0.95
GENE.2 289  0.81  0.82 0.766  0.536  1.7 0.89
GENE.3 289  0.80  0.80 0.734  0.472  1.9 0.99

```

```

Non missing response frequency for each item
      1      2      3      4      5 miss
GENE.1 0.05 0.12 0.47 0.27 0.09      0
GENE.2 0.51 0.32 0.12 0.03 0.01      0
GENE.3 0.47 0.29 0.18 0.06 0.01      0

```

- The internal consistency of the preferences subscale of the SFGENE-7 is quite low ( $\alpha = .523$ ; standardized = .528). However, the removal of Item 1 improves the alpha statistic drastically (to  $\alpha = .801$ ; standardized = .804). Because of this, we will run the analysis with both the full Preferences Subscale and only Items 2 and 3. I will calculate the scores for the latter variable and call it `GENE_p_2`.

```

# Creating the GENE_p_2 variable as the sum of items 2 and 3
data <- data %>%
  mutate(
    GENE_p_2 = rowSums(select(., c(GENE.2, GENE.3)))
  )

```

## Descriptive Statistics

### Sample Size and Time Taken

First, we will see how many participants we have and how long it took participants to complete the survey.

```

# Calculate the number of participants in the sample
nrow(data)

```

```
[1] 289
```

```

# Calculate descriptive statistics for the difference between the finish and start time (in minutes)
describe(as.numeric(difftime(data$finish_time, data$start_time, units = "mins")))

```

```

      vars    n  mean   sd median trimmed  mad min max range  skew kurtosis   se
X1      1 289 14.39 58.1      8    8.94 2.97   3 821   818 12.25  154.08 3.42

```



- In total, we have  $N = 289$  participants with complete data.
- Although the mean time taken is around 14.39, the SD is huge, and the maximum is a massive 821 minutes (over 13 hours). The trimmed mean of 8.94 is therefore a better estimate.

## Categorical Variables

Next, we will describe the sample in terms of the categorical variables. The following generates frequency tables for all categorical variables.

```
# Create a frequency table for each factor variable
lapply(data[sapply(data, is.factor)], table)
```

\$sex

Male	Female
146	143

\$gender

Agender	Female	Gender Fluid	Gender Queer	Male	Non-binary
2	135	1	1	144	4

\$ethnicity

African	Black or African American
52	4
Caribbean	East Asian
0	14
Latino or Hispanic	Middle Eastern
62	4
Mixed Native American or Alaskan Native	
11	0
South Asian	White or Causasian
8	122
White or Sapharic Jew	Black British
6	1
White Mexican	Romani or Traveller
0	0
South East Asian	Rather Not Say
2	3

# \$nationality

Algeria	Australia	Austria	Belgium	Brazil
1	7	1	1	3
Canada	Chile	China	Columbia	Czech Republic
30	13	1	2	5
Egypt	Eritrea	Finland	France	Germany
1	1	2	3	4
Ghana	Greece	Hungary	India	Ireland
1	5	3	1	1
Israel	Italy	Japan	Kenya	Mexico
2	5	1	6	39
New Zealand	Poland	Portugal	Romania	Saudi Arabia
6	14	25	1	1
Slovakia	South Africa	Spain	Sweden	Tunisia
1	46	7	5	1
Turkey	United Kingdom	United States	Venezuela	Vietnam
2	21	17	1	1
Zimbabwe				
1				

# \$religious\_affiliation

Agnostic	Anti-religious	Atheist	Buddhist	Christian
6	1	2	2	123
Deist	Hindu	Muslim	None	Pagan
1	1	11	139	1
Spiritual	Spiritualist			
1	1			

# \$christian\_affiliation

Anglican	Apostolic	Baptist
1	1	11
Catholic	Congregationalist	Evangelical
68	1	4
Jehovah's Witness	Lutheran	Methodist
1	6	9
Non-religious Christian	None	Orthodox
1	4	4
Pentecostal	Presbyterian	Protestant
1	2	8

The output is likely difficult to read, so I will break it down here.

- **Sex:**

- Due to our use of the balance sex tool on Prolific, we have a very similar number of males ( $n = 146$ , 50.5%) and females ( $n = 143$ , 49.5%).

- **Gender:**

- Male: 144 (49.8%)
- Female: 135 (46.7%)
- Non-Binary: 4 (1.4%)
- Agender: 2 (.7%)
- Gender Fluid: 1 (.3%)
- Gender Queer: 1 (.3%)

- **Ethnicity:**

- The sample is predominantly White or Caucasian ( $n = 122$ , 42.2%), followed by Latino or Hispanic ( $n = 62$ , 21.5%), African ( $n = 52$ , 18.0%), East Asian ( $n = 14$ , 4.8%), and Mixed ( $n = 11$ , 3.8%). All other ethnic categories have  $n = 8$  or fewer participants in them ( 2.8%).

- **Nationality:**

- There are a large number of nationalities represented in our sample. The most common nationality is South Africa ( $n = 46$ , 15.9%), followed by Mexico ( $n = 39$ , 13.5%), Canada ( $n = 30$ , 10.4%), Portugal ( $n = 25$ , 8.7%), the United Kingdom ( $n = 21$ , 7.3%), and the United States ( $n = 17$ , 5.9%). All other nationalities have 14 or fewer participants ( 4.8%).

- **Religious Affiliation:**

- The modal response for Religious Affiliation in this sample is “None” ( $n = 139$ , 48.1%). The most common religious affiliation other than this is, by far, Christian ( $n = 123$ , 42.6%). The next most common categories is Muslim ( $n = 11$ , 3.8%) and Agnostic ( $n = 6$ , 2.1%). All other categories have two or fewer participants ( .7%).

- **Christian Affiliation:**

- Of those who identified as Christian, most are Catholic ( $n = 68$ , 55.3%). There are also a substantial number of Baptists ( $n = 11$ , 8.9%), Methodists ( $n = 9$ , 7.3%), and Protestants ( $n = 8$ , 6.5%). All other categories have four or fewer participants (3.3%).

## Numeric Variables

Now we will describe our numeric variables starting with our only demographic numeric variable, age.

```
# Generate descriptive statistics for age
describe(data$age)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
X1	1	289	30.85	10.85	28	29.18	7.41	18	86	68	1.74	3.84	0.64

- The average age in our sample is 30.85 ( $SD = 10.85$ ), with a trimmed mean of 29.18, indicating little skew. This indicates that our sample tends to be younger adults, but with quite a bit of variability.

Now we will describe the important scales for our analysis.

```
# Describe important scales in our analysis
describe(data %>% select("religious_importance", "CRS", "TDDS_f", "TDDS_p", "SOI_f", "SOI_a"))
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range
religious_importance	1	289	30.70	35.63	12.00	26.08	17.79	0	100	100
CRS	2	289	2.59	1.26	2.20	2.50	1.19	1	5	4
TDDS_f	3	289	60.12	16.05	59.00	59.82	16.31	12	100	88
TDDS_p	4	289	21.26	6.35	22.00	21.30	5.93	4	35	31
SOI_f	5	289	3.58	1.62	3.22	3.48	1.81	1	8	7
SOI_a	6	289	5.16	2.40	5.00	5.19	2.47	1	9	8
GENE_p	7	289	6.81	2.03	7.00	6.72	1.48	3	15	12
GENE_p_2	8	289	3.57	1.71	3.00	3.34	1.48	2	10	8
CONV_f	9	289	35.55	10.30	36.00	35.78	10.38	0	65	65

	skew	kurtosis	se
religious_importance	0.87	-0.83	2.10
CRS	0.57	-1.06	0.07
TDDS_f	0.17	-0.29	0.94
TDDS_p	-0.07	-0.49	0.37
SOI_f	0.53	-0.59	0.10

SOI_a	-0.03	-1.07	0.14
GENE_p	0.50	0.36	0.12
GENE_p_2	1.02	0.64	0.10
CONV_f	-0.24	0.44	0.61

- Other than a bit of skew for `religious_importance` and the `GENE_p_2`, there does not seem to be a problem here. Variables will be inspected more closely when testing assumptions for statistical tests.

Before moving on, we will calculate correlations between each of these variables.

```
# Calculate correlations for numerical variables
corr.test(data %>% select("age", "religious_importance", "CRS", "TDDS_f", "TDDS_p", "TDDS_s"
```

```
Call:corr.test(x = data %>% select("age", "religious_importance",
  "CRS", "TDDS_f", "TDDS_p", "TDDS_s", "SOI_f", "SOI_a", "GENE_p",
  "GENE_p_2", "CONV_f"))
```

Correlation matrix

	age	religious_importance	CRS	TDDS_f	TDDS_p	TDDS_s
age	1.00	-0.07	-0.09	0.02	-0.08	-0.09
religious_importance	-0.07	1.00	0.92	0.46	0.22	0.52
CRS	-0.09	0.92	1.00	0.47	0.24	0.53
TDDS_f	0.02	0.46	0.47	1.00	0.76	0.82
TDDS_p	-0.08	0.22	0.24	0.76	1.00	0.47
TDDS_s	-0.09	0.52	0.53	0.82	0.47	1.00
SOI_f	0.12	-0.28	-0.27	-0.38	-0.16	-0.54
SOI_a	0.14	-0.39	-0.41	-0.45	-0.22	-0.59
GENE_p	-0.06	0.07	0.07	0.02	0.01	0.07
GENE_p_2	0.03	0.02	0.00	-0.02	-0.03	0.04
CONV_f	0.08	0.33	0.31	0.12	0.01	0.13

	SOI_f	SOI_a	GENE_p	GENE_p_2	CONV_f
age	0.12	0.14	-0.06	0.03	0.08
religious_importance	-0.28	-0.39	0.07	0.02	0.33
CRS	-0.27	-0.41	0.07	0.00	0.31
TDDS_f	-0.38	-0.45	0.02	-0.02	0.12
TDDS_p	-0.16	-0.22	0.01	-0.03	0.01
TDDS_s	-0.54	-0.59	0.07	0.04	0.13
SOI_f	1.00	0.88	-0.11	-0.10	-0.10
SOI_a	0.88	1.00	-0.15	-0.11	-0.15
GENE_p	-0.11	-0.15	1.00	0.88	0.08
GENE_p_2	-0.10	-0.11	0.88	1.00	0.03
CONV_f	-0.10	-0.15	0.08	0.03	1.00

Sample Size

[1] 289

Probability values (Entries above the diagonal are adjusted for multiple tests.)

	age	religious_importance	CRS	TDDS_f	TDDS_p	TDDS_s	SOI_f
age	0.00	1.00	1.00	1.00	1.00	1.00	1.00
religious_importance	0.21	0.00	0.00	0.00	0.01	0.00	0.00
CRS	0.15	0.00	0.00	0.00	0.00	0.00	0.00
TDDS_f	0.79	0.00	0.00	0.00	0.00	0.00	0.00
TDDS_p	0.15	0.00	0.00	0.00	0.00	0.00	0.21
TDDS_s	0.14	0.00	0.00	0.00	0.00	0.00	0.00
SOI_f	0.05	0.00	0.00	0.00	0.01	0.00	0.00
SOI_a	0.01	0.00	0.00	0.00	0.00	0.00	0.00
GENE_p	0.35	0.20	0.25	0.75	0.81	0.21	0.06
GENE_p_2	0.66	0.79	1.00	0.72	0.67	0.55	0.09
CONV_f	0.17	0.00	0.00	0.04	0.93	0.03	0.10

	SOI_a	GENE_p	GENE_p_2	CONV_f
age	0.42	1.00	1.00	1.00
religious_importance	0.00	1.00	1.00	0.00
CRS	0.00	1.00	1.00	0.00
TDDS_f	0.00	1.00	1.00	1.00
TDDS_p	0.01	1.00	1.00	1.00
TDDS_s	0.00	1.00	1.00	0.74
SOI_f	0.00	1.00	1.00	1.00
SOI_a	0.00	0.29	1.00	0.32
GENE_p	0.01	0.00	0.00	1.00
GENE_p_2	0.06	0.00	0.00	1.00
CONV_f	0.01	0.16	0.62	0.00

To see confidence intervals of the correlations, print with the short=FALSE option

- For each correlation the sample size is  $N = 289$ .
- Noteworthy for this analysis, pathogen disgust is positively and significantly related to both **religious\_importance** ( $r = .22$ ,  $p < .01$ ) and the **CRS** ( $r = .24$ ,  $p < .01$ ), but sexual disgust is more strongly related to both **religious\_importance** ( $r = .52$ ,  $p < .01$ ) and the **CRS** ( $r = .53$ ,  $p < .01$ ). This replicates previous work and provides initial support for the sexual strategies hypothesis.
  - In addition, there is a very strong correlation between **religious\_importance** and the **CRS** ( $r = .92$ ,  $p < .01$ ).
  - However, neither the **GENE\_p** nor the **GENE\_p\_2** are significantly related to either **religious\_importance** ( $r = .07$ ,  $p = .20$ ;  $r = .02$ ,  $p = .02$ ; respectively) or the **CRS** ( $r = .07$ ,  $p = .25$ ;  $r = .00$ ,  $p = .00$ ; respectively), suggesting that it could not mediate the relationship between disgust sensitivity and religiosity.

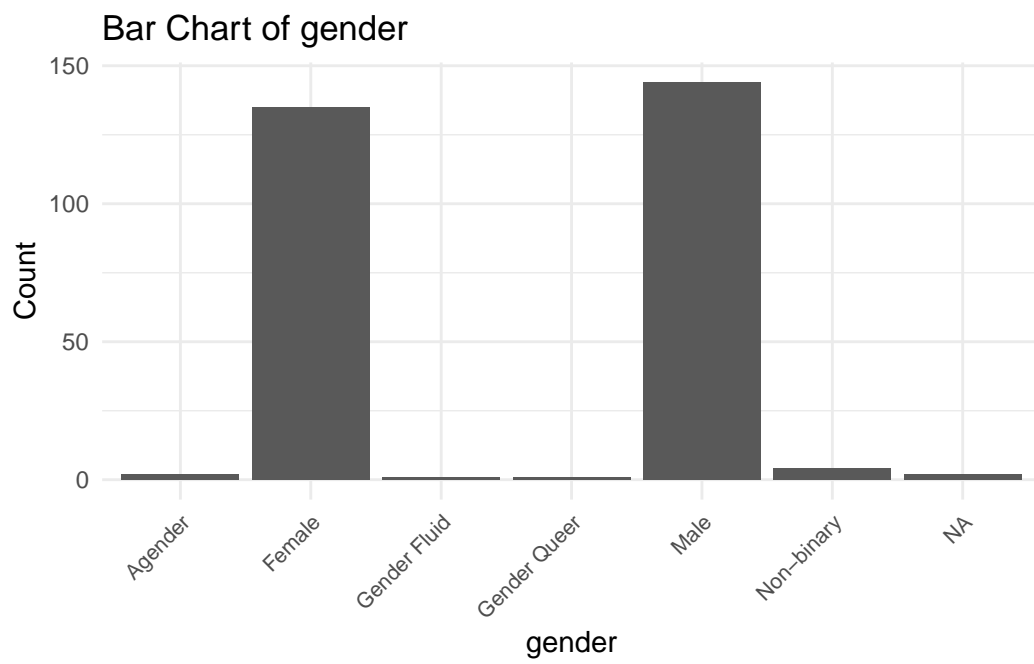
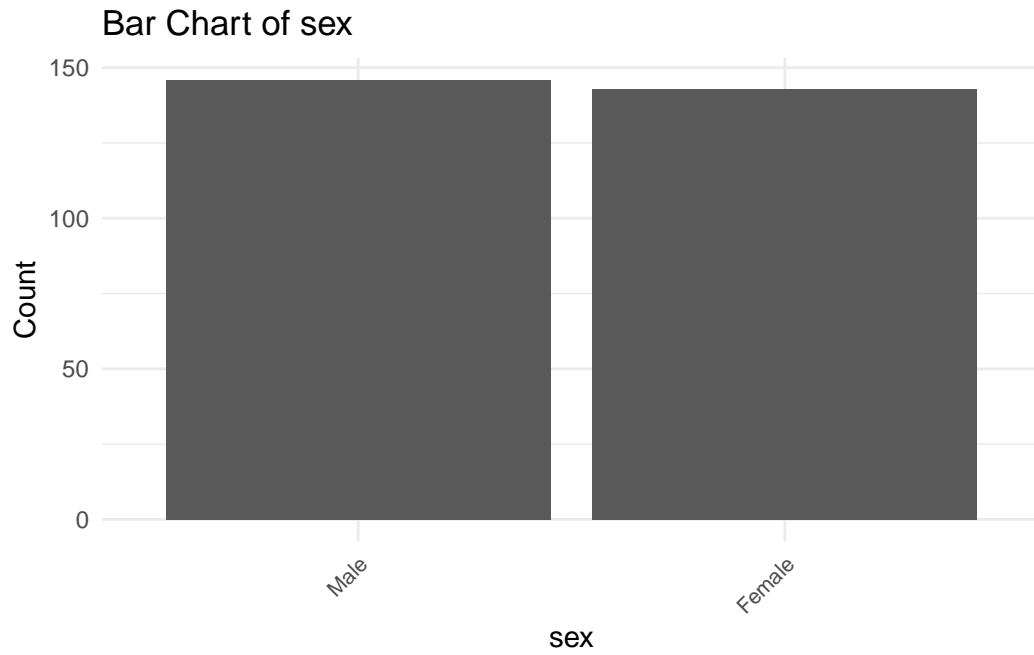
- Unlike the **GENE** variables, **CONV\_f** is significantly related to both **religious\_importance** ( $r = .33, p < .01$ ) and the **CRS** ( $r = .31, p < .01$ ).
- Next, the only potential mediator that is significantly related to pathogen disgust is the **SOI\_a**:
  - **SOI\_a**:  $r = -.22, p = .01$
  - **GENE\_p**:  $r = .01, p = 1$ 
    - \* **GENE\_p\_2**:  $r = -.03, p = 1$
  - **CONV\_f**:  $r = .01, p = 1$
  - This suggests that the only plausible mediator (here, that is) of the relationship between disgust and religiosity could be monogamous mating strategies.

## Data Visualization

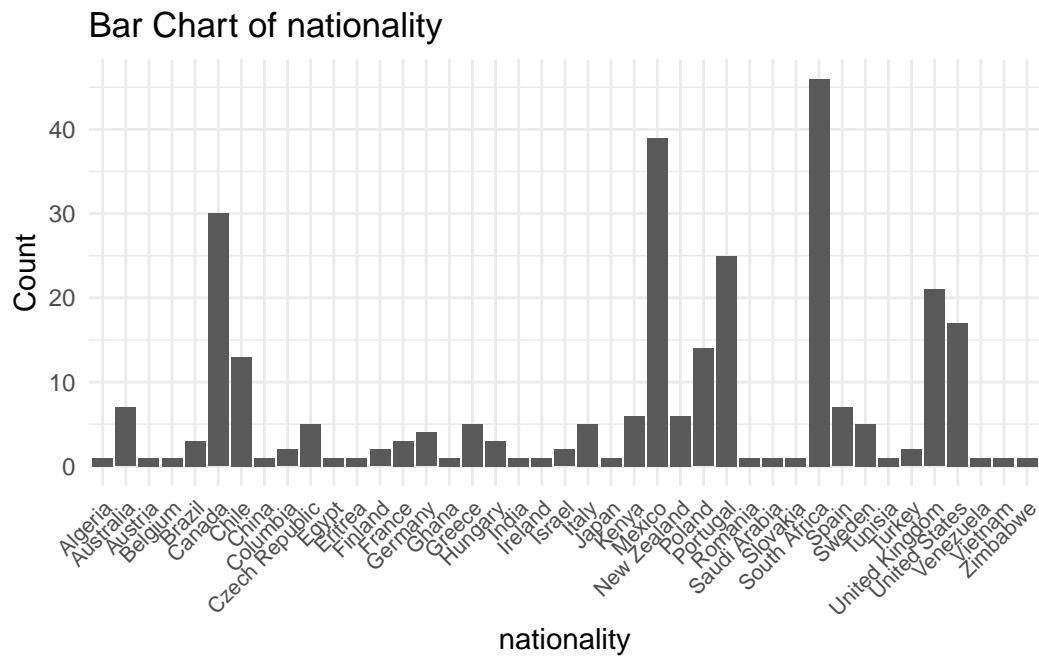
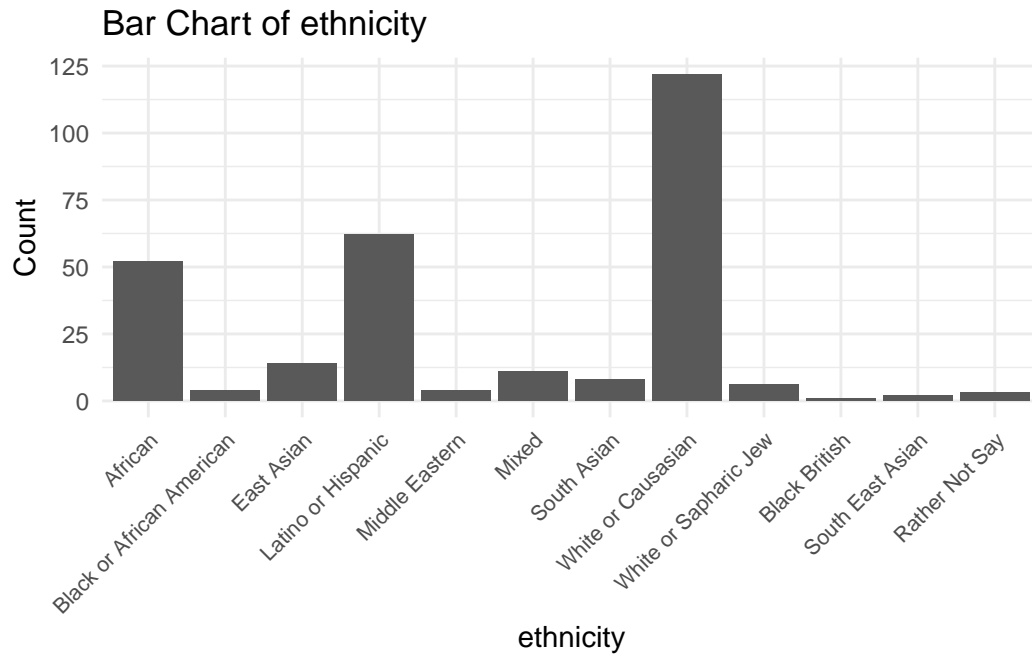
Now we will visualize the distribution of our variables by creating bar plots and histograms (for categorical and numerical variables, respectively).

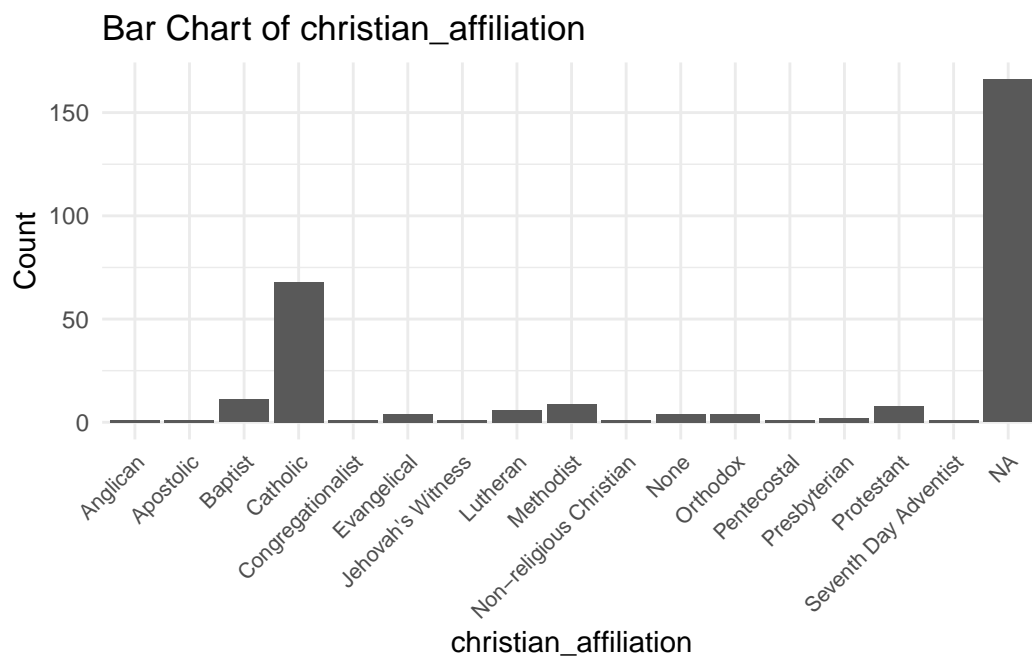
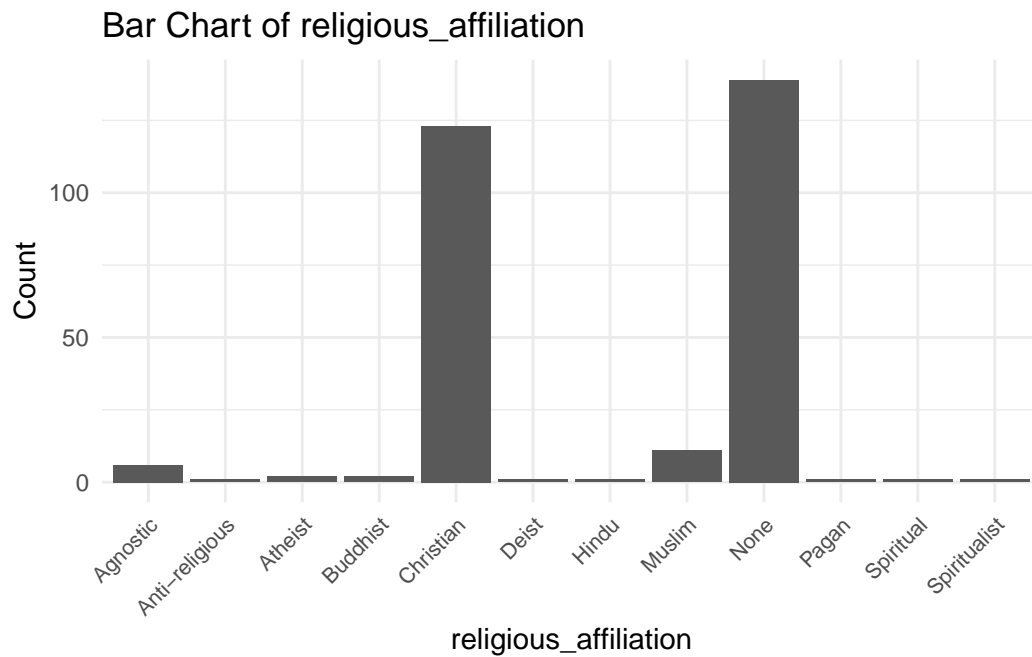
### Categorical Variables

```
# Loop through each categorical variable in the data frame and create a bar chart
for (var in names(data[, sapply(data, is.factor)])) {
  p <- ggplot(data[, sapply(data, is.factor)], aes(x = !!sym(var))) +
    geom_bar() +
    theme_minimal() +
    labs(title = paste("Bar Chart of", var), x = var, y = "Count") +
    theme(axis.text.x = element_text(angle = 45, hjust = 1, size = 8)) # Adjust x axis labels
  # Print each bar chart
  print(p)
  # Remove the plot object so it doesn't clutter the environment
  rm(p)
}
```







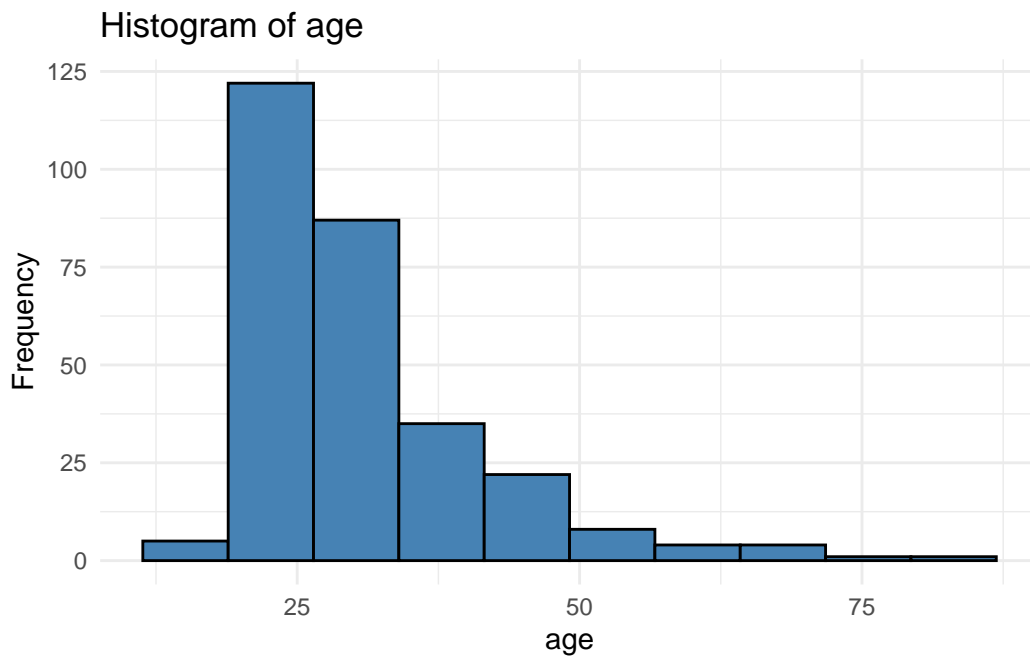


```
# Remove the var object
rm(var)
```

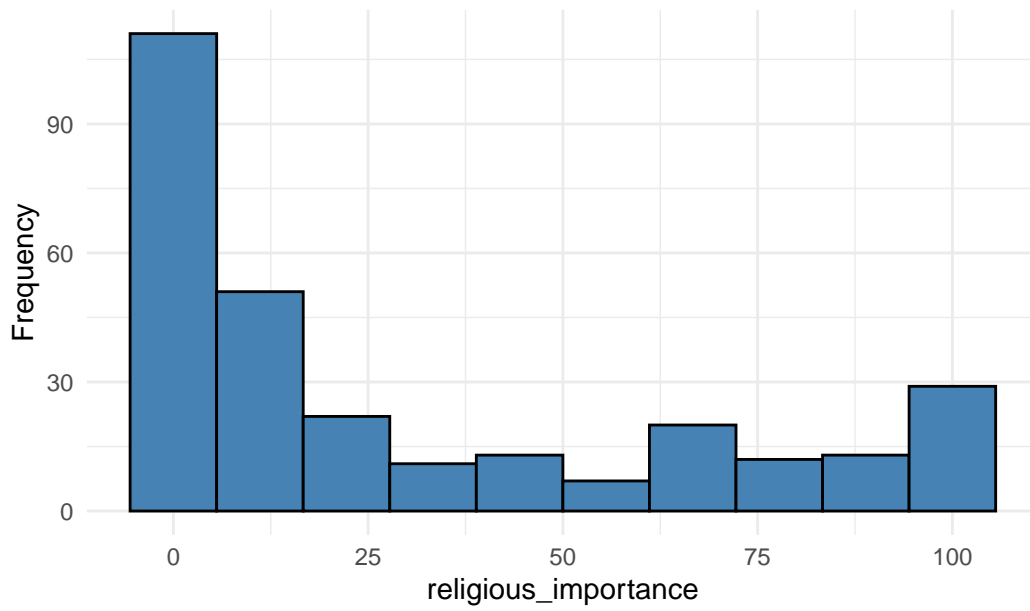
## Numeric Variables

```
# Loop through numeric variable identified to produce a histogram
for (var in names(data %>% select("age", "religious_importance", "CRS", "TDDS_f", "TDDS_p",
  p <- ggplot(data %>% select("age", "religious_importance", "CRS", "TDDS_f", "TDDS_p", "SOI
    geom_histogram(bins = 10, fill = "steelblue", color = "black") +
    theme_minimal() +
    labs(title = paste("Histogram of", var), x = var, y = "Frequency")

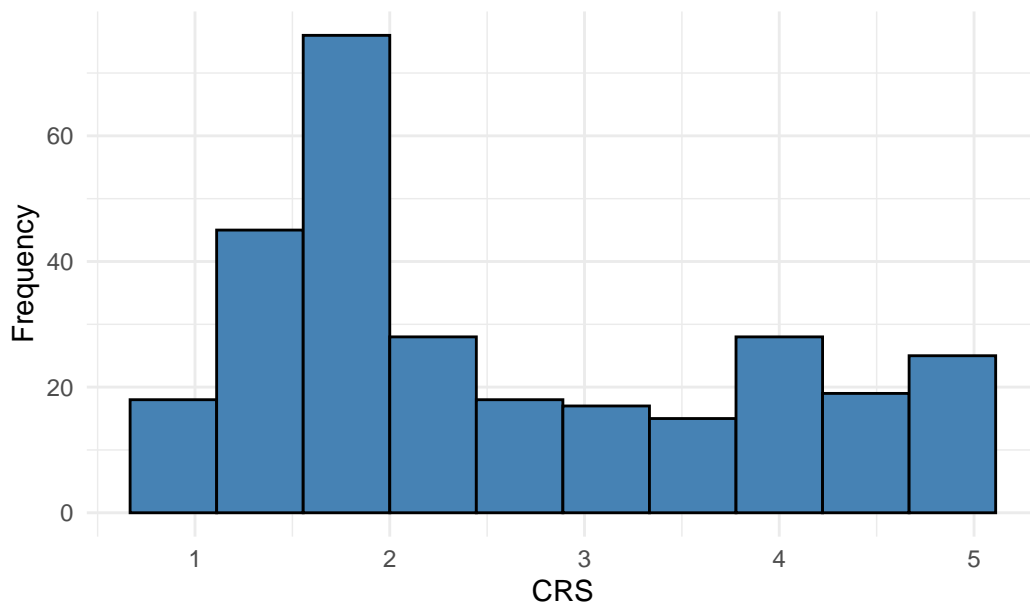
# Print each histogram
print(p)
# Remove the p object so it doesn't clutter up the environment
rm(p)
}
```

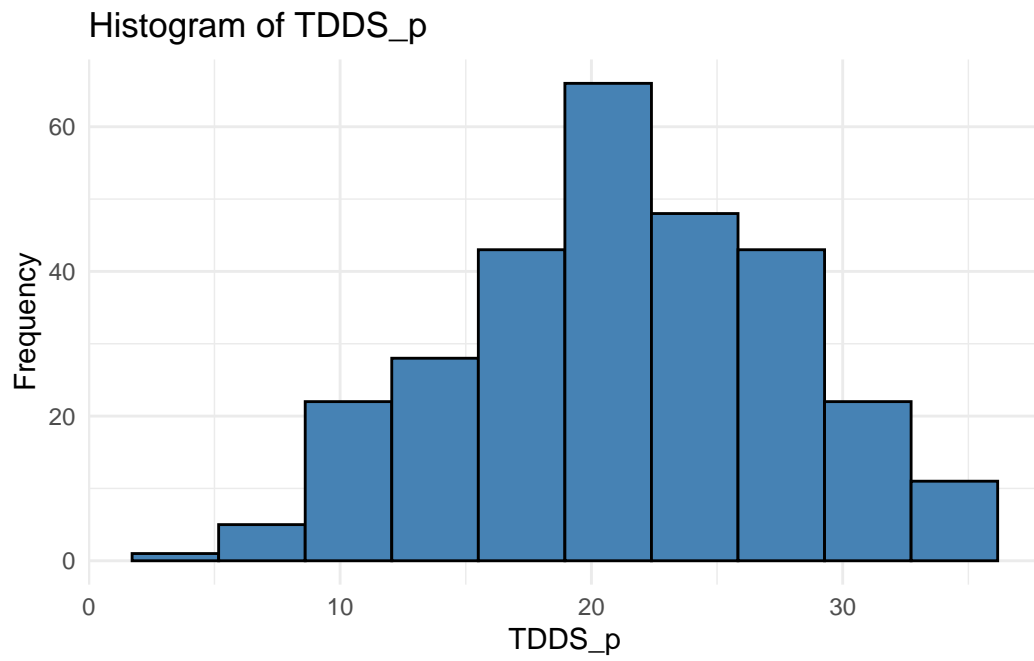
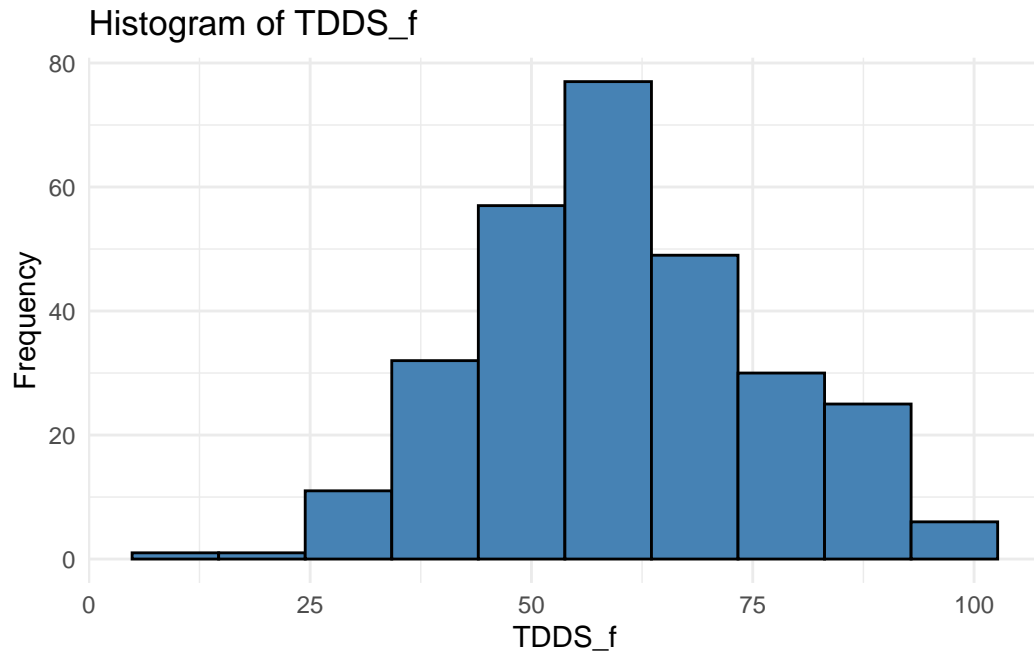


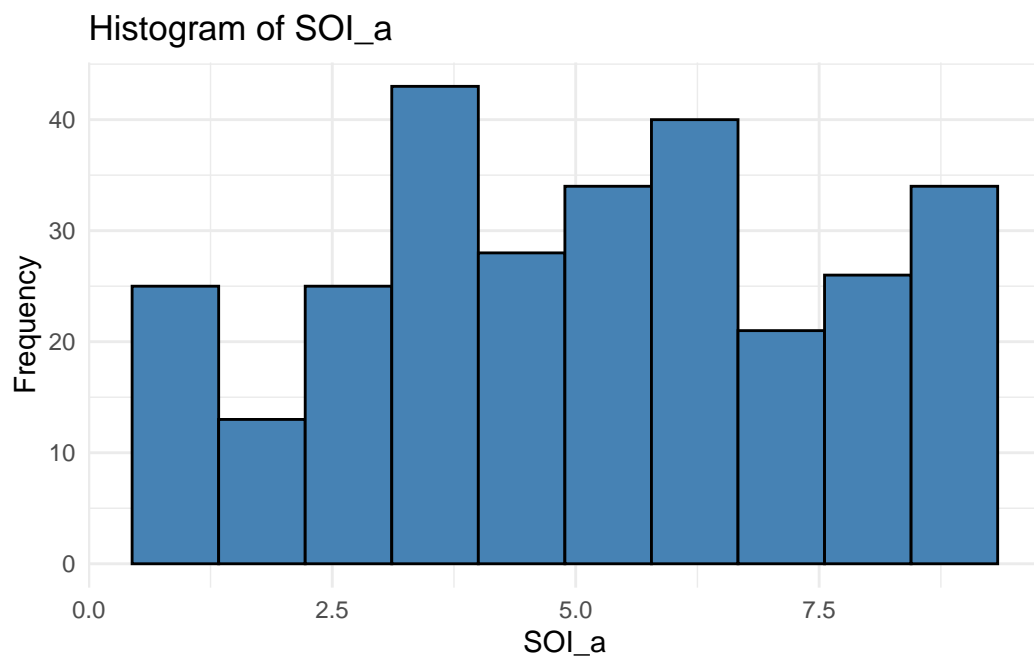
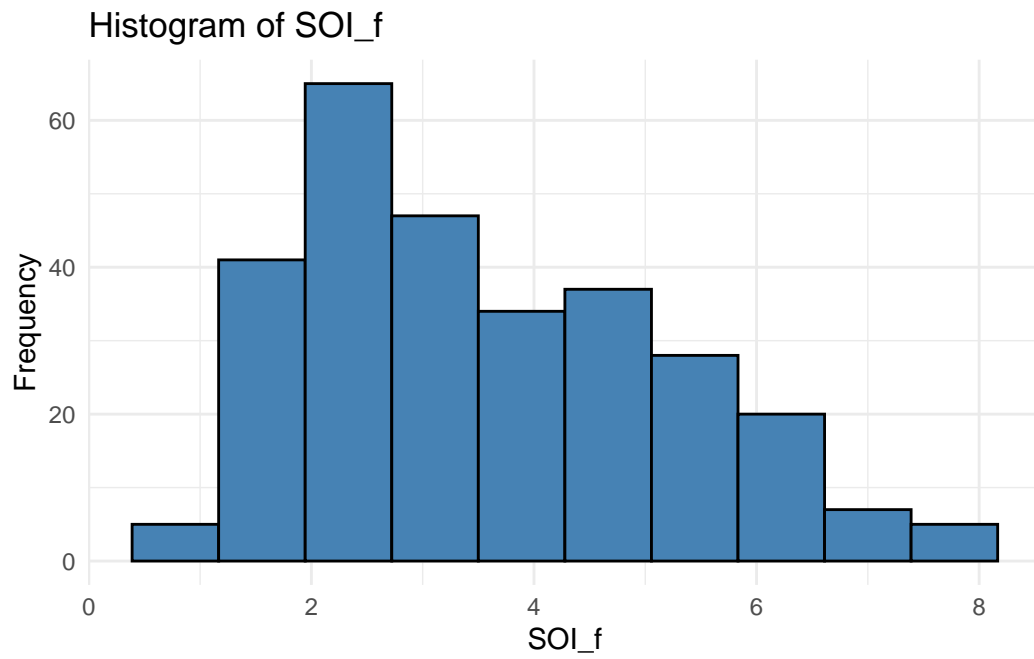
Histogram of religious\_importance

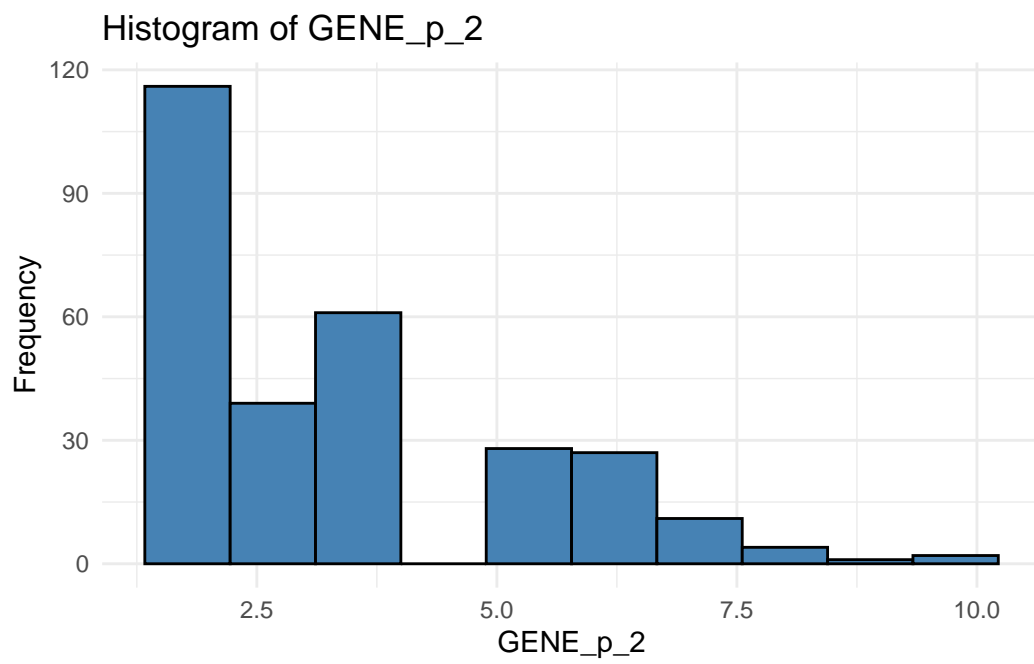
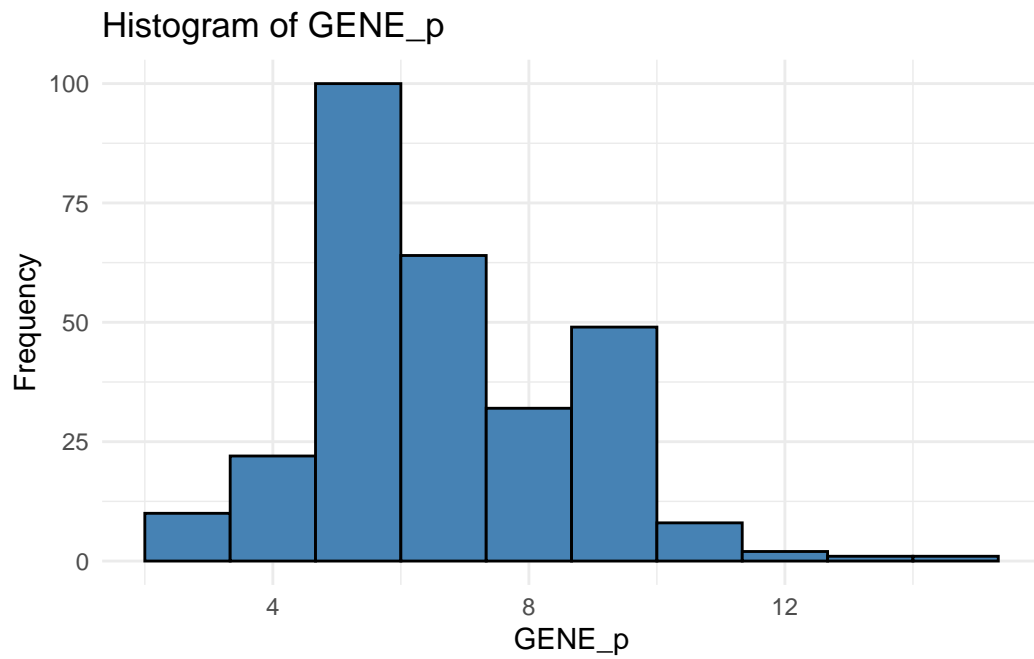


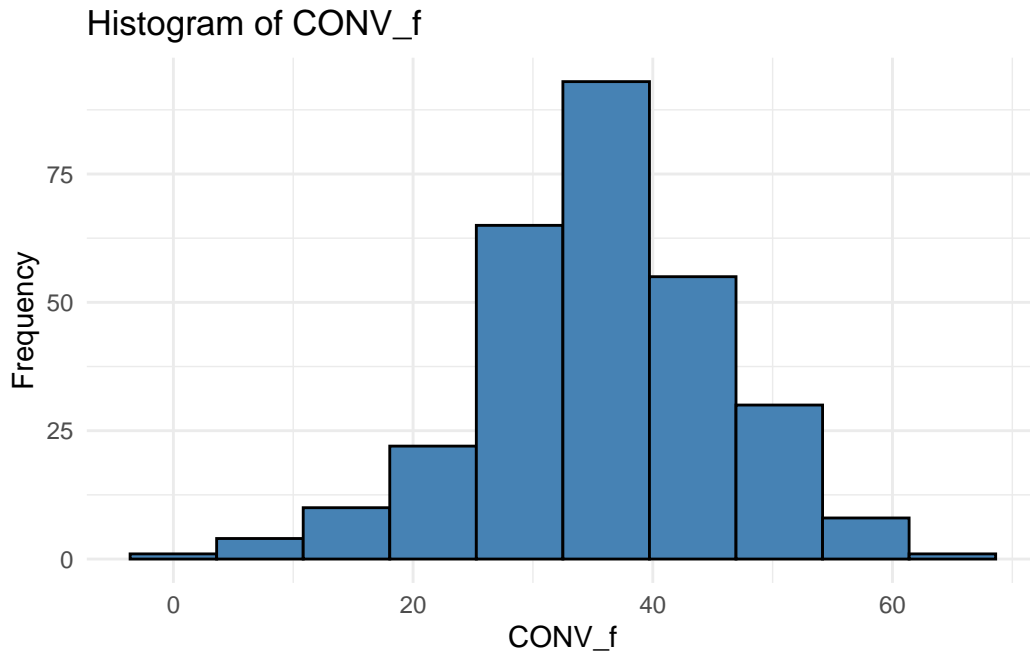
Histogram of CRS











```
# Remove the var object from the environment  
rm(var)
```

- The distributions of some of these variables indicate that they are non-normally distributed. With such a large sample ( $N = 289$ ), this should not be much of a problem, particularly because we will be using bootstrapped confidence intervals for much of the analysis.

## Testing Hypotheses

As above, the following are our hypotheses.

1. The relationship between disgust sensitivity and religiosity is motivated by adherence to traditional practices.
  - Strong version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be fully mediated by conventionalism.
  - Weak version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be partially mediated by conventionalism.



2. The relationship between disgust sensitivity and religiosity is motivated by out-group avoidance.
  - Strong version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be fully mediated by ethnocentrism.
  - Weak version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be partially mediated by ethnocentrism.
3. The relationship between disgust sensitivity and religiosity is motivated by a monogamous mating strategy.
  - Strong version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be fully mediated by restricted sociosexual attitudes.
  - Weak version of statistical hypothesis:
    - The positive relationship between parasite disgust sensitivity and religiosity will be partially mediated by restricted sociosexual attitudes.

The three hypotheses are each broken down into two statistical hypotheses, which represent their strong and weak version. In the strong version there is full mediation, and in the weak version there is only partial mediation. We have defined and preregistered inference criteria for both of these versions of each hypothesis.

#### **Strong Version Inference Criterion:**

For each model (each hypothesis) we will use the “K” method outlined by Beribisky et al. (2020) to infer full mediation. That is, if the proportion of variance explained by the indirect effect compared to the total effect is 80% or higher, we will infer full mediation. This will provide support for the strong version of the respective hypothesis.

#### **Weak Version Inference Criterion:**

Where there is not full mediation, for each model (each hypothesis) if the 95% confidence interval for the indirect effect of parasite disgust sensitivity on religiosity through the mediator does not contain zero, then we will infer partial mediation. This will provide support for the weak version of the respective hypothesis.

### **Data Preparation**

Before testing our hypotheses, we will standardize each variable. We will also reverse the scores of `SOI_a` before standardizing it, because higher scores currently represent more unrestricted sociosexual attitudes, whereas we would like higher scores to represent more restricted sociosexual attitudes (as a proxy for a monogamous mating strategy).

```

# Reversing the scoring for the SOI_a (saving it as SOI_a_r) and standardizing all variables
data <- data %>%
  mutate(
    SOI_a_r = (max(data$SOI_a) + 1) - SOI_a,
    SOI_a_r_z = scale(SOI_a_r),
    CRS_z = scale(CRS),
    religious_importance_z = scale(religious_importance),
    GENE_p_z = scale(GENE_p),
    GENE_p_2_z = scale(GENE_p_2),
    CONV_f_z = scale(CONV_f),
    TDDS_p_z = scale(TDDS_p)
  )

# Ensure that the variables are plain numeric vectors
data$TDDS_p_z <- as.numeric(data$TDDS_p_z)
data$CONV_f_z <- as.numeric(data$CONV_f_z)
data$CRS_z <- as.numeric(data$CRS_z)
data$SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
data$religious_importance_z <- as.numeric(data$religious_importance_z)
data$GENE_p_z <- as.numeric(data$GENE_p_z)
data$GENE_p_2_z <- as.numeric(data$GENE_p_2_z)

# Because the boot function was throwing a fit about this earlier
CRS_z <- as.numeric(data$CRS_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

```

We will also load the packages involved in the mediation analysis.

```

# Defining the packages to use to do the mediation analysis
pkg <- c("car", "mediation")
# car for testing for multicollinearity and bootstrapping
# mediation for testing the indirect effect with bootstrapped CIs and accessing

# Reading in the packages with the groundhog package for reproducibility
# Message suppressed in order to allow for rendering of quarto document
suppressMessages(groundhog.library(pkg = pkg, date = "2024-06-01"))

# Dropping the pkg vector
rm(pkg)

```

## Testing the Baseline Model

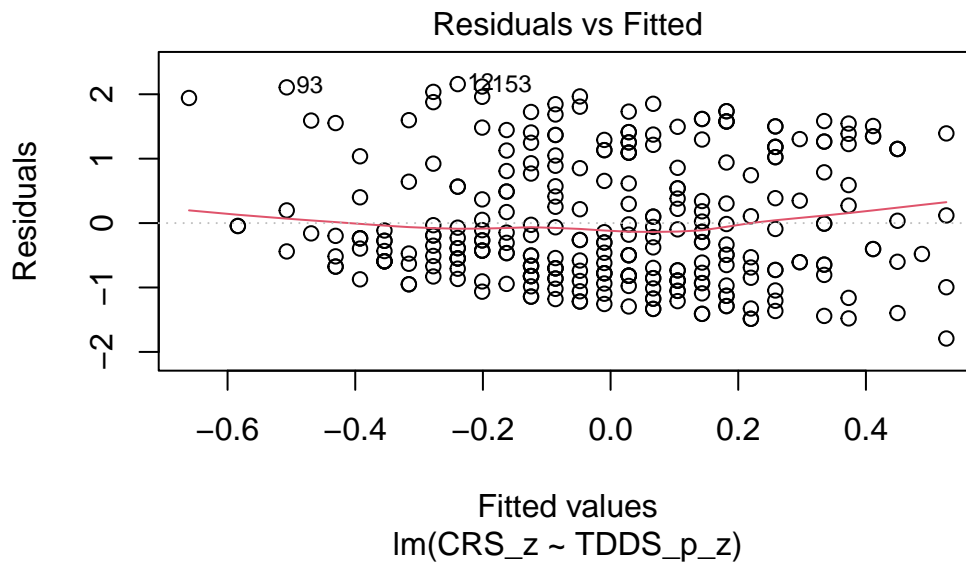
Before moving on, we will first test the baseline model where we regress the dependent variable (CRS) on the independent variable (TDDS\_p), because each mediation analysis requires that there is a significant positive relationship between these variables. We will also test each of the assumptions.

```
# Fitting the model
baseline_model <- lm(CRS_z ~ TDDS_p_z, data = data)
```

## Assumptions

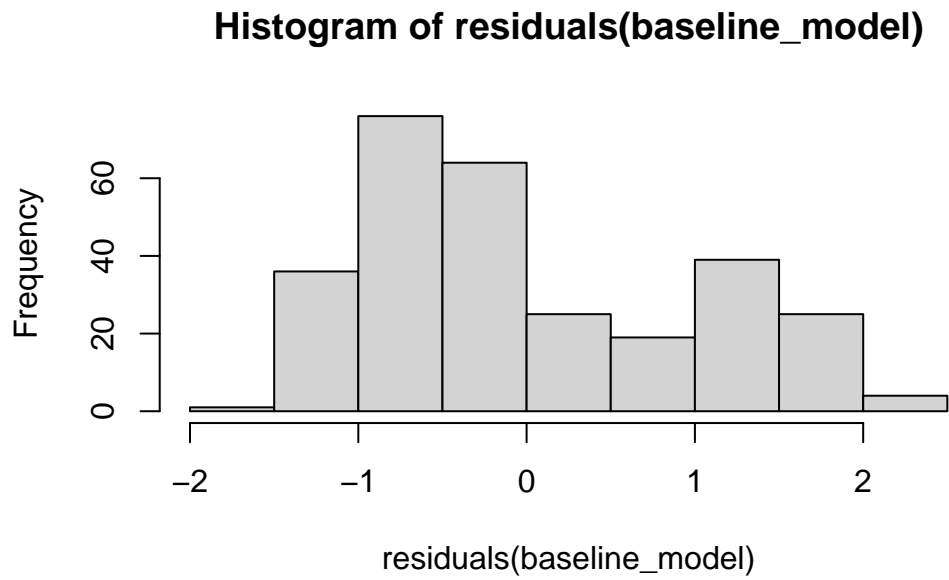
1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(baseline_model, which = 1)
```

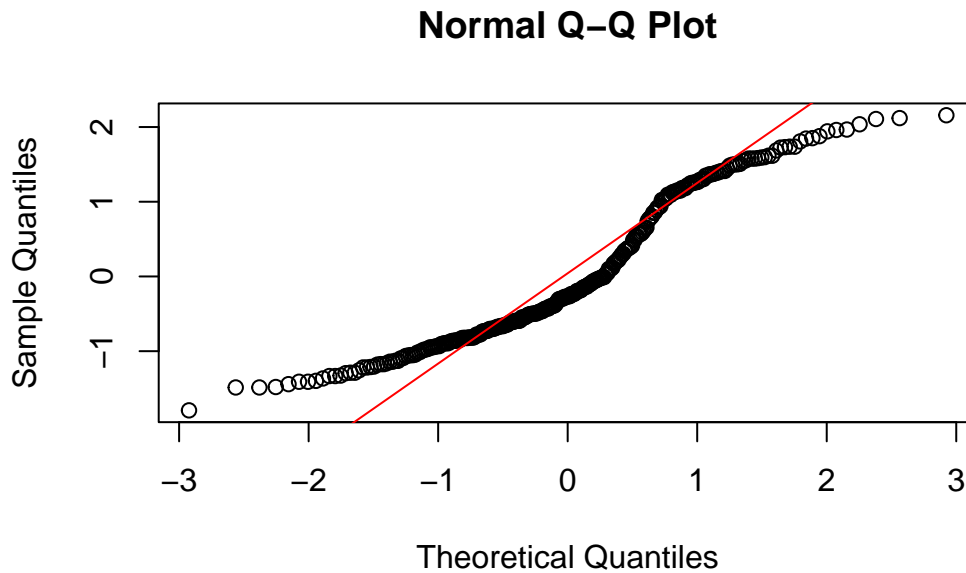


- The residuals vs. fitted plot does not stand out as violating the assumption of linearity or heteroscedasticity.
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

```
# Histogram for normality of residuals  
hist(residuals(baseline_model))
```



```
# QQ-plot for normality of residuals  
qqnorm(residuals(baseline_model))  
qqline(residuals(baseline_model), col = "red")
```



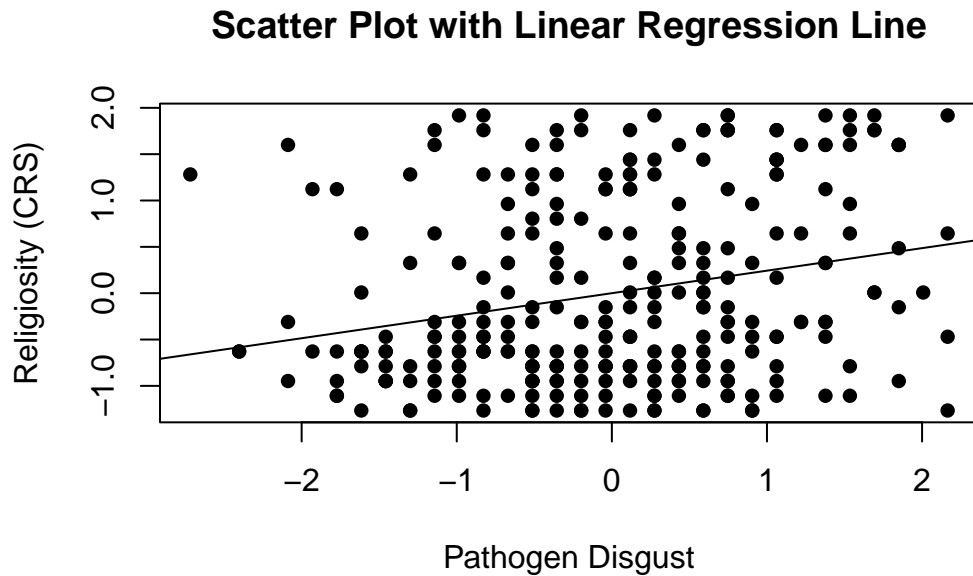
- The histogram of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just around -1 SD and around 2 SD. The two clusters of similar errors in prediction may be due to sex differences in disgust and religiosity. In exploratory analyses, we will add `sex` as a control variable.
- Because our inferential criteria do not require parametric tests, we should not be in too much trouble with a non-normality of residuals. We were relying on a parametric t-test to assess whether pathogen disgust is related to religiosity in the first place, however. Because we have failed the assumption of normality of residuals, we will use a bootstrapped test to make our inference here as well.

3. **Multicollinearity:** Not applicable because only one predictor

### Summarizing the Model

```
# Create a scatterplot with the regression line to visualize
plot(data$TDDS_p_z, data$CRS_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Pathogen Disgust",
     ylab = "Religiosity (CRS)",
     pch = 16, # For solid circle points)
```

```
col = "black")
# Add regression line
abline(baseline_model, col = "black", lwd = 1)
```



```
# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_baseline_model <- Boot(baseline_model, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_baseline_model) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
      original    bootBias  bootSE  bootMed
(Intercept) 1.6683e-17 -0.00066792 0.058534 -0.001485
TDDS_p_z    2.4299e-01  0.00007957 0.059718  0.245364
```

```
confint(boot_baseline_model) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.1064481	0.1239702
TDDS_p_z	0.1206344	0.3523252

```
# Removing the CRS_z and TDDS_p_z vectors in the environment that the Boot package required  
rm(CRS_z, TDDS_p_z)
```

- We can see from the Scatter Plot that there is a positive relationship between pathogen disgust and religiosity.
- Based on the estimates, the standardized coefficient for pathogen disgust is around .24. This aligns with estimates from a meta-analysis of this relationship (Yu et al., 2022).
  - In addition, this relationship is significant, as the BCIs do not contain zero (upper = .35, lower = .12).

This shows that the IV effects the DV, as in the second regression equation for testing mediation in Baron & Kenny (1986). To establish mediation, we will still need to establish that the IV affects the mediator and the mediator affects the DV. Before we test the indirect effect and proportion of variance explained by each effect for each mediation model, we will run these two models.

## Hypothesis 1: Adherence to Traditional Practices

### Preliminary Simple Linear Regression Models

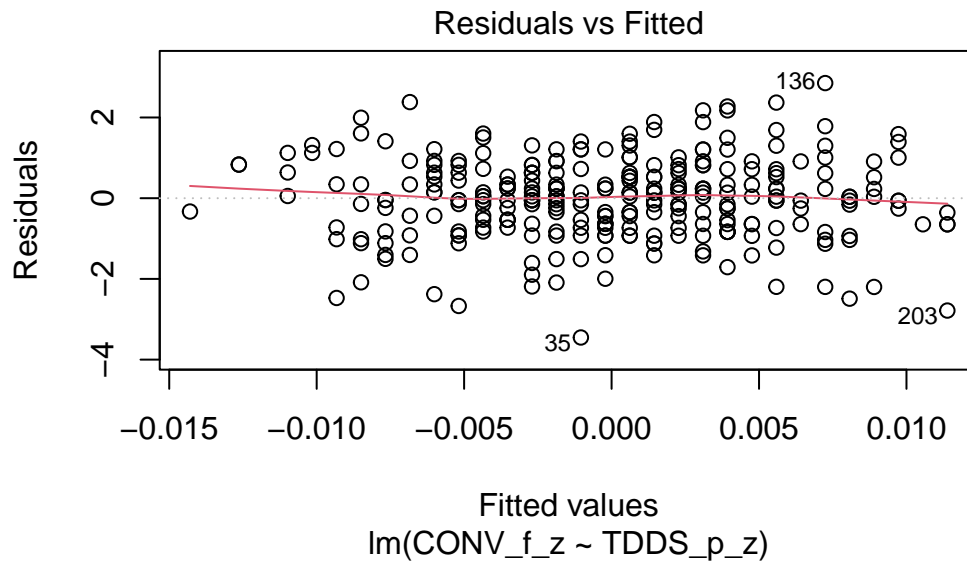
As mentioned above, to test Hypothesis 1, we must first establish that pathogen disgust affects adherence to traditional practices (i.e., CONV\_f\_z) and that adherence to traditional practices influences religiosity.

```
# Fitting the models  
H1_M1 <- lm(CONV_f_z ~ TDDS_p_z, data = data)  
H1_M2 <- lm(CRS_z ~ CONV_f_z, data = data)
```

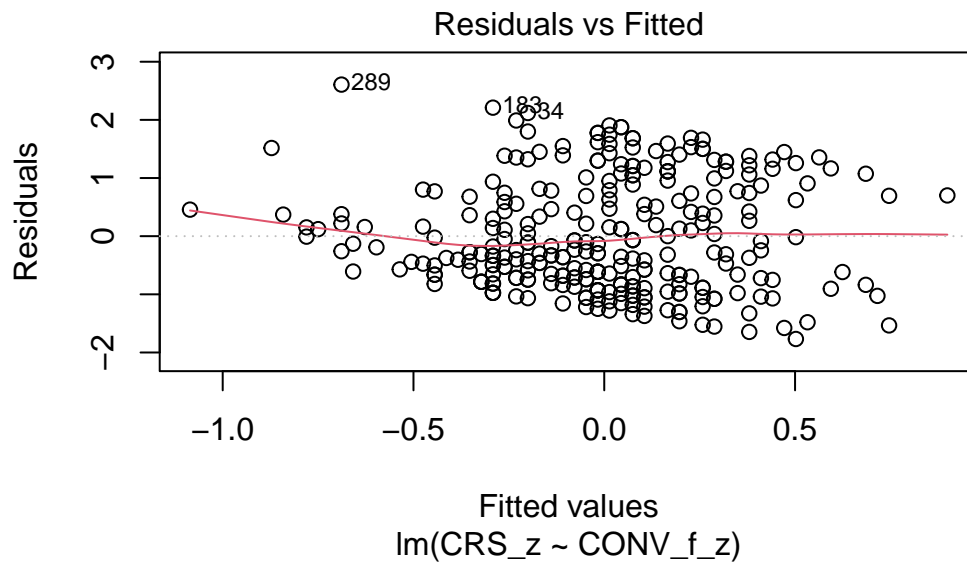
### Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H1_M1, which = 1) # For model 1
```



```
plot(H1_M2, which = 1) # For model 2
```

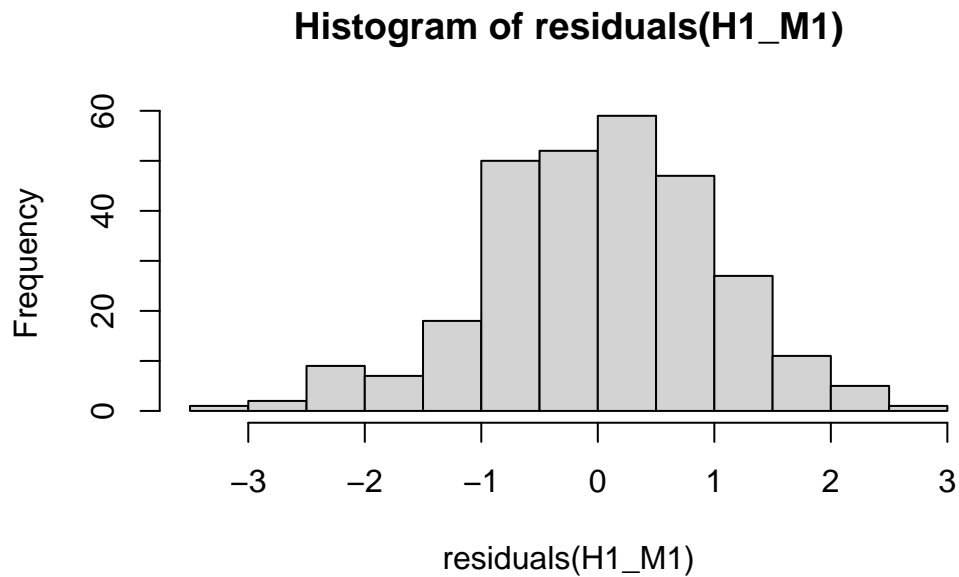




- The residuals vs. fitted plot does not stand out as violating the assumption of linearity or heteroscedasticity for the first model, but there is a trend towards more negative residual values at the higher fitted values of the model for model 2. This may be problematic.

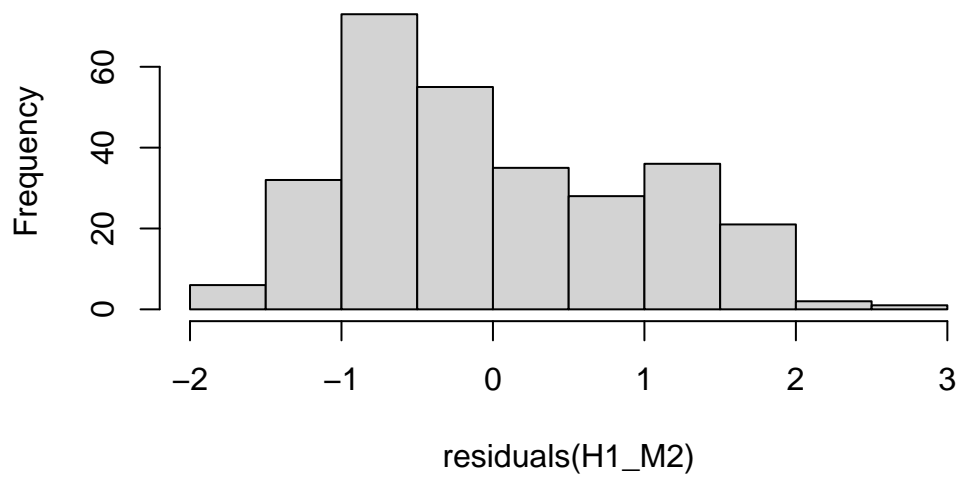
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

```
# Histogram for normality of residuals
hist(residuals(H1_M1)) # For model 1
```



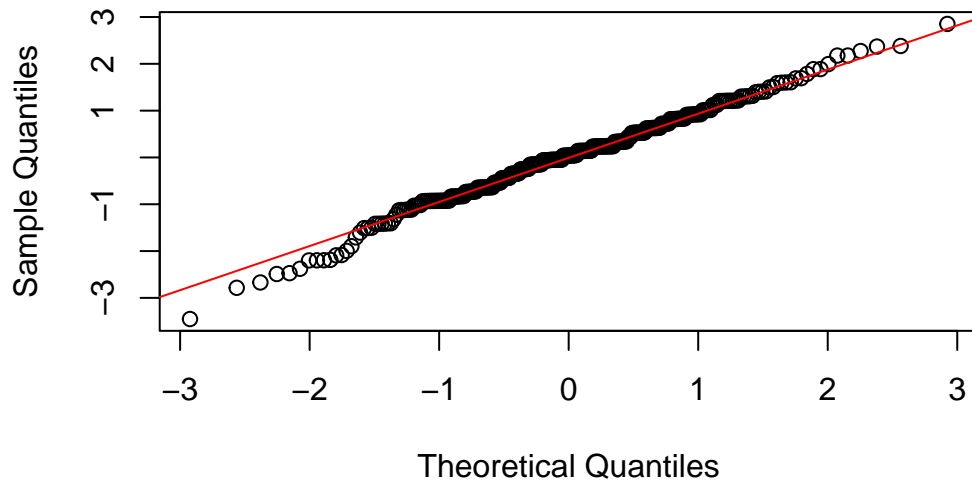
```
hist(residuals(H1_M2)) # For model 2
```

**Histogram of residuals(H1\_M2)**

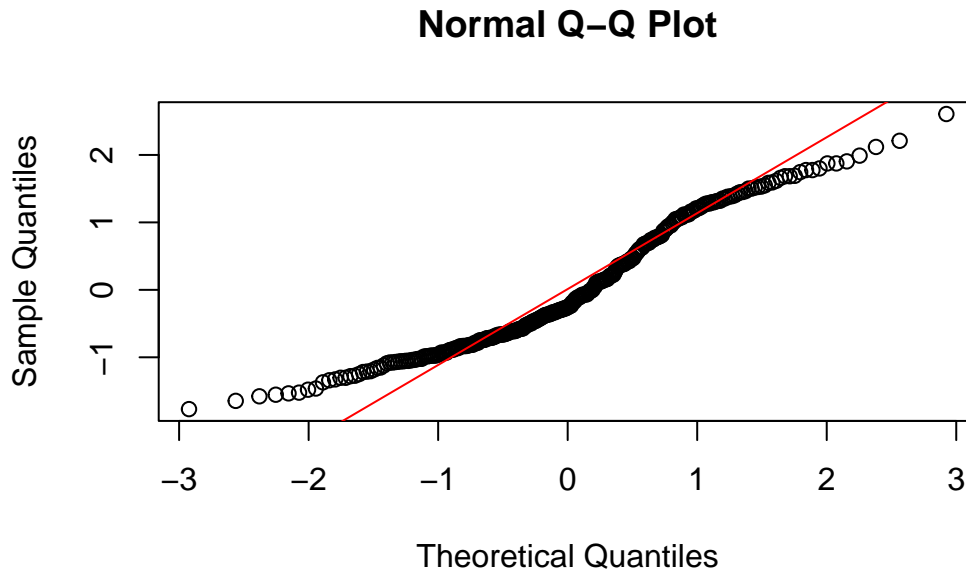


```
# QQ-plot for normality of residuals of Model 1  
qqnorm(residuals(H1_M1))  
qqline(residuals(H1_M1), col = "red")
```

**Normal Q-Q Plot**



```
# QQ-plot for normality of residuals of Model 2
qqnorm(residuals(H1_M2))
qqline(residuals(H1_M2), col = "red")
```



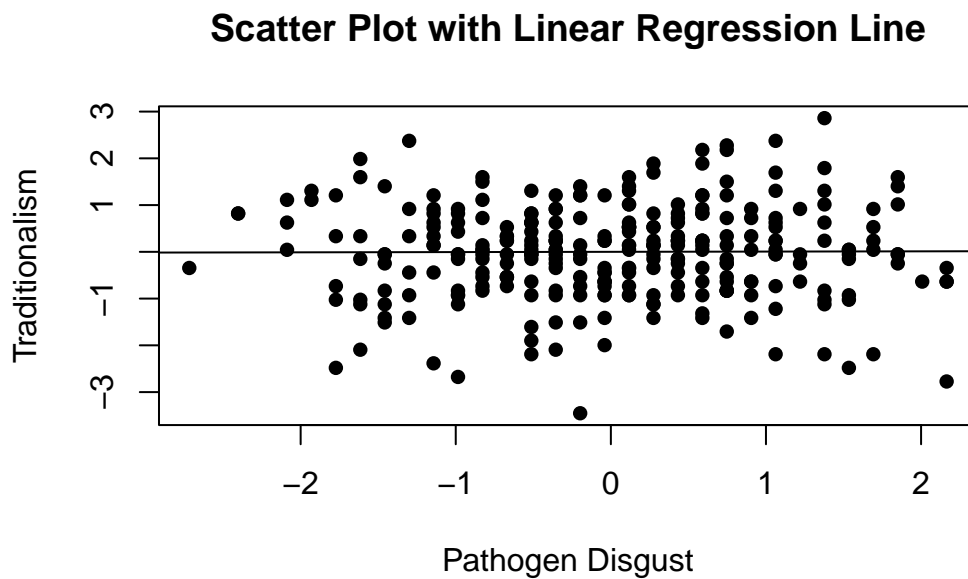
- **Model 1:**
  - The histogram and QQ-plot suggest that the residuals are normally distributed for Model 1.
- **Model 2:**
  - The histogram and QQ-plot of the residuals of Model 2 indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 SD. The two clusters of similar errors in prediction may be due to sex differences in religiosity. In exploratory analyses, we will add **sex** as a control variable.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictor for Model 2.

3. **Multicollinearity:** Not applicable because each model only has one predictor.

### Summarizing the Models

## Model 1

```
# Create a scatterplot with the regression line to visualize
plot(data$TDDS_p_z, data$CONV_f_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Pathogen Disgust",
     ylab = "Traditionalism",
     pch = 16, # For solid circle points
     col = "black")
# Add regression line
abline(H1_M1, col = "black", lwd = 1)
```



```
# Summarizing the model
summary(H1_M1)
```

Call:

```
lm(formula = CONV_f_z ~ TDDS_p_z, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.4517	-0.6429	0.0402	0.6287	2.8524

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-7.860e-17	5.893e-02	0.000	1.000
TDDS_p_z	5.260e-03	5.903e-02	0.089	0.929

Residual standard error: 1.002 on 287 degrees of freedom

Multiple R-squared: 2.766e-05, Adjusted R-squared: -0.003457

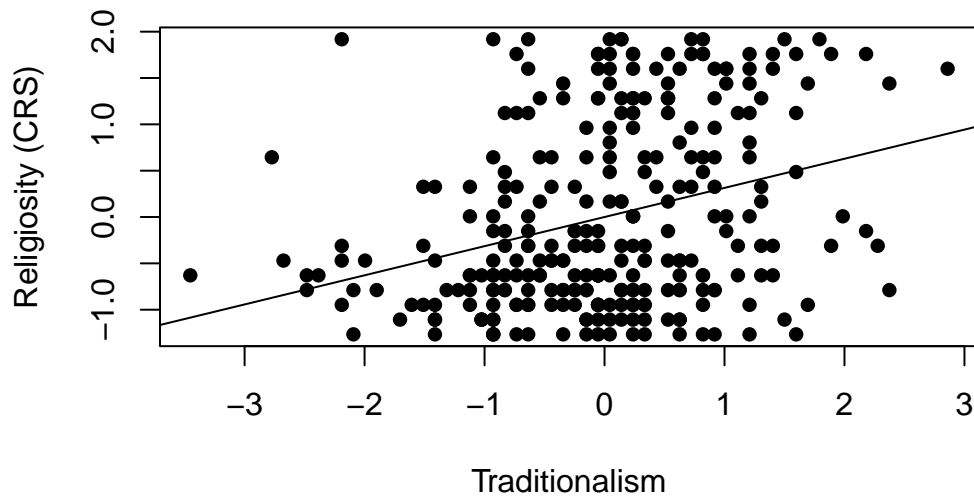
F-statistic: 0.00794 on 1 and 287 DF, p-value: 0.9291

- We can see from the Scatter Plot that there is no relationship between pathogen disgust and traditionalism to speak of ( $B = .005$ ,  $t(287) = .089$ ,  $p = .929$ ), just as with the correlation above.
  - Although this precludes the possibility that adherence to tradition mediates the relationship between disgust and religiosity, we will continue with the analysis for completions sake.

## Model 2

```
# Create a scatterplot with the regression line to visualize
plot(data$CONV_f_z, data$CRS_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Traditionalism",
     ylab = "Religiosity (CRS)",
     pch = 16, # For solid circle points
     col = "black")
# Add regression line
abline(H1_M2, col = "black", lwd = 1)
```

## Scatter Plot with Linear Regression Line



```
# Adding the variables to the broader environment so the Boot function will run
CONV_f_z <- as.numeric(data$CONV_f_z)
CRS_z <- as.numeric(data$CRS_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H1_M2 <- Boot(H1_M2, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H1_M2) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
      original    bootBias   bootSE    bootMed
(Intercept) 5.1338e-17 -6.9336e-04 0.056239 -0.00035888
CONV_f_z    3.1443e-01  5.1009e-05 0.054894  0.31405077
```

```
confint(boot_H1_M2) # Getting the BCIs
```

Bootstrap bca confidence intervals

```

                2.5 %    97.5 %
(Intercept) -0.1074245 0.1127866
CONV_f_z      0.2047469 0.4243772

```

```

# Removing the vectors in the environment that the Boot package required outside of the data
rm(CONV_f_z, CRS_z)

```

- We can see from the Scatter Plot that there is a positive relationship between traditionalism and religiosity.
- Based on the estimates, the standardized coefficient for traditionalism is around .31.
  - In addition, this relationship is significant, as the BCIs do not contain zero (upper = .42, lower = .20).

## Full Mediation Model

The fact that pathogen disgust does not predict traditionalism indicates that traditionalism could not mediate the relationship between pathogen disgust and religiosity, but I will run the analysis that we planned anyway just to have the results.

```

# Fitting the full mediation model
H1_full <- lm(CRS_z ~ TDDS_p_z + CONV_f_z, data = data)

```

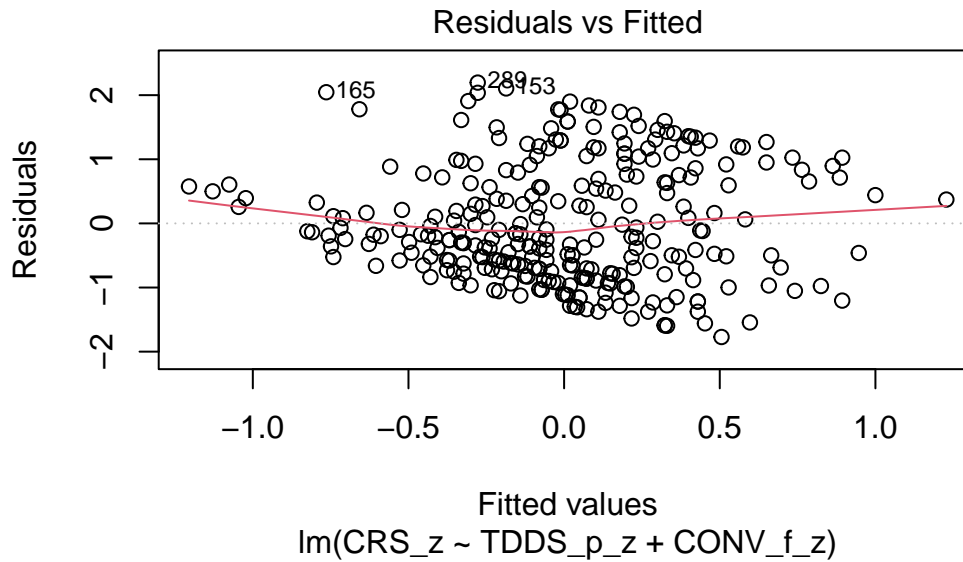
## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```

# Plotting the residual vs. fitted values
plot(H1_full, which = 1)

```



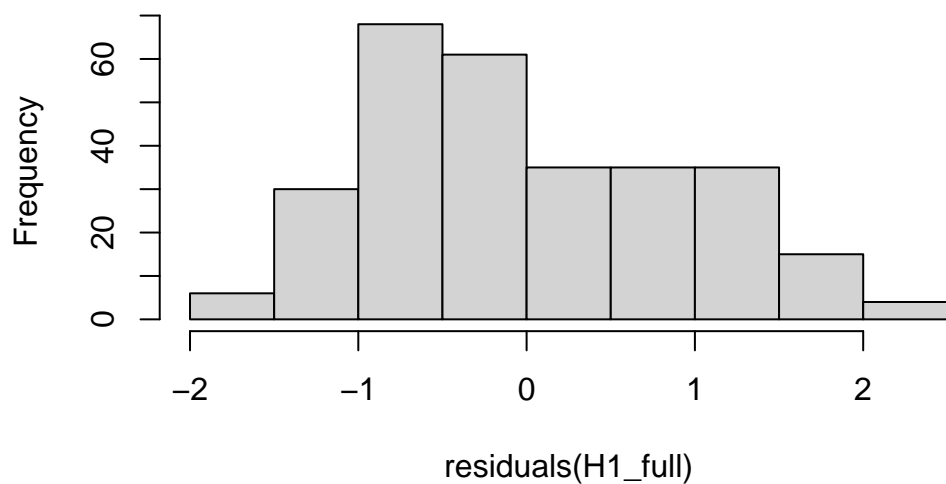
- There is a trend towards more negative residual values at the higher fitted values of the model. This may be problematic.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

```
# Histogram for normality of residuals
hist(residuals(H1_full))
```

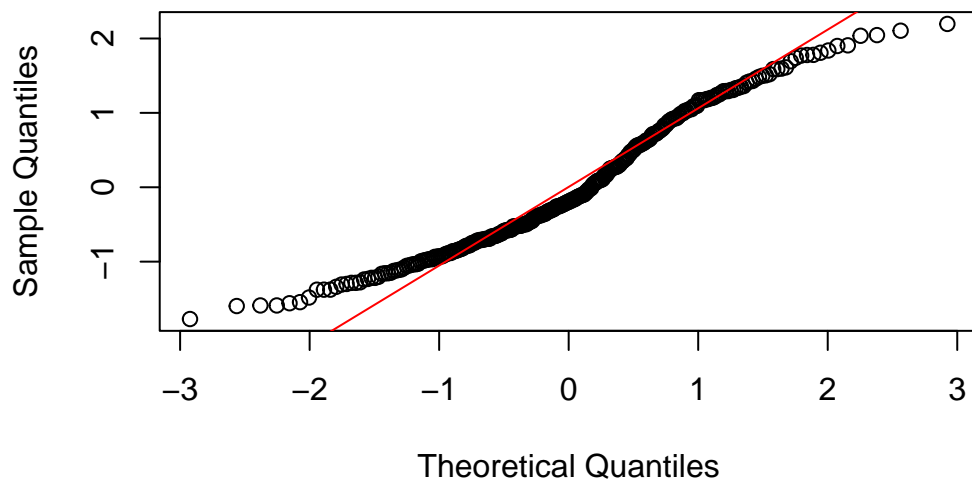


### Histogram of residuals(H1\_full)



```
# QQ-plot for normality of residuals of the full model
qqnorm(residuals(H1_full))
qqline(residuals(H1_full), col = "red")
```

### Normal Q-Q Plot



- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 SD. The two clusters of similar errors in prediction may be due to sex differences in religiosity. In exploratory analyses, we will add **sex** as a control variable.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictors for the model.

### 3. **Multicollinearity:** Variance Inflation Factor (VIF) from **car** package

```
# Calculating VIF
vif(H1_full)
```

```
TDDS_p_z CONV_f_z
1.000028 1.000028
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
CONV_f_z <- as.numeric(data$CONV_f_z)
CRS_z <- as.numeric(data$CRS_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H1_full <- Boot(H1_full, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H1_full) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original    bootBias    bootSE    bootMed
(Intercept) 4.1299e-17 -0.00139011 0.054836 -0.0025004
TDDS_p_z    2.4134e-01 -0.00021190 0.057757 0.2419430
CONV_f_z    3.1316e-01 0.00080321 0.050325 0.3138690
```

```
confint(boot_H1_full) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.1009681	0.1150922
TDDS_p_z	0.1269087	0.3490115
CONV_f_z	0.2120557	0.4121981

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(CONV_f_z, CRS_z, TDDS_p_z)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H1_mediate_results <- mediate(model.m = H1_M1,
  model.y = H1_full,
  treat = "TDDS_p_z",
  mediator = "CONV_f_z",
  boot = TRUE, # Bootstrap the CI for the indirect effect
  sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results for the mediation analysis
summary(H1_mediate_results)
```

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.00165	-0.03517	0.04	0.96
ADE	0.24134	0.13151	0.35	<2e-16 ***
Total Effect	0.24299	0.13360	0.36	<2e-16 ***
Prop. Mediated	0.00678	-0.18265	0.18	0.96

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- **Multiple Regression Results:**

- Based on the bootstrapped CIs not containing zero, both pathogen disgust and traditionalism are significantly positively related to religiosity.
- The coefficients for pathogen disgust (.24) and traditionalism (.31) are virtually identical in this model to the models above where they are stand-alone predictors. This suggests that the unique effects of pathogen disgust and traditionalism on religiosity over-and-above each other are completely independent effects.

- **Mediation Analysis Results:**

- The ACME indicates that there is no significant indirect effect of pathogen disgust on religiosity through traditionalism ( $B = .001$ ,  $BCI = [-.035, .04]$ ).
- This result is inconsistent with both the weak and strong version of Hypothesis 1. That is, there is no evidence here that the relationship between pathogen disgust and religiosity is mediated by traditionalism.

## Hypothesis 2: Out-Group Avoidance

### Preliminary Simple Linear Regression Models

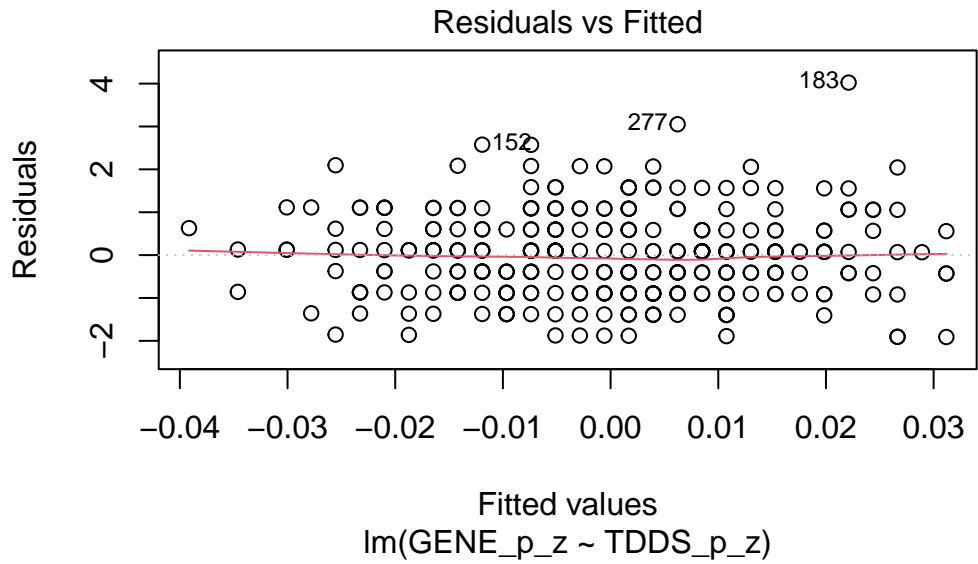
As mentioned above, to test Hypothesis 2, we must first establish that pathogen disgust affects out-group avoidance (i.e., GENE\_p\_z) and that out-group avoidance influences religiosity.

```
# Fitting the models
H2_M1 <- lm(GENE_p_z ~ TDDS_p_z, data = data)
H2_M2 <- lm(CRS_z ~ GENE_p_z, data = data)
```

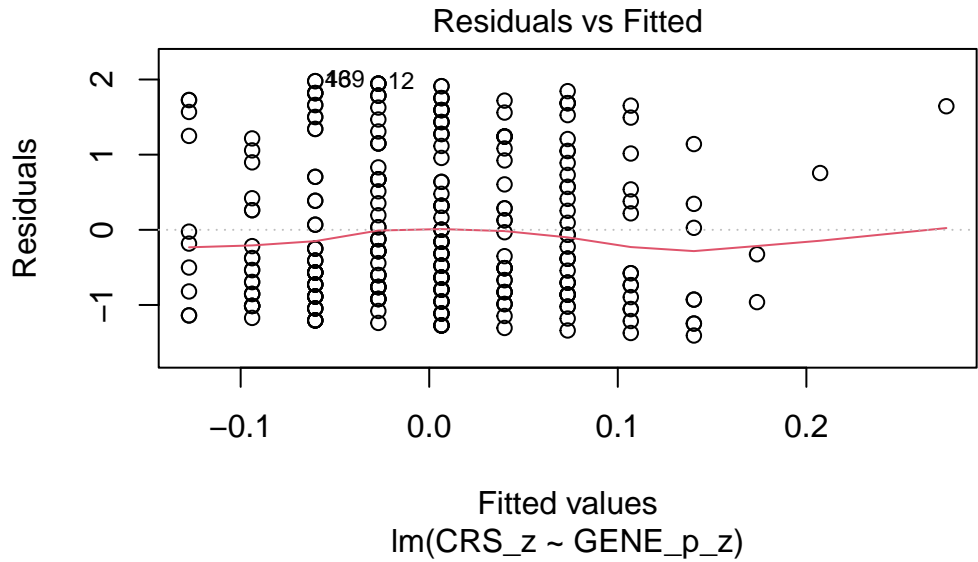
### Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H2_M1, which = 1) # For model 1
```



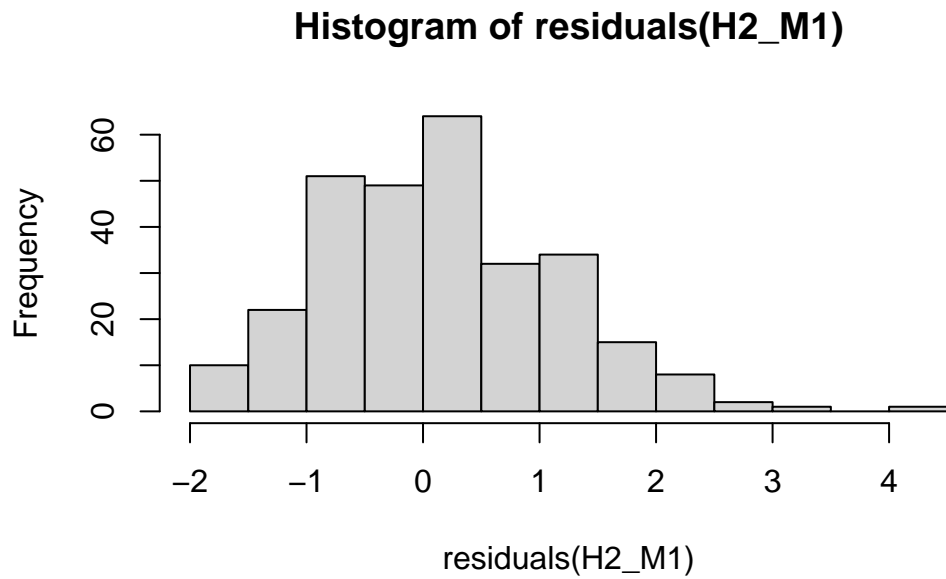
```
plot(H2_M2, which = 1) # For model 2
```



- The residuals vs. fitted plot does not stand out as violating the assumption of linearity or heteroscedasticity for the either model.

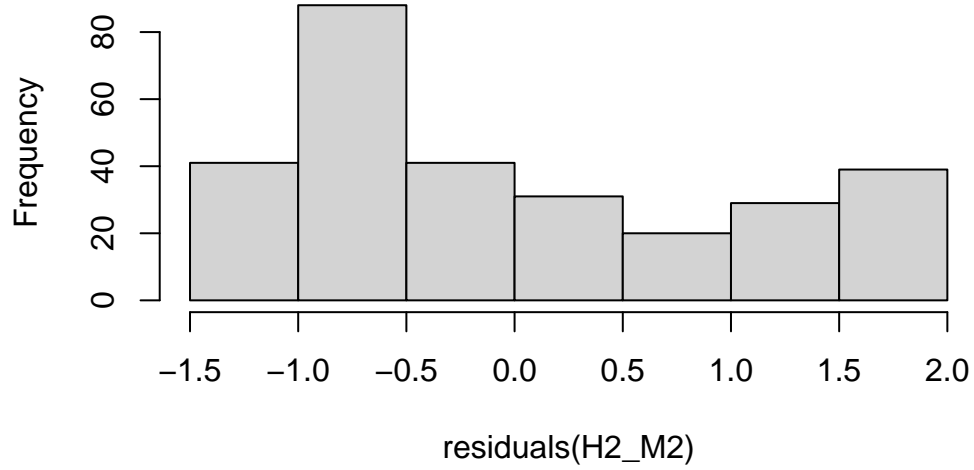
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

```
# Histogram for normality of residuals  
hist(residuals(H2_M1)) # For model 1
```



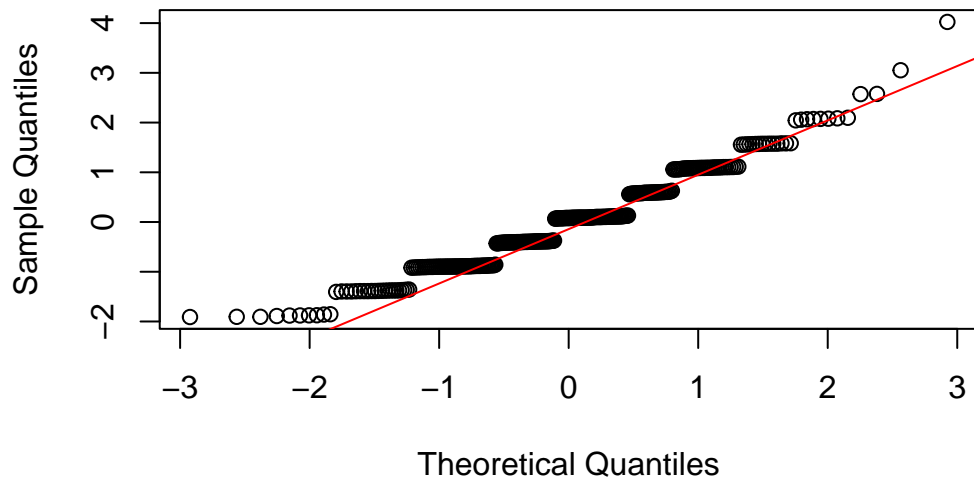
```
hist(residuals(H2_M2)) # For model 2
```

**Histogram of residuals(H2\_M2)**

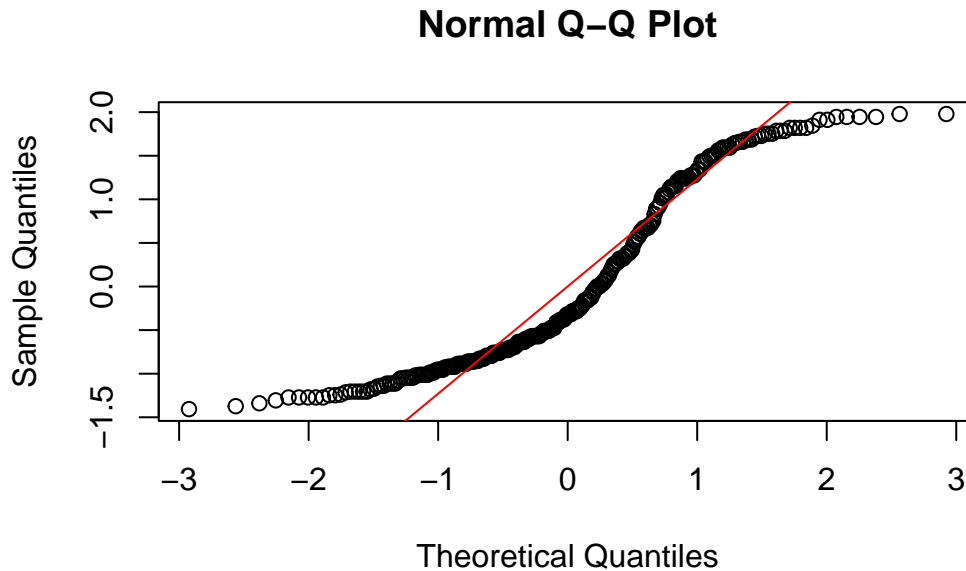


```
# QQ-plot for normality of residuals of Model 1
qqnorm(residuals(H2_M1))
qqline(residuals(H2_M1), col = "red")
```

**Normal Q-Q Plot**



```
# QQ-plot for normality of residuals of Model 2
qqnorm(residuals(H2_M2))
qqline(residuals(H2_M2), col = "red")
```



- **Model 1:**
  - The histogram and QQ-plot suggest that the residuals are relatively normally distributed for Model 1.
- **Model 2:**
  - The histogram and QQ-plot of the residuals of Model 2 indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 or 1.5 SD. The two clusters of similar errors in prediction may be due to sex differences in religiosity. In exploratory analyses, we will add **sex** as a control variable.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictor for Model 2.

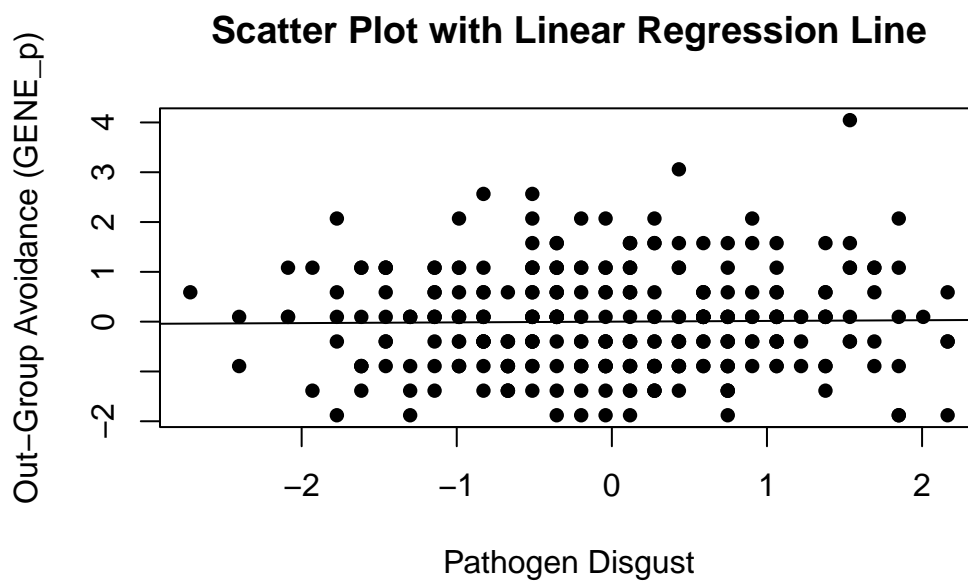
3. **Multicollinearity:** Not applicable because each model only has one predictor.

### Summarizing the Models



## Model 1

```
# Create a scatterplot with the regression line to visualize
plot(data$TDDS_p_z, data$GENE_p_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Pathogen Disgust",
     ylab = "Out-Group Avoidance (GENE_p)",
     pch = 16, # For solid circle points
     col = "black")
# Add regression line
abline(H2_M1, col = "black", lwd = 1)
```



```
# Summarizing the model
summary(H2_M1)
```

Call:

```
lm(formula = GENE_p_z ~ TDDS_p_z, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.9107	-0.8800	0.0804	0.5946	4.0239

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.593e-16	5.892e-02	0.000	1.000
TDDS_p_z	1.440e-02	5.902e-02	0.244	0.807

Residual standard error: 1.002 on 287 degrees of freedom

Multiple R-squared: 0.0002075, Adjusted R-squared: -0.003276

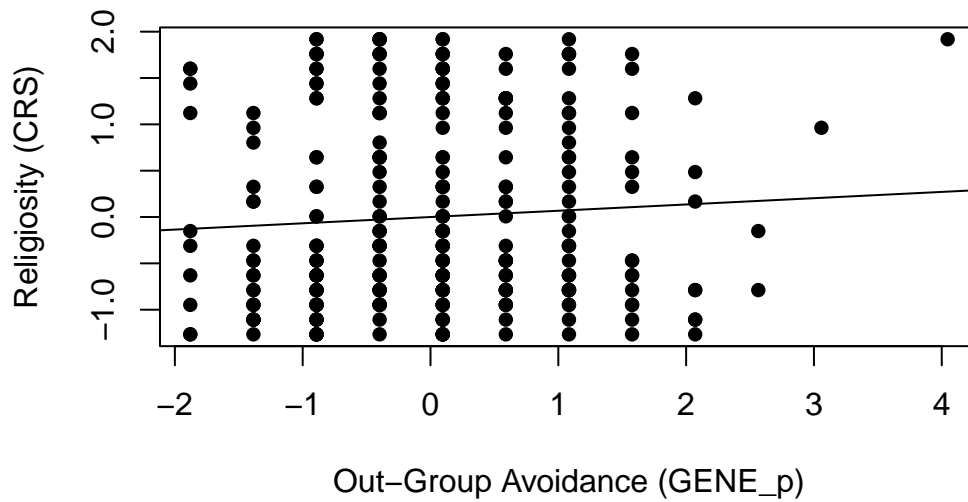
F-statistic: 0.05956 on 1 and 287 DF, p-value: 0.8074

- We can see from the Scatter Plot that there is no relationship between pathogen disgust and out-group avoidance to speak of ( $B = .014$ ,  $t(287) = .244$ ,  $p = .807$ ), just as with the correlation above.
  - Although this precludes the possibility that out-group avoidance mediates the relationship between disgust and religiosity, we will continue with the analysis for completions sake.

## Model 2

```
# Create a scatterplot with the regression line to visualize
plot(data$GENE_p_z, data$CRS_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Out-Group Avoidance (GENE_p)",
     ylab = "Religiosity (CRS)",
     pch = 16, # For solid circle points
     col = "black")
# Add regression line
abline(H2_M2, col = "black", lwd = 1)
```

## Scatter Plot with Linear Regression Line



```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_z <- as.numeric(data$GENE_p_z)
CRS_z <- as.numeric(data$CRS_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_M2 <- Boot(H2_M2, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H2_M2) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
      original    bootBias  bootSE  bootMed
(Intercept) 1.5852e-17  0.00034356 0.059932 0.0020055
GENE_p_z    6.7778e-02 -0.00134646 0.062326 0.0655162
```

```
confint(boot_H2_M2) # Getting the BCIs
```

```
Bootstrap bca confidence intervals
```

```

                2.5 %    97.5 %
(Intercept) -0.11444697 0.1182643
GENE_p_z    -0.04812802 0.1936607

```

```

# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_z, CRS_z)

```

- We can see from the Scatter Plot that there is a positive trend in the relationship between out-group avoidance and religiosity.
- Based on the estimates, the standardized coefficient for out-group avoidance is around .068.
  - However, this relationship is not significant, as the BCI contains zero (upper = .19, lower = -.04).

## Full Mediation Model

The fact that pathogen disgust does not predict out-group avoidance and, in turn, out-group avoidance does not predict religiosity indicates that out-group avoidance could not mediate the relationship between pathogen disgust and religiosity, but I will run the analysis that we planned anyway just to have the results.

```

# Fitting the full mediation model
H2_full <- lm(CRS_z ~ TDDS_p_z + GENE_p_z, data = data)

```

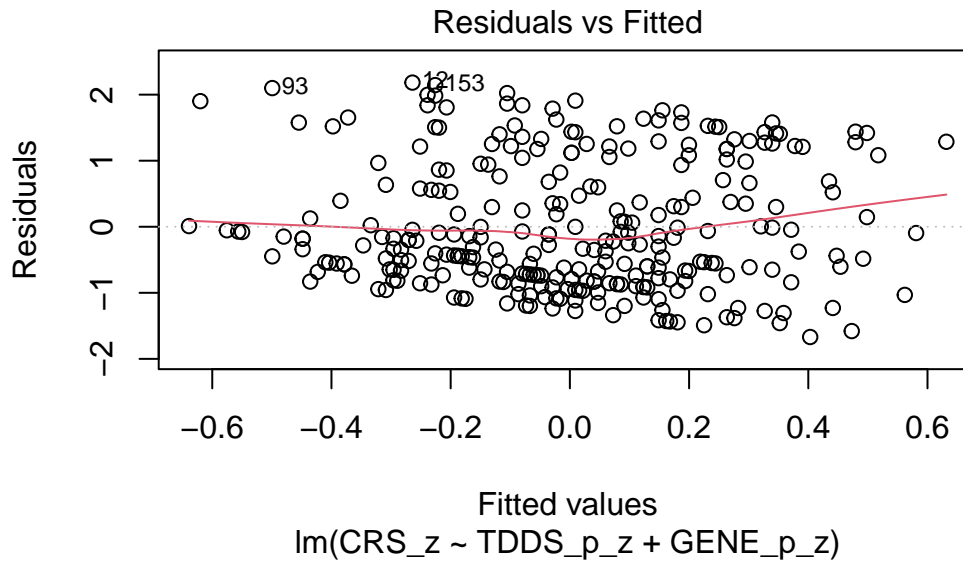
## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```

# Plotting the residual vs. fitted values
plot(H2_full, which = 1)

```

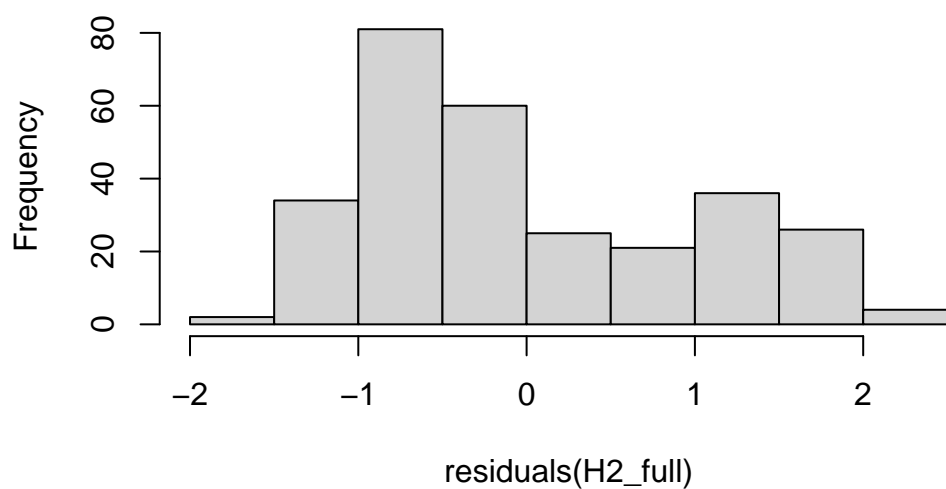


- There is a trend towards more negative residual values at the higher fitted values of the model. This may be problematic.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

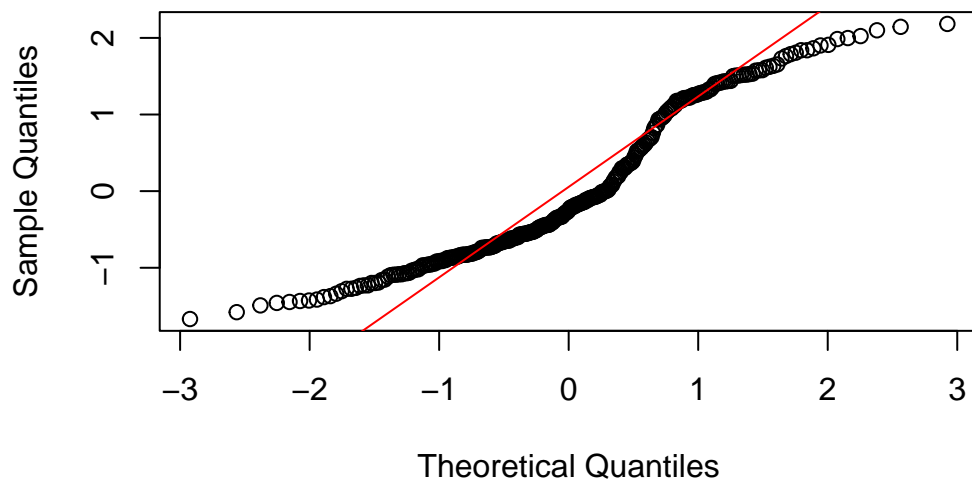
```
# Histogram for normality of residuals
hist(residuals(H2_full))
```

**Histogram of residuals(H2\_full)**



```
# QQ-plot for normality of residuals of the full model  
qqnorm(residuals(H2_full))  
qqline(residuals(H2_full), col = "red")
```

**Normal Q-Q Plot**



- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 SD. The two clusters of similar errors in prediction may be due to sex differences in religiosity. In exploratory analyses, we will add **sex** as a control variable.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictors for the model.

### 3. **Multicollinearity:** Variance Inflation Factor (VIF) from **car** package

```
# Calculating VIF
vif(H2_full)
```

```
TDDS_p_z GENE_p_z
1.000208 1.000208
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_z <- as.numeric(data$GENE_p_z)
CRS_z <- as.numeric(data$CRS_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_full <- Boot(H2_full, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H2_full) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original    bootBias   bootSE   bootMed
(Intercept) 6.4410e-18 -0.00036098 0.058493 -0.0014392
TDDS_p_z    2.4206e-01 -0.00012574 0.059348  0.2420111
GENE_p_z    6.4291e-02 -0.00114233 0.059533  0.0630945
```

```
confint(boot_H2_full) # Getting the BCIs
```

Bootstrap bca confidence intervals

```

                2.5 %    97.5 %
(Intercept) -0.10737353 0.1269200
TDDS_p_z     0.12066998 0.3534012
GENE_p_z     -0.04731359 0.1858159

```

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_z, CRS_z, TDDS_p_z)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H2_mediate_results <- mediate(model.m = H2_M1,
                              model.y = H2_full,
                              treat = "TDDS_p_z",
                              mediator = "GENE_p_z",
                              boot = TRUE, # Bootstrap the CI for the indirect effect
                              sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results of the mediation analysis
summary(H2_mediate_results)
```

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.000926	-0.008892	0.01	0.81
ADE	0.242061	0.132686	0.36	<2e-16 ***
Total Effect	0.242987	0.133597	0.36	<2e-16 ***
Prop. Mediated	0.003811	-0.039966	0.06	0.81

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289



Simulations: 1000

- **Multiple Regression Results:**

- Based on the bootstrapped CIs, pathogen disgust, but not out-group avoidance, is significantly positively related to religiosity.
- The coefficients for pathogen disgust (.24) and out-group avoidance (.06) are virtually identical in this model to the models above where they are stand-alone predictors. This suggests that the unique effect of pathogen disgust (and not out-group avoidance) on religiosity over-and-above each other are completely independent.

- **Mediation Analysis Results:**

- The ACME indicates that there is no significant indirect effect of pathogen disgust on religiosity through out-group avoidance ( $B = .001$ ,  $BCI = [-.009, .01]$ ).
- This result is inconsistent with both the weak and strong version version of Hypothesis 2. That is, there is no evidence here that the relationship between pathogen disgust and religiosity is mediated by out-group avoidance.

### Hypothesis 3: Sexual Strategies

#### Preliminary Simple Linear Regression Models

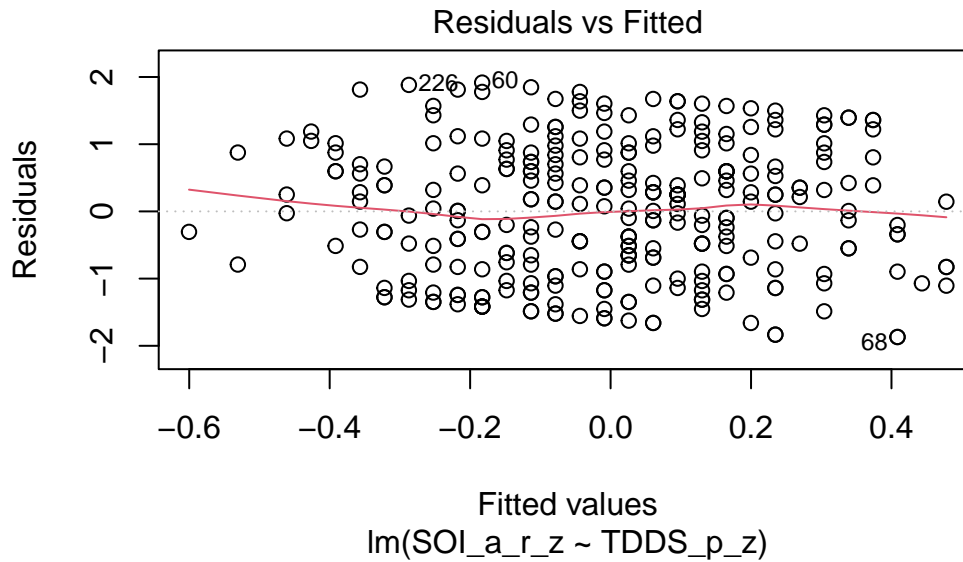
As mentioned above, to test Hypothesis 3, we must first establish that pathogen disgust affects one's mating strategy (i.e., make it more monogamous; as measured by more restricted sociosexuality; SOI\_a\_r\_z) and that a monogamous mating strategy influences religiosity.

```
# Fitting the models
H3_M1 <- lm(SOI_a_r_z ~ TDDS_p_z, data = data)
H3_M2 <- lm(CRS_z ~ SOI_a_r_z, data = data)
```

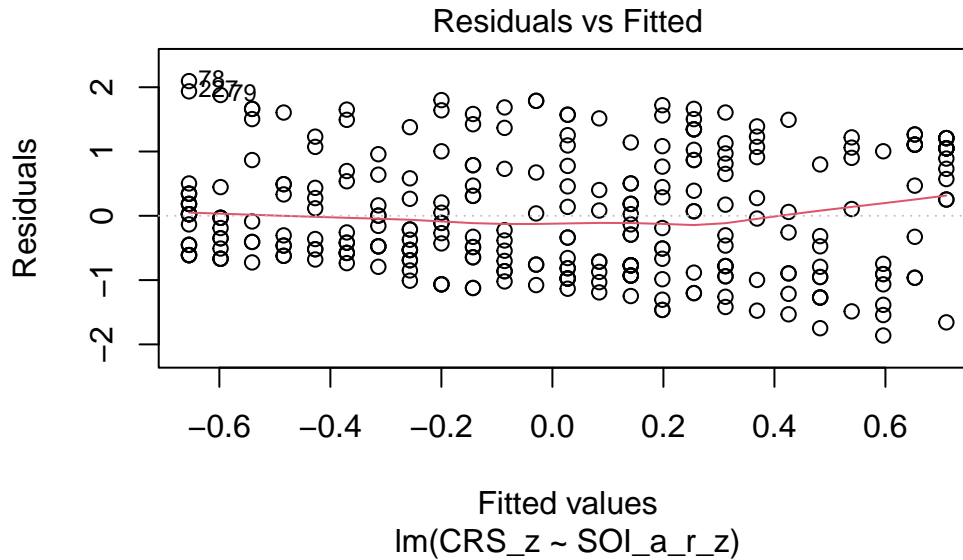
#### Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H3_M1, which = 1) # For model 1
```



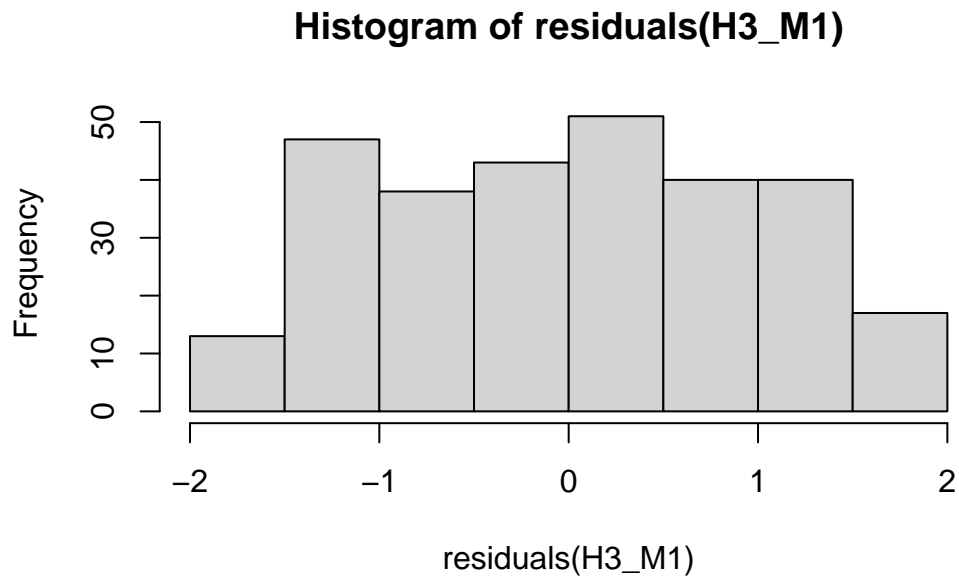
```
plot(H3_M2, which = 1) # For model 2
```



- The residuals vs. fitted plot does not stand out as violating the assumption of linearity or heteroscedasticity for the either model.

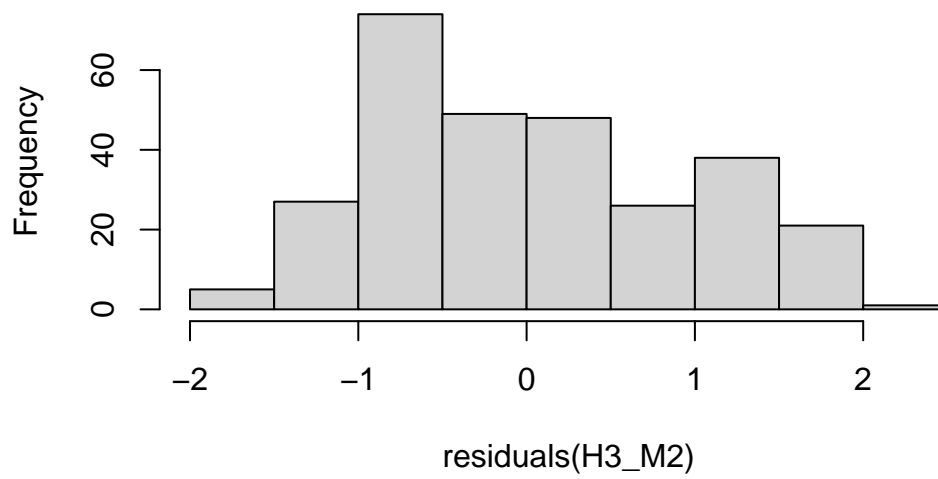
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

```
# Histogram for normality of residuals  
hist(residuals(H3_M1)) # For model 1
```



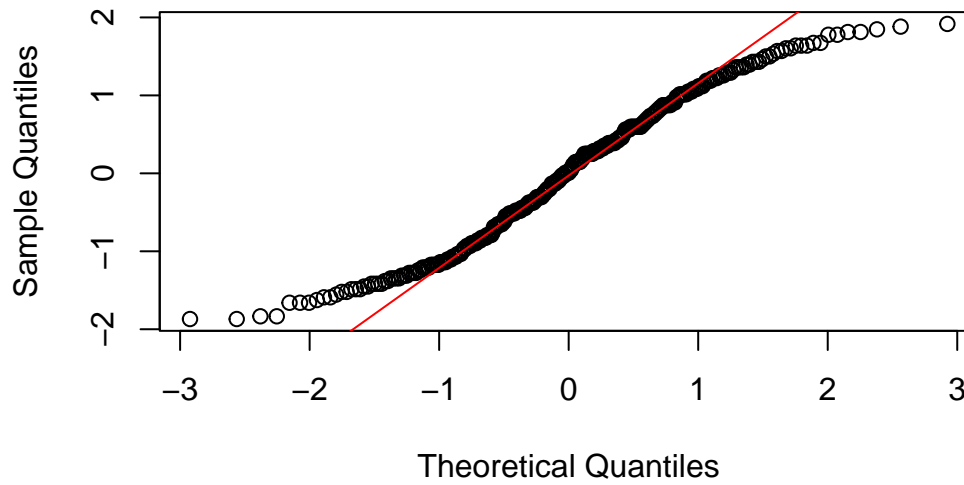
```
hist(residuals(H3_M2)) # For model 2
```

**Histogram of residuals(H3\_M2)**

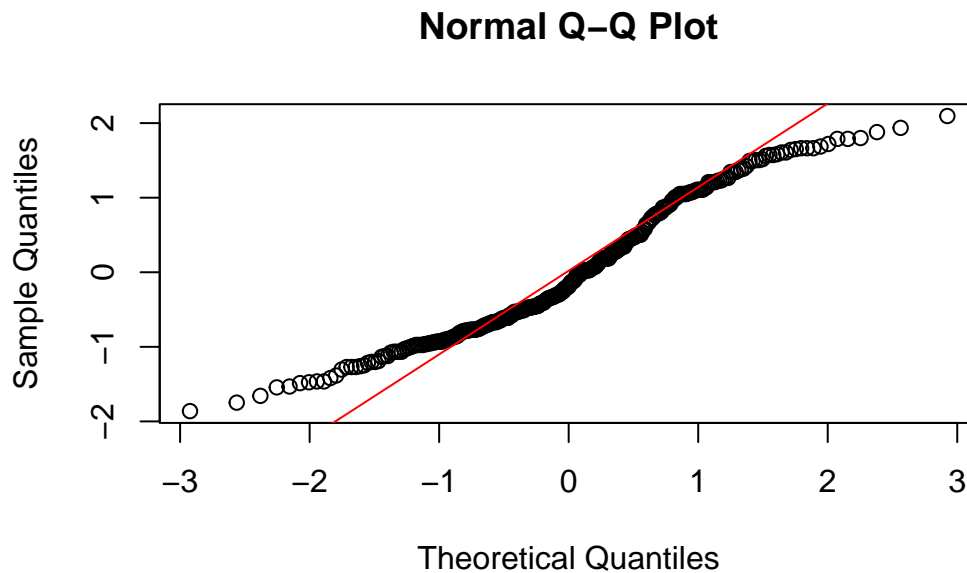


```
# QQ-plot for normality of residuals of Model 1  
qqnorm(residuals(H3_M1))  
qqline(residuals(H3_M1), col = "red")
```

**Normal Q-Q Plot**



```
# QQ-plot for normality of residuals of Model 2
qqnorm(residuals(H3_M2))
qqline(residuals(H3_M2), col = "red")
```



- **Both Models:**

- The QQ-plots for both models do not indicate that the residuals are normally distributed.
- Because of this failure of this assumption, we will conduct a bootstrapped test of the predictor for both models.

3. **Multicollinearity:** Not applicable because each model only has one predictor.

## Summarizing the Models

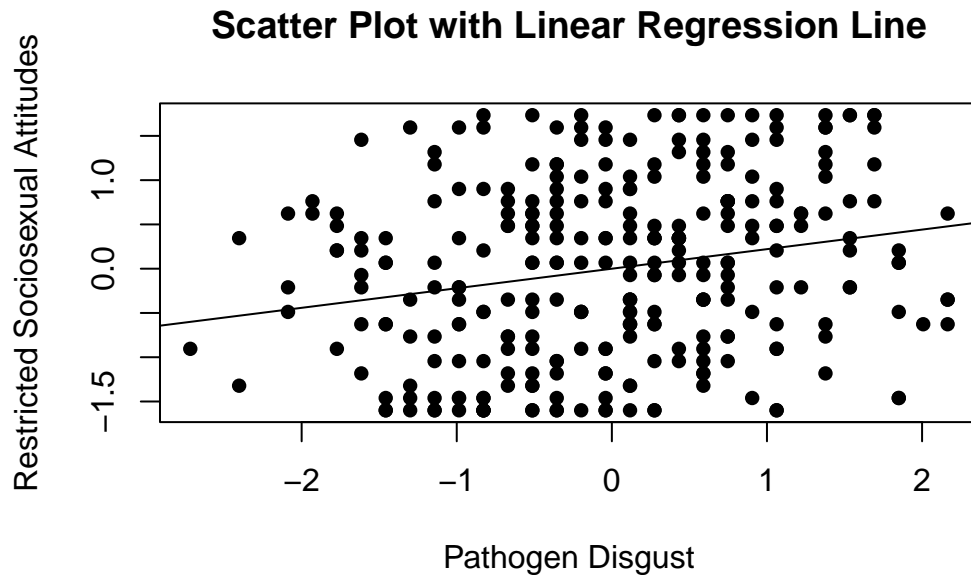
### Model 1

```
# Create a scatterplot with the regression line to visualize
plot(data$TDDS_p_z, data$SOI_a_r_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Pathogen Disgust",
     ylab = "Restricted Sociosexual Attitudes",
```

```

    pch = 16, # For solid circle points
    col = "black")
# Add regression line
abline(H3_M1, col = "black", lwd = 1)

```



```

# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_M1 <- Boot(H3_M1, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H3_M1) # Getting the coefficients

```

```

Number of bootstrap replications R = 1000
      original bootBias  bootSE  bootMed
(Intercept) -1.1616e-17 0.0012152 0.059909 0.0021975

```

```
TDDS_p_z      2.2087e-01 0.0015469 0.059225 0.2236528
```

```
confint(boot_H3_M1) # Getting the BCIs
```

Bootstrap bca confidence intervals

```
          2.5 %    97.5 %  
(Intercept) -0.1256230 0.1076692  
TDDS_p_z      0.1039745 0.3299802
```

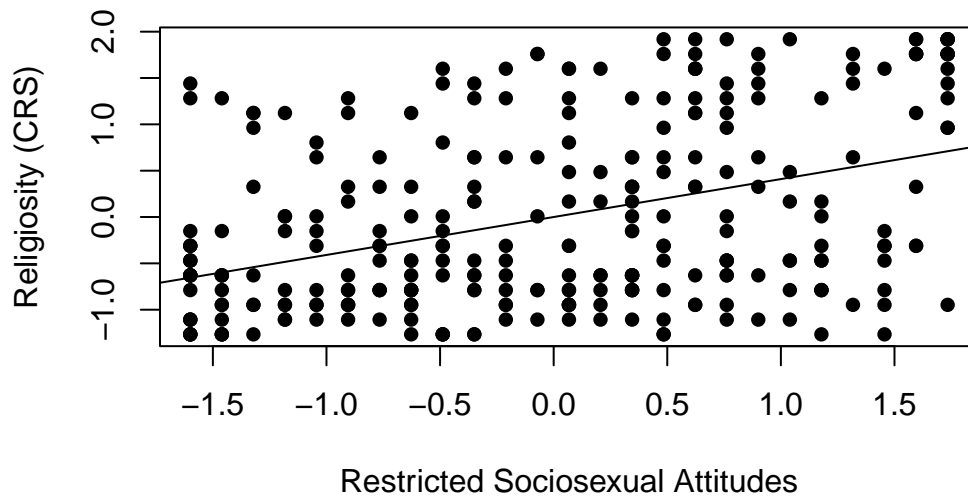
```
# Removing the vectors in the environment that the Boot package required outside of the data  
rm(SOI_a_r_z, TDDS_p_z)
```

- We can see from the Scatter Plot that there is a positive relationship between pathogen disgust and restricted sociosexual attitudes.
- Based on the estimates, the standardized coefficient for pathogen disgust is around .22.
  - In addition, this relationship is significant, as the BCI excludes zero (upper = .33, lower = .10).

## Model 2

```
# Create a scatterplot with the regression line to visualize  
plot(data$SOI_a_r_z, data$CRS_z,  
      main = "Scatter Plot with Linear Regression Line",  
      xlab = "Restricted Sociosexual Attitudes",  
      ylab = "Religiosity (CRS)",  
      pch = 16, # For solid circle points  
      col = "black")  
# Add regression line  
abline(H3_M2, col = "black", lwd = 1)
```

## Scatter Plot with Linear Regression Line



```
# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
CRS_z <- as.numeric(data$CRS_z)
```

```
# Setting the seed
set.seed(1042557)
```

```
# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_M2 <- Boot(H3_M2, R = 1000) # 1,000 resamples
```

```
# Summarizing the estimates and CIs
summary(boot_H3_M2) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
      original    bootBias  bootSE  bootMed
(Intercept) 2.7721e-17 -0.00099204 0.054590 -0.00413
SOI_a_r_z   4.0934e-01  0.00020098 0.055098  0.41078
```

```
confint(boot_H3_M2) # Getting the BCIs
```

```
Bootstrap bca confidence intervals
```



```

                2.5 %    97.5 %
(Intercept) -0.09552607 0.1229143
SOI_a_r_z    0.29501332 0.5104241

```

```

# Removing the vectors in the environment that the Boot package required outside of the data
rm(SOI_a_r_z, CRS_z)

```

- We can see from the Scatter Plot that there is a positive trend in the relationship between restricted sociosexual attitudes and religiosity.
- Based on the estimates, the standardized coefficient for restricted sociosexual attitudes is around .41.
  - In addition, this relationship is significant, as the BCI excludes zero (upper = .51, lower = .30).

## Full Mediation Model

The fact that pathogen disgust predicts a monogamous mating strategy and, in turn, a monogamous mating strategy predicts religiosity indicates that a monogamous mating strategy could potentially mediate the relationship between pathogen disgust and religiosity. Now we will run the mediation analysis.

```

# Fitting the full mediation model
H3_full <- lm(CRS_z ~ TDDS_p_z + SOI_a_r_z, data = data)

```

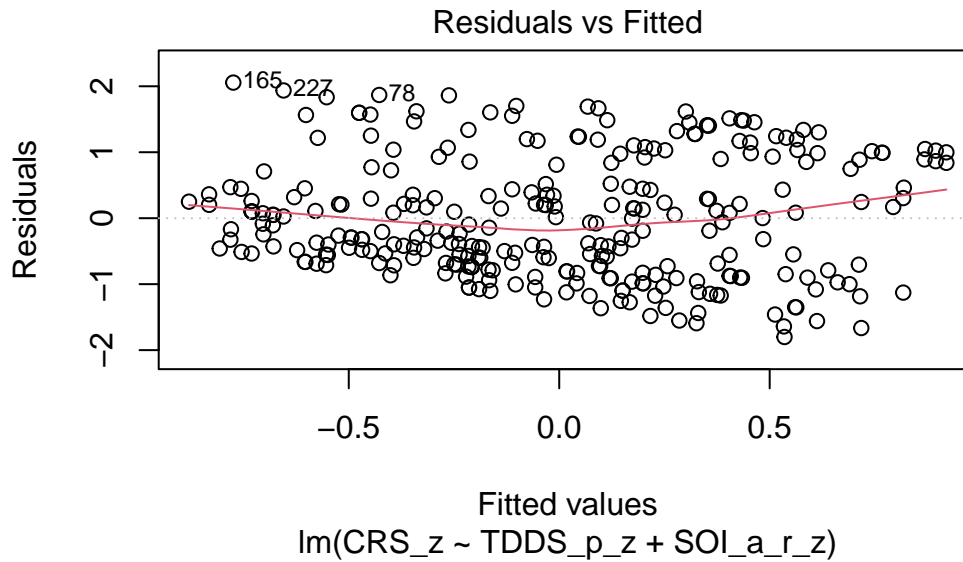
## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```

# Plotting the residual vs. fitted values
plot(H3_full, which = 1)

```

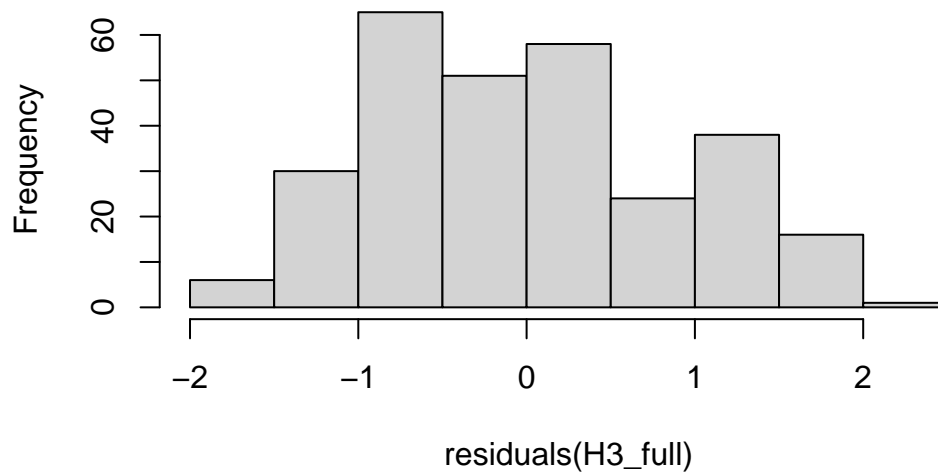


- There is a trend towards more negative residual values at the higher fitted values of the model. This may be problematic.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

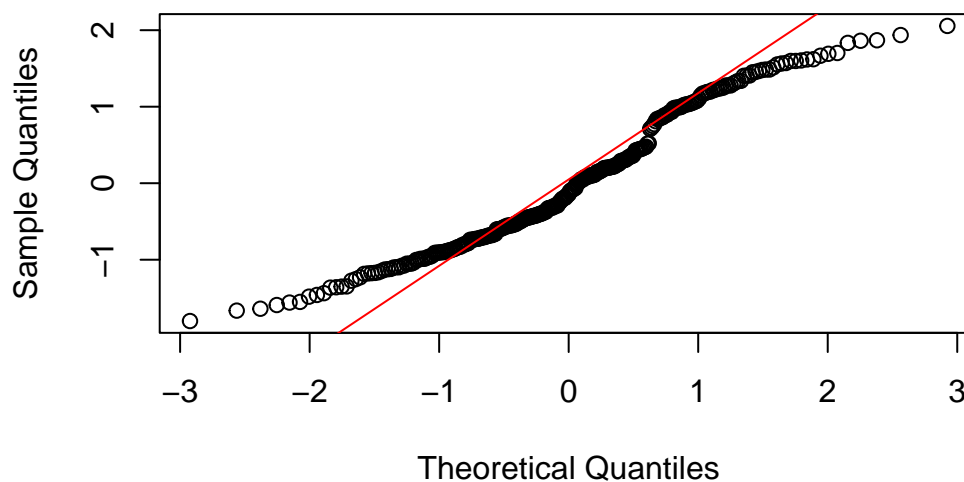
```
# Histogram for normality of residuals
hist(residuals(H3_full))
```

**Histogram of residuals(H3\_full)**



```
# QQ-plot for normality of residuals of the full model
qqnorm(residuals(H3_full))
qqline(residuals(H3_full), col = "red")
```

**Normal Q-Q Plot**



- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 SD. The two clusters of similar errors in prediction may be due to sex differences in religiosity. In exploratory analyses, we will add **sex** as a control variable.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictors for the model.

### 3. Multicollinearity: Variance Inflation Factor (VIF) from **car** package

```
# Calculating VIF
vif(H3_full)
```

```
TDDS_p_z SOI_a_r_z
1.051286  1.051286
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
CRS_z <- as.numeric(data$CRS_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_full <- Boot(H3_full, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H3_full) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original    bootBias   bootSE   bootMed
(Intercept) 2.1027e-17 -0.00136655 0.054173 -0.0038041
TDDS_p_z    1.6040e-01 -0.00045433 0.056386  0.1585189
SOI_a_r_z    3.7391e-01  0.00013298 0.057102  0.3743473
```

```
confint(boot_H3_full) # Getting the BCIs
```

Bootstrap bca confidence intervals

```

                2.5 %    97.5 %
(Intercept) -0.09343557 0.1242681
TDDS_p_z     0.05467911 0.2719632
SOI_a_r_z    0.25544697 0.4836007

```

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(SOI_a_r_z, CRS_z, TDDS_p_z)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H3_mediate_results <- mediate(model.m = H3_M1,
                              model.y = H3_full,
                              treat = "TDDS_p_z",
                              mediator = "SOI_a_r_z",
                              boot = TRUE, # Bootstrap the CI for the indirect effect
                              sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results of the mediation analysis
summary(H3_mediate_results)
```

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.0826	0.0384	0.13	<2e-16 ***
ADE	0.1604	0.0563	0.27	0.002 **
Total Effect	0.2430	0.1336	0.36	<2e-16 ***
Prop. Mediated	0.3399	0.1636	0.63	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- **Multiple Regression Results:**

- Based on the bootstrapped CIs, both pathogen disgust and restricted sociosexual attitudes are significantly positively related to religiosity, with the unique effect of restricted sociosexual attitudes being substantially larger ( $B = .374$ ,  $BCI = [.255, .484]$ ) than the unique effect of pathogen disgust ( $B = .160$ ,  $BCI = [.055, .272]$ ).

- **Mediation Analysis Results:**

- The ACME indicates that there is a significant indirect effect of pathogen disgust on religiosity through a monogamous mating strategy ( $B = .08$ ,  $BCI = [.038, .13]$ ).
  - \* Given this result, we have evidence for the weak version of Hypothesis 3. That is, the relationship between pathogen disgust and religiosity is partially mediated by a monogamous mating strategy.
- The proportion of the total effect mediated (the ratio of the indirect effect to the direct effect) is .34. That is, 34% of the covariance between pathogen disgust and religiosity seems to be able to be accounted for by restricted sociosexual attitudes. Our inferential criterion for concluding full mediation was that this ratio would be at least .80, with the total effect greater than a  $B = .2$ . Even the upper bound of the bootstrapped confidence interval for this proportion mediated does not cross this threshold ( $BCI = [.164, .63]$ ).
  - \* Given this result, we do not have evidence for the strong version of Hypothesis 3. That is, the relationship between pathogen disgust and religiosity is not fully mediated by a monogamous mating strategy.

## Exploratory Analyses

### Hypothesis 1: Controlling for Sex

#### Preliminary Simple Linear Regression Models

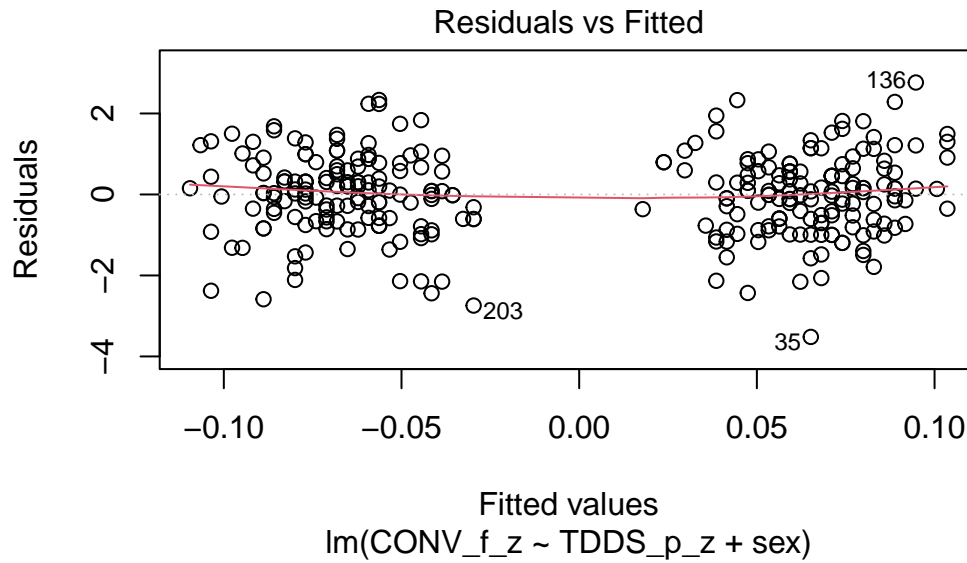
We will now re-test Hypothesis 1 while controlling for sex.

```
# Fitting the models
H1_M1_s <- lm(CONV_f_z ~ TDDS_p_z + sex, data = data)
H1_M2_s <- lm(CRS_z ~ CONV_f_z + sex, data = data)
```

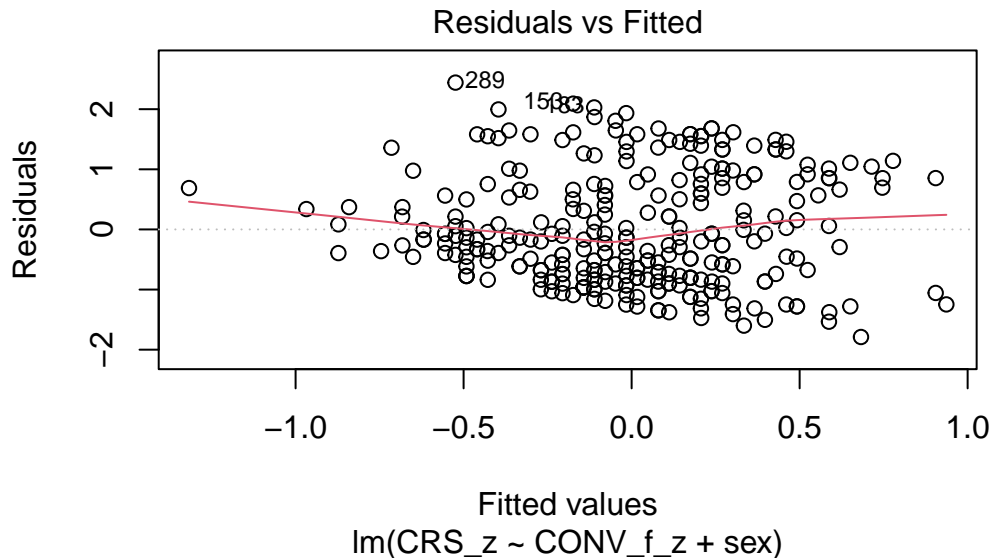
## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values  
plot(H1_M1_s, which = 1) # For model 1
```



```
plot(H1_M2_s, which = 1) # For model 2
```



- The residuals vs. fitted plot for the first model is quite strange, and it suggests that there are strongly different predictions for males and females. I will differentiate the sexes by making their points different colors, to see if this is the case.
- The residuals vs. fitted plot for the second model has a pretty clear trend towards more negative residuals at higher fitted values. We should conduct bootstrapped resampling inferences for these models.

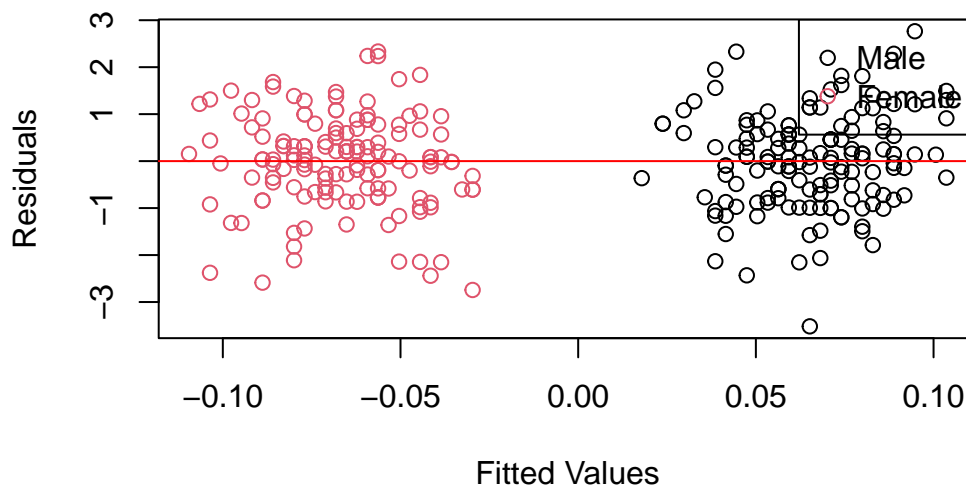
```
# Extract residuals and fitted values for model 1
residuals_m1 <- residuals(H1_M1_s)
fitted_m1 <- fitted(H1_M1_s)

# Plot residuals vs fitted values for model 1 with different colors for sex
plot(fitted_m1, residuals_m1, col = as.numeric(data$sex),
     main = "Residuals vs Fitted Values (Model 1)",
     xlab = "Fitted Values", ylab = "Residuals")
abline(h = 0, col = "red")

# Add a legend to the plot
legend("topright", legend = levels(data$sex), col = 1:length(levels(data$sex)), pch = 1)
```



## Residuals vs Fitted Values (Model 1)



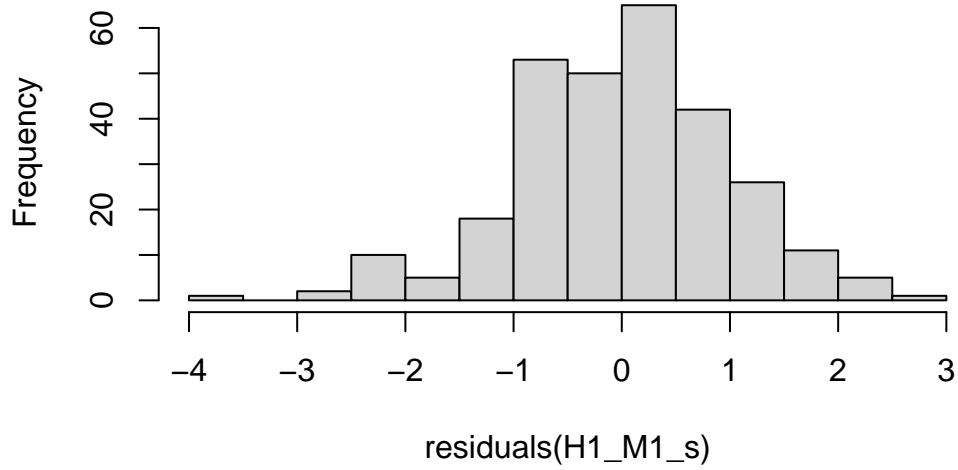
```
# Drop the residuals and fitted values from the environment  
rm(residuals_m1, fitted_m1)
```

- Indeed, the different fitted values represent different predictions for each sex.

### 2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

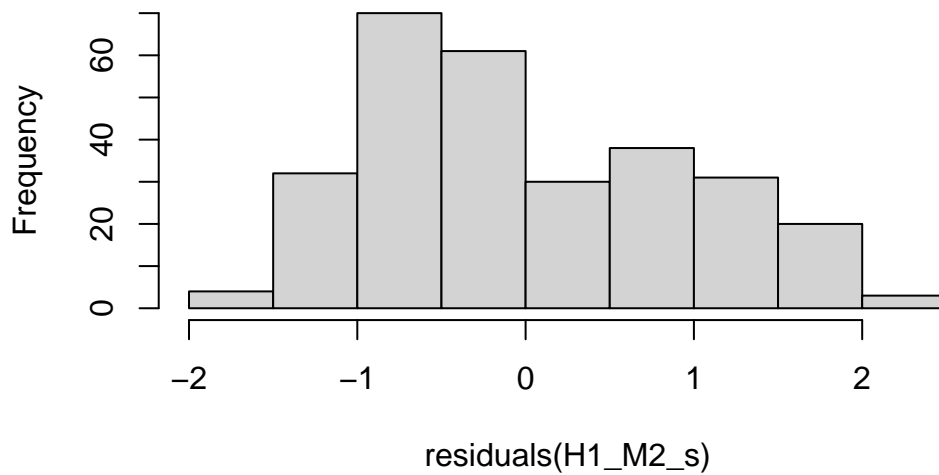
```
# Histogram for normality of residuals  
hist(residuals(H1_M1_s)) # For model 1
```

**Histogram of residuals(H1\_M1\_s)**



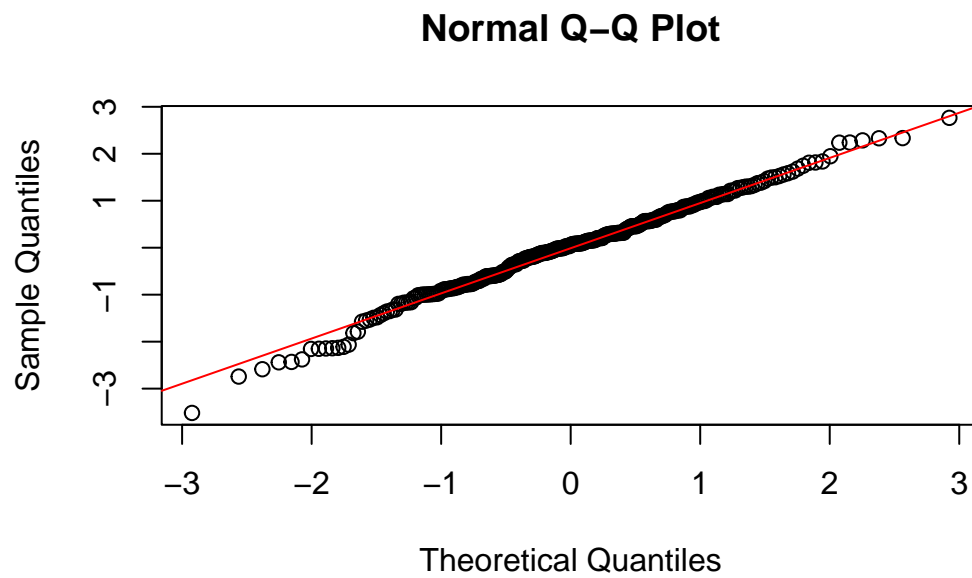
```
hist(residuals(H1_M2_s)) # For model 2
```

**Histogram of residuals(H1\_M2\_s)**

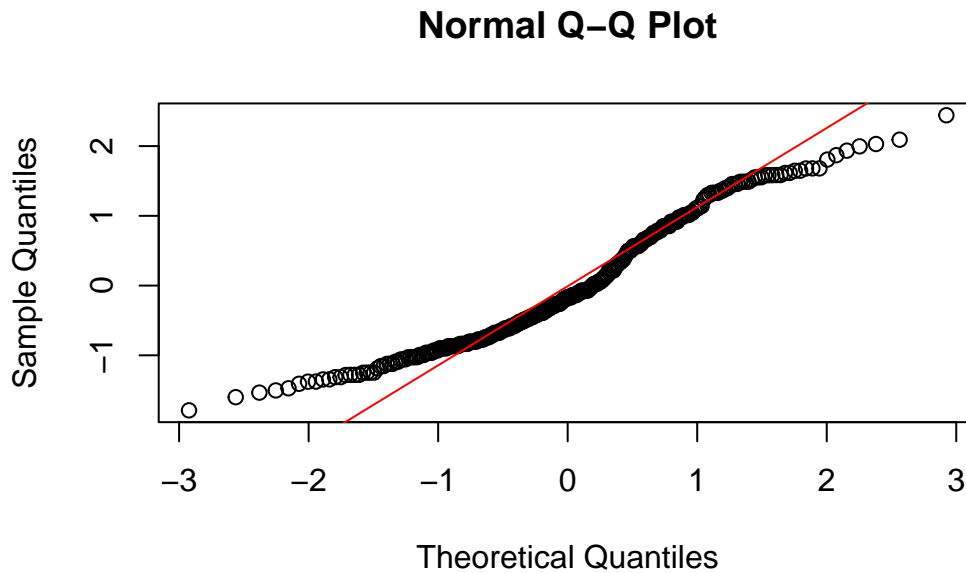


```
# QQ-plot for normality of residuals of Model 1  
qqnorm(residuals(H1_M1_s))
```

```
qqline(residuals(H1_M1_s), col = "red")
```



```
# QQ-plot for normality of residuals of Model 2  
qqnorm(residuals(H1_M2_s))  
qqline(residuals(H1_M2_s), col = "red")
```



- **Model 1:**

- The histogram and QQ-plot suggest that the residuals are normally distributed for Model 1.

- **Model 2:**

- The histogram and QQ-plot of the residuals of Model 2 indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just below 1 SD.

- We will conduct a bootstrapped tests of the predictors for both models.

### 3. **Multicollinearity:** Variance inflation factor (VIF) for both models

```
# Calculating the VIF for both models
vif(H1_M1_s)
```

```
TDDS_p_z      sex
1.038981 1.038981
```

```
vif(H1_M2_s)
```

```
CONV_f_z      sex
1.004388 1.004388
```

- The VIF values are good here.

## Summarizing the Models

### Model 1

```
# Adding the variables to the broader environment so the Boot function will run
CONV_f_z <- as.numeric(data$CONV_f_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)
sex <- data$sex

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H1_M1_s <- Boot(H1_M1_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H1_M1_s) # Getting the coefficients
```

Number of bootstrap replications R = 1000

	original	bootBias	bootSE	bootMed
(Intercept)	0.068890	9.5639e-04	0.083729	0.071961
TDDS_p_z	0.018766	-8.7770e-04	0.067435	0.017903
sexFemale	-0.139225	3.8768e-06	0.118828	-0.140871

```
confint(boot_H1_M1_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.1076349	0.22374169
TDDS_p_z	-0.1137392	0.15325593
sexFemale	-0.3708032	0.09485534

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(CONV_f_z, TDDS_p_z, sex)
```

- Based on the estimate and BCI for pathogen disgust ( $B = .018$ ,  $BCI = [-.114, .153]$ ), controlling for sex does not change whether there is a positive relationship between pathogen disgust and traditionalism.

- Although this precludes the possibility that adherence to tradition mediates the relationship between disgust and religiosity, we will continue with the analysis for completions sake.

## Model 2

```
# Adding the variables to the broader environment so the Boot function will run
CONV_f_z <- as.numeric(data$CONV_f_z)
CRS_z <- as.numeric(data$CRS_z)
sex <- data$sex

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H1_M2_s <- Boot(H1_M2_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H1_M2_s) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
```

	original	bootBias	bootSE	bootMed
(Intercept)	-0.18812	-0.00399574	0.074103	-0.19341
CONV_f_z	0.32702	0.00040075	0.053030	0.32711
sexFemale	0.38019	0.00631972	0.113040	0.38812

```
confint(boot_H1_M2_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.3170446	-0.02384113
CONV_f_z	0.2199292	0.42953209
sexFemale	0.1614658	0.60697545

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(CONV_f_z, CRS_z, sex)
```

- Based on the estimates, traditionalism is significantly related to religiosity independent of sex ( $B = .327$ ,  $BCI = [.22, .43]$ ), and the estimate is very similar to the previous estimate without controlling for sex, suggesting that the effects are primarily independent.

## Full Mediation Model

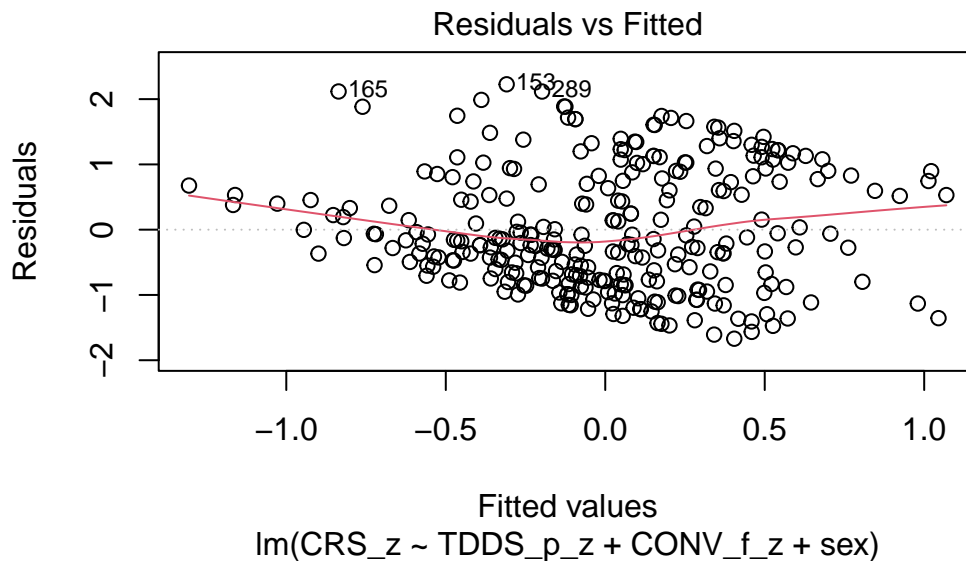
The fact that pathogen disgust does not predict traditionalism indicates that traditionalism could not mediate the relationship between pathogen disgust and religiosity, but I will run the analysis that we planned anyway just to have the results.

```
# Fitting the full mediation model
H1_full_s <- lm(CRS_z ~ TDDS_p_z + CONV_f_z + sex, data = data)
```

## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

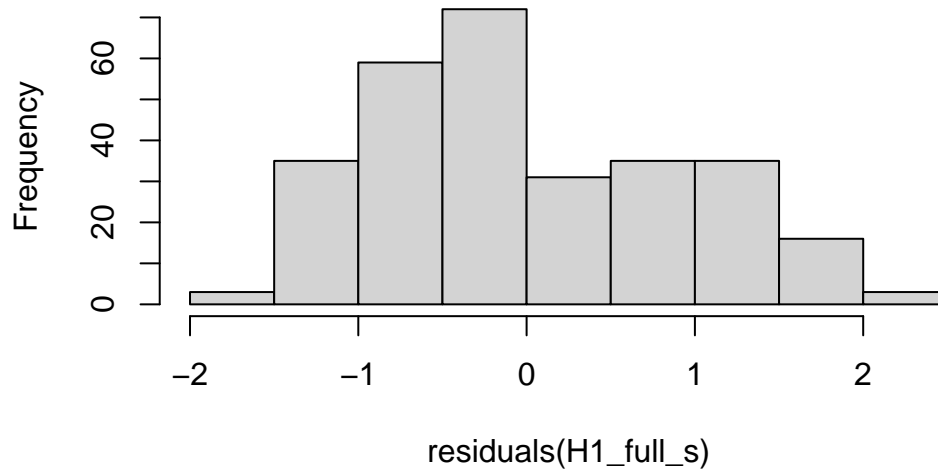
```
# Plotting the residual vs. fitted values
plot(H1_full_s, which = 1)
```



- There is a trend towards more negative residual values at the higher fitted values of the model. This would be problematic for parametric tests.
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

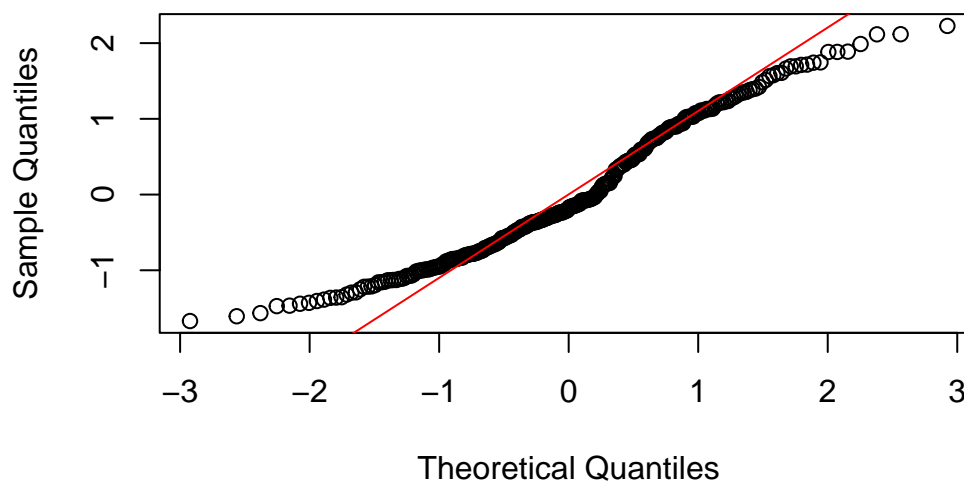
```
# Histogram for normality of residuals
hist(residuals(H1_full_s))
```

**Histogram of residuals(H1\_full\_s)**



```
# QQ-plot for normality of residuals of the full model
qqnorm(residuals(H1_full_s))
qqline(residuals(H1_full_s), col = "red")
```

**Normal Q-Q Plot**





- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a tendency for higher frequencies of residuals at just around 1 SD.
- We will conduct bootstrapping for confidence intervals to make inferences.

### 3. Multicollinearity: Variance Inflation Factor (VIF) from `car` package

```
# Calculating VIF
vif(H1_full_s)
```

```
TDDS_p_z CONV_f_z      sex
1.039334 1.004730 1.043866
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
CONV_f_z <- as.numeric(data$CONV_f_z)
CRS_z <- as.numeric(data$CRS_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)
sex <- data$sex

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H1_full_s <- Boot(H1_full_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H1_full_s) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original  bootBias  bootSE  bootMed
(Intercept) -0.14722 -0.00473112 0.074568 -0.15385
TDDS_p_z      0.21242 -0.00125691 0.057959  0.21105
CONV_f_z      0.32316  0.00096636 0.049858  0.32491
sexFemale     0.29753  0.00669264 0.110568  0.30430
```

```
confint(boot_H1_full_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.27878504	0.02449477
TDDS_p_z	0.09861968	0.32435475
CONV_f_z	0.21945136	0.41880569
sexFemale	0.07606250	0.50767084

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(CONV_f_z, CRS_z, TDDS_p_z, sex)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H1_s_mediate_results <- mediate(model.m = H1_M1_s,
  model.y = H1_full_s,
  treat = "TDDS_p_z",
  mediator = "CONV_f_z",
  covariates = "sex",
  boot = TRUE, # Bootstrap the CI for the indirect effect
  sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results for the mediation analysis
summary(H1_s_mediate_results)
```

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

(Inference Conditional on the Covariate Values Specified in `covariates`)

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.00606	-0.03225	0.05	0.76
ADE	0.21242	0.10577	0.33	<2e-16 ***
Total Effect	0.21849	0.10907	0.35	<2e-16 ***
Prop. Mediated	0.02776	-0.19673	0.23	0.76

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- **Multiple Regression Results:**

- Based on the bootstrapped CIs not containing zero, all three variables are significantly positively related to religiosity (with the dummy coding of sex being 1 = Female).

- **Mediation Analysis Results:**

- The ACME indicates that there is no significant indirect effect of pathogen disgust on religiosity through traditionalism ( $B = .006$ ,  $BCI = [-.032, .05]$ ).
- This result is inconsistent with both the weak and strong version version of Hypothesis 1. That is, there is no evidence here that the relationship between pathogen disgust and religiosity is mediated by traditionalism, even when controlling for sex.

## Hypothesis 2: Controlling for Sex

### Preliminary Simple Linear Regression Models

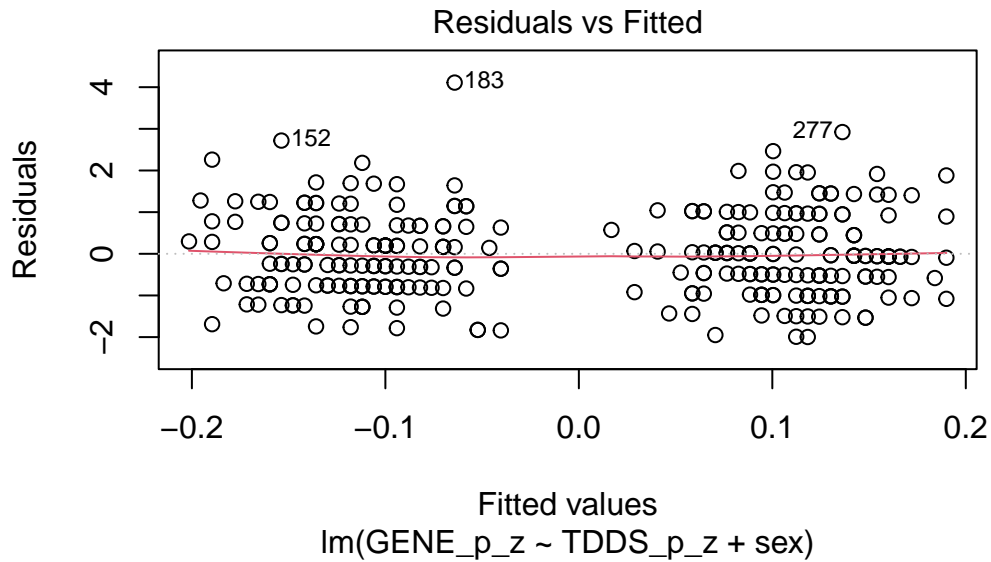
We will now re-test Hypothesis 2 while controlling for sex.

```
# Fitting the models
H2_M1_s <- lm(GENE_p_z ~ TDDS_p_z + sex, data = data)
H2_M2_s <- lm(CRS_z ~ GENE_p_z + sex, data = data)
```

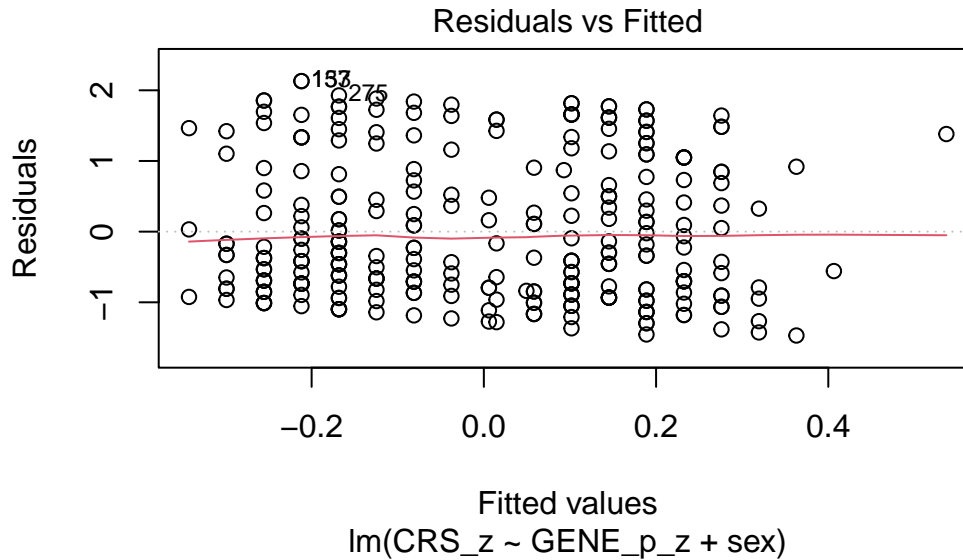
### Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H2_M1_s, which = 1) # For model 1
```



```
plot(H2_M2_s, which = 1) # For model 2
```



- The residuals vs. fitted plot for the first model seems to be partitioned by sex like the previous plot, and it suggests that there are strongly different predictions for

males and females. I will differentiate the sexes by making their points different colors, to see if this is the case.

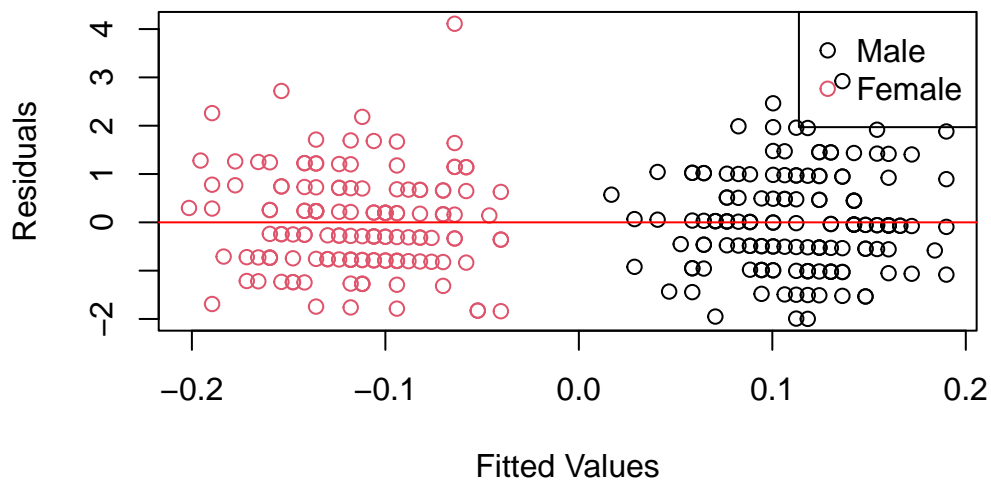
- The second plot looks quite good.

```
# Extract residuals and fitted values for model 1
residuals_m1 <- residuals(H2_M1_s)
fitted_m1 <- fitted(H2_M1_s)

# Plot residuals vs fitted values for model 1 with different colors for sex
plot(fitted_m1, residuals_m1, col = as.numeric(data$sex),
     main = "Residuals vs Fitted Values (Model 1)",
     xlab = "Fitted Values", ylab = "Residuals")
abline(h = 0, col = "red")

# Add a legend to the plot
legend("topright", legend = levels(data$sex), col = 1:length(levels(data$sex)), pch = 1)
```

### Residuals vs Fitted Values (Model 1)



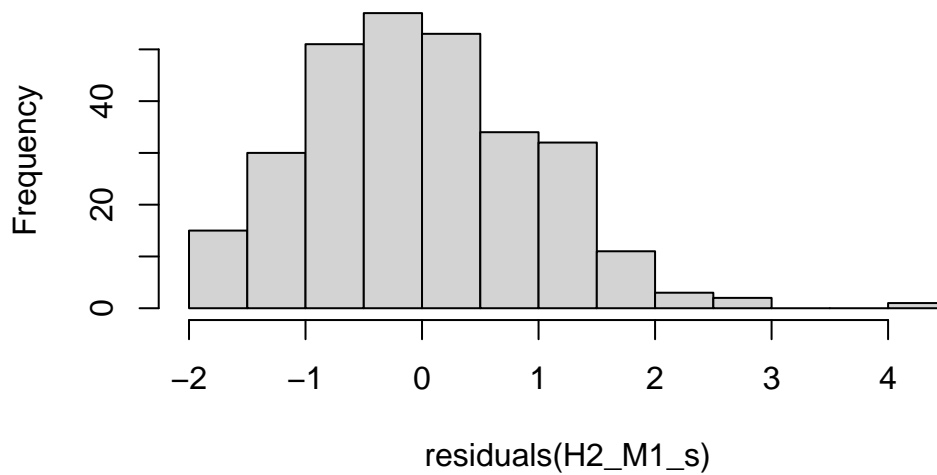
```
# Drop the residuals and fitted values from the environment
rm(residuals_m1, fitted_m1)
```

- Indeed, the different fitted values represent different predictions for each sex.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

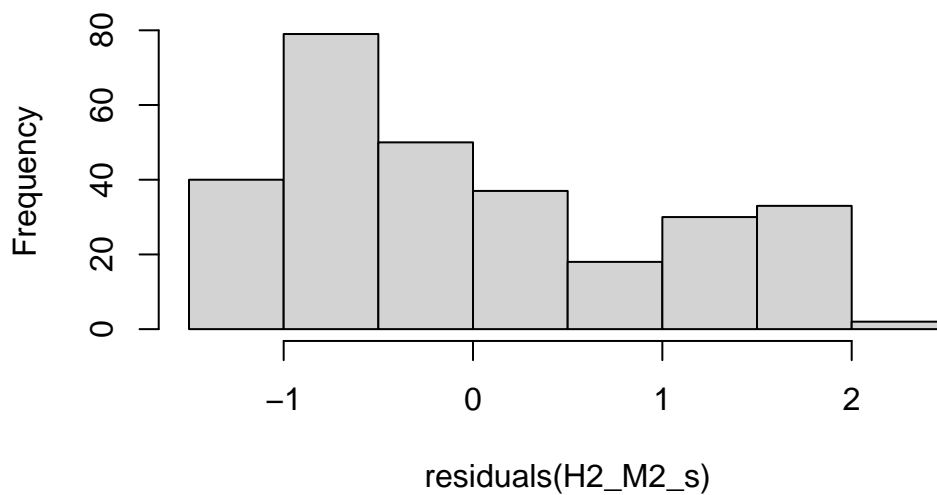
```
# Histogram for normality of residuals  
hist(residuals(H2_M1_s)) # For model 1
```

**Histogram of residuals(H2\_M1\_s)**



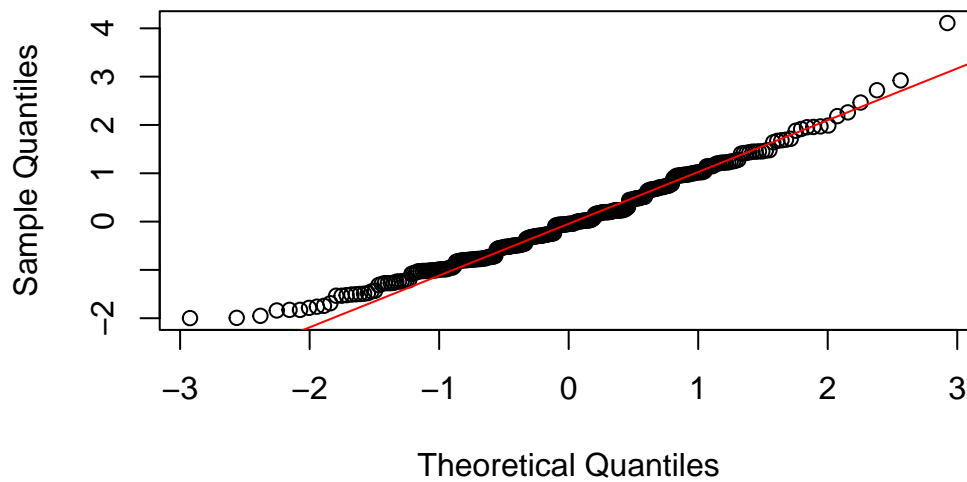
```
hist(residuals(H2_M2_s)) # For model 2
```

**Histogram of residuals(H2\_M2\_s)**

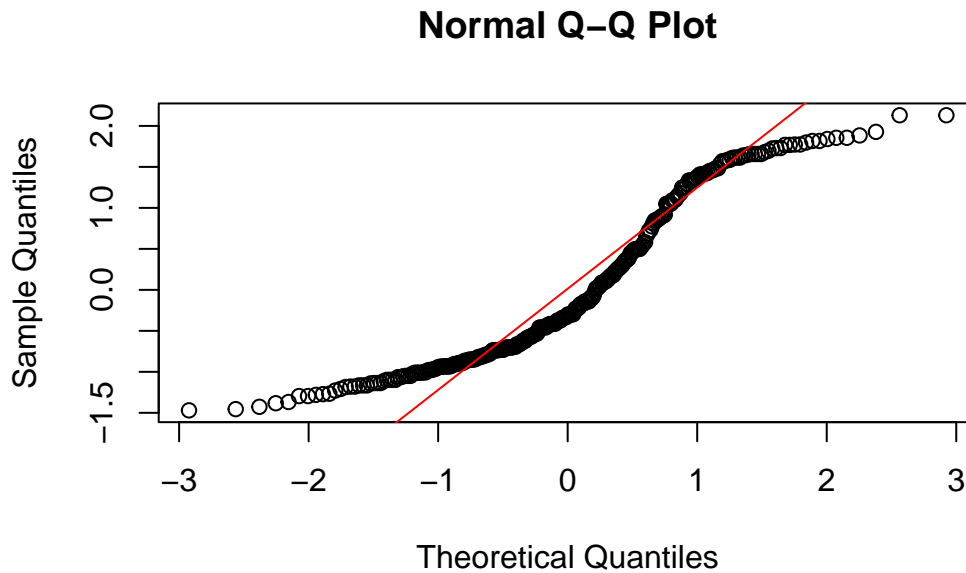


```
# QQ-plot for normality of residuals of Model 1
qqnorm(residuals(H2_M1_s))
qqline(residuals(H2_M1_s), col = "red")
```

**Normal Q-Q Plot**



```
# QQ-plot for normality of residuals of Model 2
qqnorm(residuals(H2_M2_s))
qqline(residuals(H2_M2_s), col = "red")
```



- **Model 1:**

- The histogram and QQ-plot suggest that the residuals are normally distributed for Model 1.

- **Model 2:**

- The histogram and QQ-plot of the residuals of Model 2 indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just below 2 SD.

- We will conduct a bootstrapped tests of the predictors for both models.

### 3. **Multicollinearity:** Variance inflation factor (VIF) for both models

```
# Calculating the VIF for both models
vif(H2_M1_s)
```

```
TDDS_p_z      sex
1.038981 1.038981
```

```
vif(H2_M2_s)
```

```
GENE_p_z      sex
1.013159 1.013159
```

- The VIF values are good here.



## Summarizing the Models

### Model 1

```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_z <- as.numeric(data$GENE_p_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)
sex <- data$sex

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_M1_s <- Boot(H2_M1_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H2_M1_s) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original    bootBias    bootSE    bootMed
(Intercept)  0.119846 -0.00150248  0.087176   0.118491
TDDS_p_z     0.037901 -0.00194585  0.061237   0.036581
sexFemale    -0.242207 -0.00069038  0.120490  -0.243567
```

```
confint(boot_H2_M1_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

```
              2.5 %    97.5 %
(Intercept) -0.04583828 0.29032402
TDDS_p_z     -0.08378796 0.15652933
sexFemale    -0.47809343 0.01058525
```

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_z, TDDS_p_z, sex)
```

- Based on the estimate and BCI for pathogen disgust ( $B = .038$ ,  $BCI = [-.084, .157]$ ), while controlling for sex there is not a relationship between pathogen disgust and out-group avoidance.

- Although this precludes the possibility that out-group avoidance mediates the relationship between disgust and religiosity, we will continue with the analysis for completions sake.

## Model 2

```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_z <- as.numeric(data$GENE_p_z)
CRS_z <- as.numeric(data$CRS_z)
sex <- data$sex

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_M2_s <- Boot(H2_M2_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H2_M2_s) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
```

	original	bootBias	bootSE	bootMed
(Intercept)	-0.17669	-0.00266419	0.078782	-0.183267
GENE_p_z	0.08816	-0.00091022	0.061117	0.085621
sexFemale	0.35710	0.00584804	0.118083	0.362284

```
confint(boot_H2_M2_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.30605613	0.0008240374
GENE_p_z	-0.02455483	0.2142069266
sexFemale	0.13007365	0.5932166829

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_z, CRS_z, sex)
```

- Based on the estimates, out-group avoidance is not significantly related to religiosity independent of sex ( $B = .088$ ,  $BCI = [-.025, .214]$ ).

## Full Mediation Model

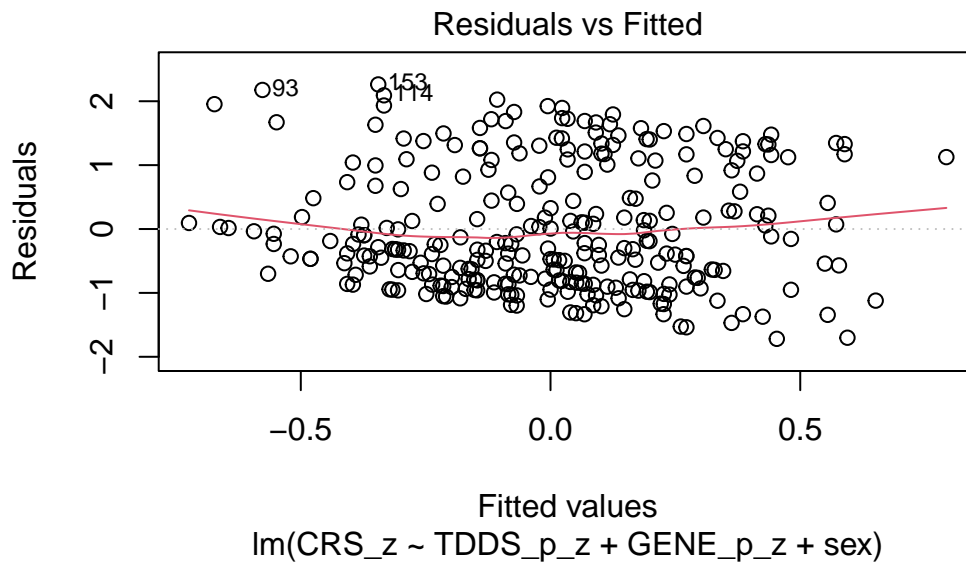
The fact that pathogen disgust does not predict out-group avoidance and that out-group avoidance does not predict religiosity indicates that traditionalism could not mediate the relationship between pathogen disgust and religiosity, but I will run the analysis that we planned anyway just to have the results.

```
# Fitting the full mediation model
H2_full_s <- lm(CRS_z ~ TDDS_p_z + GENE_p_z + sex, data = data)
```

## Assumptions

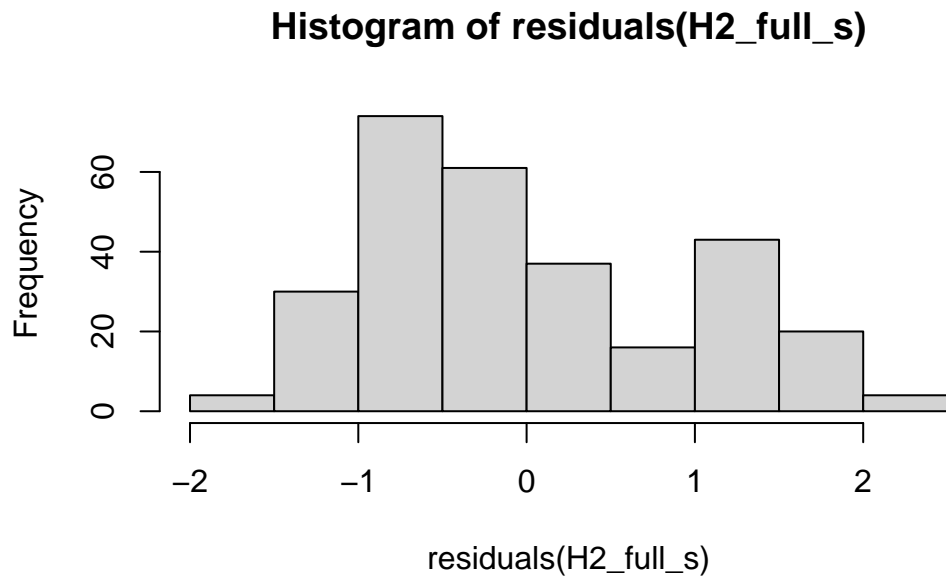
1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H2_full_s, which = 1)
```



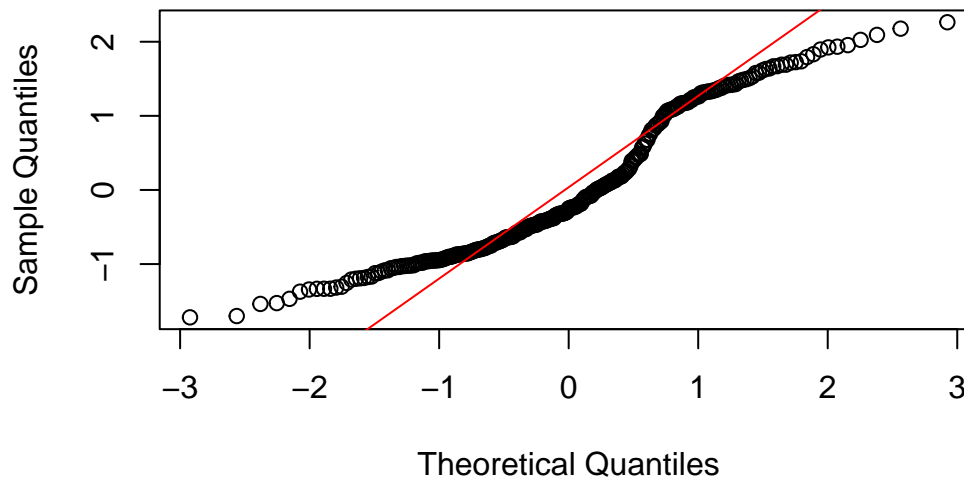
- There is a trend towards more negative residual values at the higher fitted values of the model. This would be problematic for parametric tests.
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

```
# Histogram for normality of residuals  
hist(residuals(H2_full_s))
```



```
# QQ-plot for normality of residuals of the full model  
qqnorm(residuals(H2_full_s))  
qqline(residuals(H2_full_s), col = "red")
```

## Normal Q-Q Plot



- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a tendency for higher frequencies of residuals at just above -1 SD and just above 1 SD.
- We will conduct bootstrapping for confidence intervals to make inferences.

### 3. Multicollinearity: Variance Inflation Factor (VIF) from `car` package

```
# Calculating VIF  
vif(H2_full_s)
```

```
TDDS_p_z  GENE_p_z      sex  
1.040438  1.014580  1.053911
```

- The VIF is good here.

## Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run  
GENE_p_z <- as.numeric(data$GENE_p_z)  
CRS_z <- as.numeric(data$CRS_z)  
TDDS_p_z <- as.numeric(data$TDDS_p_z)  
sex <- data$sex
```

```
# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_full_s <- Boot(H2_full_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H2_full_s) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
```

	original	bootBias	bootSE	bootMed
(Intercept)	-0.134569	-0.0036411	0.079129	-0.142816
TDDS_p_z	0.215448	-0.0012646	0.060393	0.214011
GENE_p_z	0.080198	-0.0005672	0.058824	0.078091
sexFemale	0.271962	0.0066281	0.116184	0.278642

```
confint(boot_H2_full_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.26815673	0.03305607
TDDS_p_z	0.09131601	0.33257881
GENE_p_z	-0.02763633	0.20025791
sexFemale	0.03676185	0.49832392

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_z, CRS_z, TDDS_p_z, sex)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H2_s_mediate_results <- mediate(model.m = H2_M1_s,
  model.y = H2_full_s,
  treat = "TDDS_p_z",
  mediator = "GENE_p_z",
  covariates = "sex",
  boot = TRUE, # Bootstrap the CI for the indirect effect
  sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results for the mediation analysis
summary(H2_s_mediate_results)
```

## Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

(Inference Conditional on the Covariate Values Specified in `covariates')

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.00304	-0.00715	0.02	0.58
ADE	0.21545	0.10555	0.34	<2e-16 ***
Total Effect	0.21849	0.10907	0.35	<2e-16 ***
Prop. Mediated	0.01391	-0.03591	0.09	0.58

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- **Multiple Regression Results:**

- Based on the bootstrapped CIs not containing zero, both pathogen disgust and sex, but not out-group avoidance, are significantly positively related to religiosity (with the dummy coding of sex being 1 = Female).

- **Mediation Analysis Results:**

- The ACME indicates that there is no significant indirect effect of pathogen disgust on religiosity through out-group avoidance ( $B = .003$ ,  $BCI = [-.007, .02]$ ).
- This result is inconsistent with both the weak and strong version version of Hypothesis 2. That is, there is no evidence here that the relationship between pathogen disgust and religiosity is mediated by out-group avoidance, even when controlling for sex.

## Hypothesis 3: Controlling for Sex

### Preliminary Simple Linear Regression Models

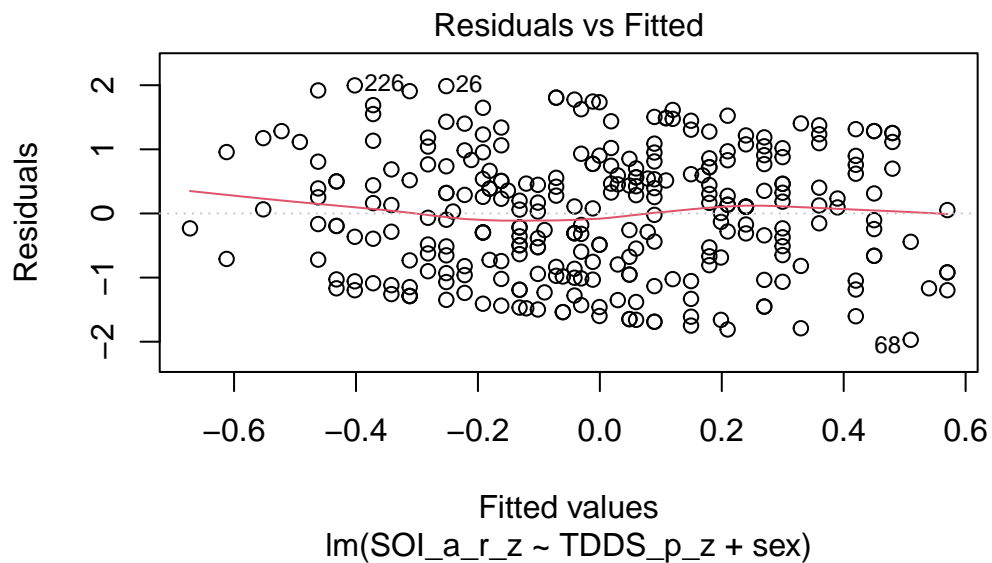
We will now re-test Hypothesis 3 while controlling for sex.

```
# Fitting the models
H3_M1_s <- lm(SOI_a_r_z ~ TDDS_p_z + sex, data = data)
H3_M2_s <- lm(CRS_z ~ SOI_a_r_z + sex, data = data)
```

### Assumptions

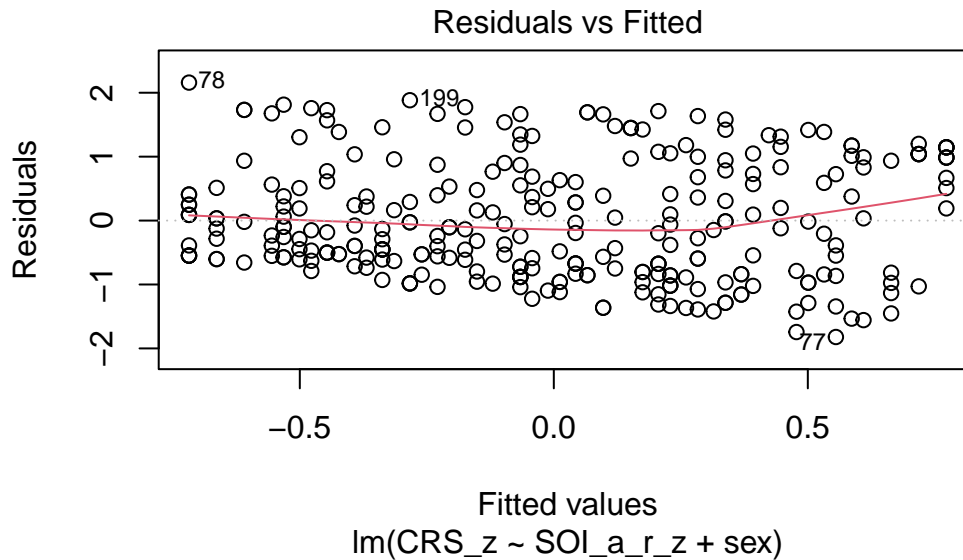
1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H3_M1_s, which = 1) # For model 1
```



```
plot(H3_M2_s, which = 1) # For model 2
```



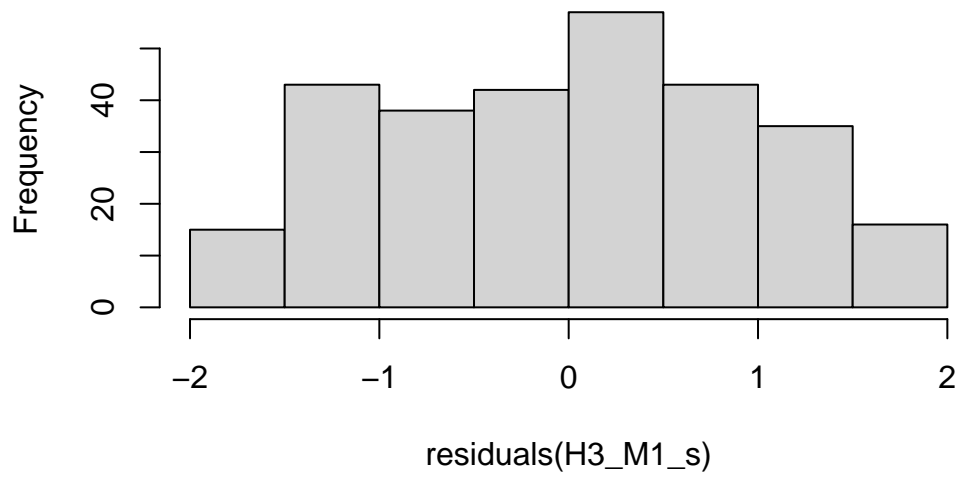


- Unlike the previous two plots, the predicted values for model 1 are not very different by sex (or at least not obviously). There is a bit of downward trend, so we will bootstrap for inferences of the coefficients.
- The second plot is quite similar to the first.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

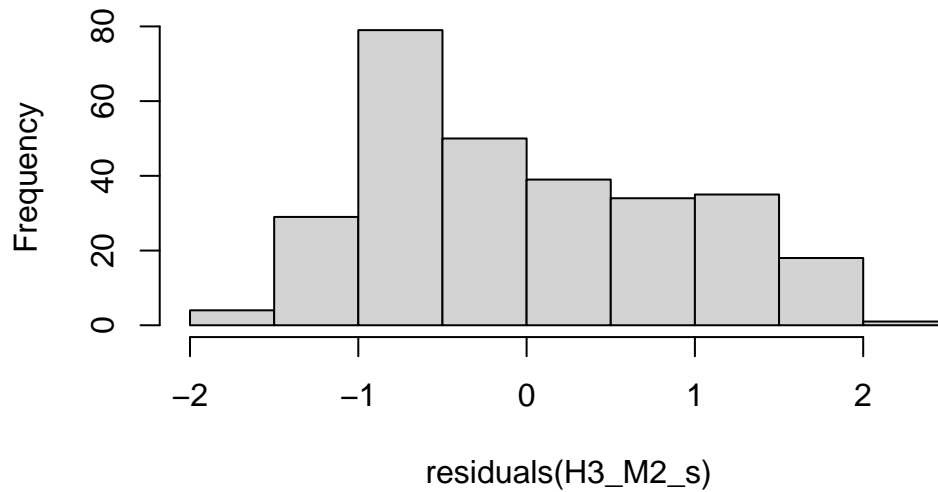
```
# Histogram for normality of residuals
hist(residuals(H3_M1_s)) # For model 1
```

**Histogram of residuals(H3\_M1\_s)**



```
hist(residuals(H3_M2_s)) # For model 2
```

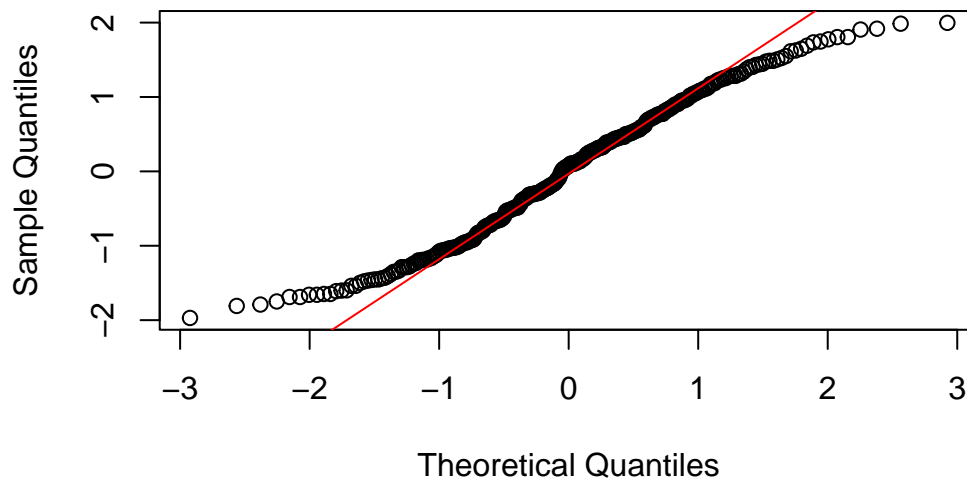
**Histogram of residuals(H3\_M2\_s)**



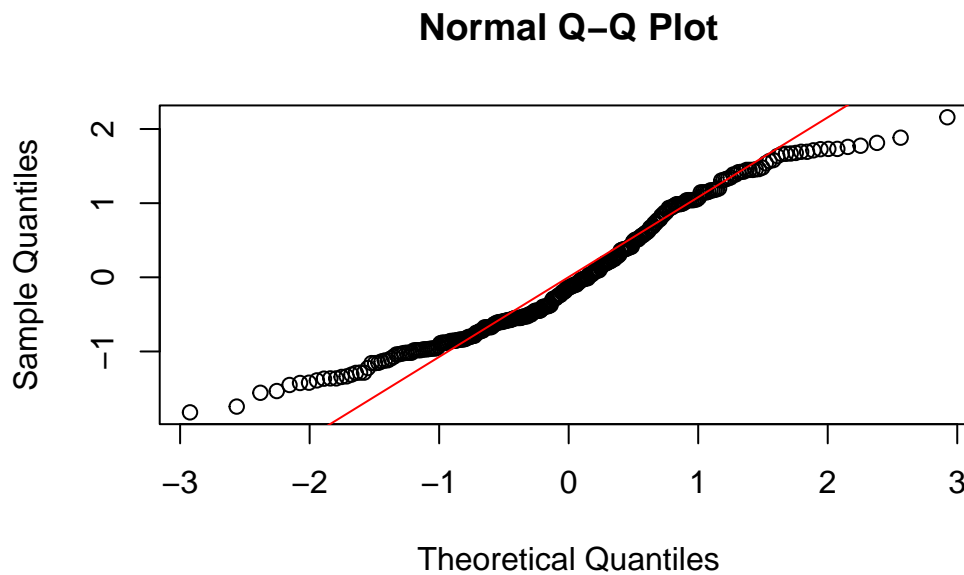
```
# QQ-plot for normality of residuals of Model 1  
qqnorm(residuals(H3_M1_s))
```

```
qqline(residuals(H3_M1_s), col = "red")
```

### Normal Q-Q Plot



```
# QQ-plot for normality of residuals of Model 2  
qqnorm(residuals(H3_M2_s))  
qqline(residuals(H3_M2_s), col = "red")
```



- **Both Models:**

- The QQ-plots for both models are not as bad as in previous models, but there is some deviation from normality towards the ends of the distributions.

- We will conduct a bootstrapped tests of the predictors for both models.

### 3. Multicollinearity: Variance inflation factor (VIF) for both models

```
# Calculating the VIF for both models
vif(H3_M1_s)
```

```
TDDS_p_z      sex
1.038981 1.038981
```

```
vif(H3_M2_s)
```

```
SOI_a_r_z      sex
1.038623 1.038623
```

- The VIF values are good here.

### Summarizing the Models

## Model 1

```
# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)
sex <- data$sex

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_M1_s <- Boot(H3_M1_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H3_M1_s) # Getting the coefficients
```

Number of bootstrap replications R = 1000

	original	bootBias	bootSE	bootMed
(Intercept)	-0.15403	-0.00051909	0.083356	-0.15251
TDDS_p_z	0.19067	0.00065110	0.059984	0.19025
sexFemale	0.31129	0.00377787	0.113956	0.31817

```
confint(boot_H3_M1_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.3198589	0.007813617
TDDS_p_z	0.0761788	0.308853893
sexFemale	0.0794630	0.540941360

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(SOI_a_r_z, TDDS_p_z, sex)
```

- Based on the estimate and BCI for pathogen disgust ( $B = .191$ ,  $BCI = [.076, .309]$ ) and sex ( $B = .311$ ,  $BCI = [.079, .541]$ ), there is a significant relationship between both variables and restricted sociosexual attitudes.

## Model 2

```
# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
CRS_z <- as.numeric(data$CRS_z)
sex <- data$sex

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_M2_s <- Boot(H3_M2_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H3_M2_s) # Getting the coefficients
```

Number of bootstrap replications R = 1000

	original	bootBias	bootSE	bootMed
(Intercept)	-0.092212	-0.00303183	0.078827	-0.10045
SOI_a_r_z	0.391337	-0.00026862	0.055955	0.39226
sexFemale	0.186359	0.00421532	0.113045	0.18804

```
confint(boot_H3_M2_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.23088030	0.08466747
SOI_a_r_z	0.27130629	0.49262933
sexFemale	-0.02599912	0.41658472

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(SOI_a_r_z, CRS_z, sex)
```

- Based on the estimates, restricted sociosexual attitudes is significantly related to religiosity independent of sex ( $B = .391$ ,  $BCI = [.271, .493]$ ), and sex drops just below significance as a predictor of religiosity ( $B = .186$ ,  $BCI = [-.026, .417]$ ).

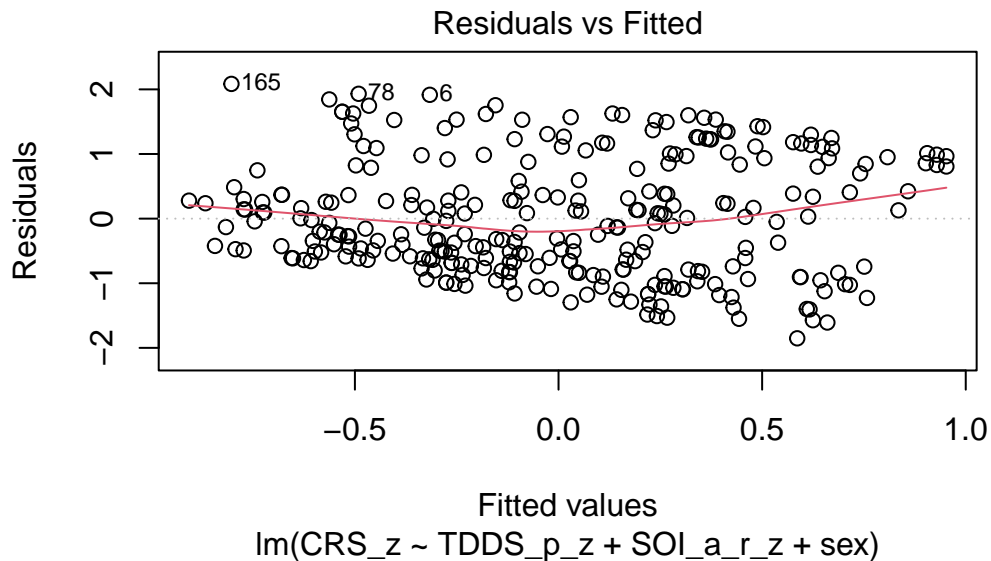
## Full Mediation Model

```
# Fitting the full mediation model
H3_full_s <- lm(CRS_z ~ TDDS_p_z + SOI_a_r_z + sex, data = data)
```

### Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

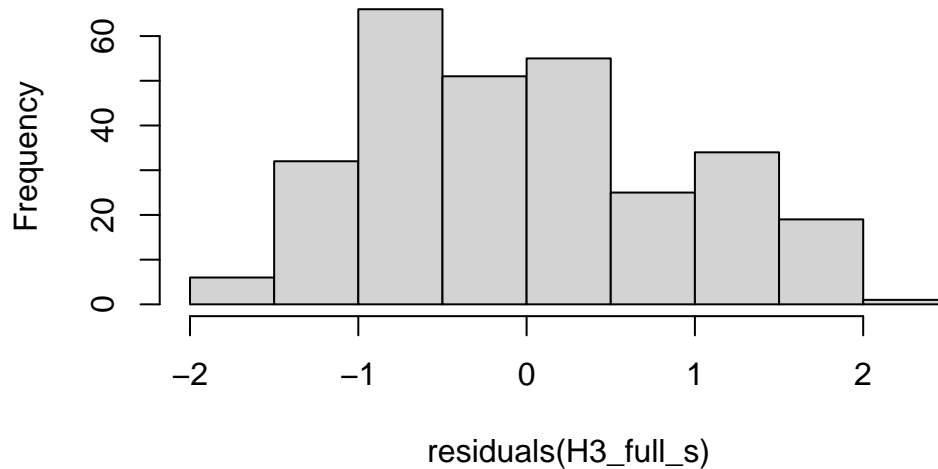
```
# Plotting the residual vs. fitted values
plot(H3_full_s, which = 1)
```



- There is a trend towards more negative residual values at the higher fitted values of the model. This would be problematic for parametric tests.
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

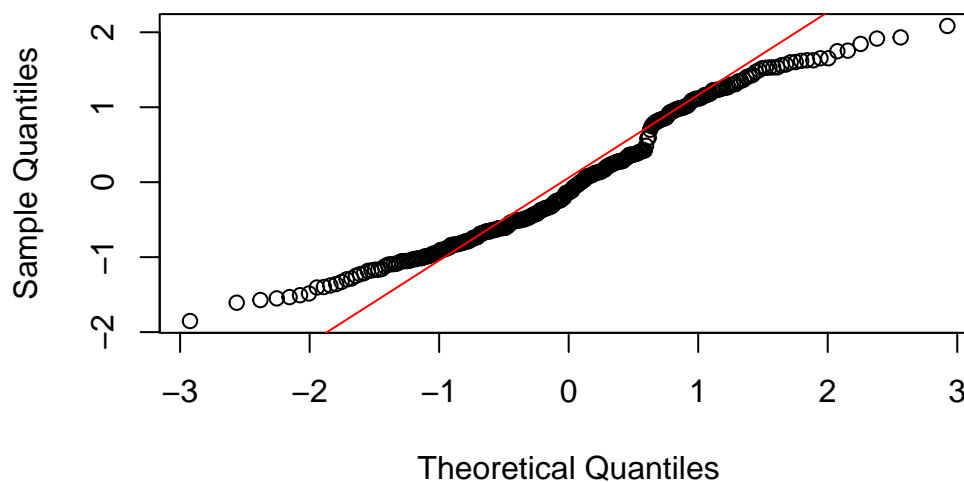
```
# Histogram for normality of residuals
hist(residuals(H3_full_s))
```

**Histogram of residuals(H3\_full\_s)**



```
# QQ-plot for normality of residuals of the full model  
qqnorm(residuals(H3_full_s))  
qqline(residuals(H3_full_s), col = "red")
```

**Normal Q-Q Plot**





- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not quite normally distributed.
- We will conduct bootstrapping for confidence intervals to make inferences.

### 3. Multicollinearity: Variance Inflation Factor (VIF) from `car` package

```
# Calculating VIF
vif(H3_full_s)
```

```
TDDS_p_z SOI_a_r_z      sex
1.078165  1.077794  1.065178
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
CRS_z <- as.numeric(data$CRS_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)
sex <- data$sex

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_full_s <- Boot(H3_full_s, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H3_full_s) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
      original  bootBias  bootSE  bootMed
(Intercept) -0.069064 -0.0038943 0.078950 -0.078932
TDDS_p_z      0.149296 -0.0010695 0.057400  0.148526
SOI_a_r_z      0.362880 -0.0001937 0.057988  0.361283
sexFemale      0.139577  0.0051513 0.113132  0.146277
```

```
confint(boot_H3_full_s) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.20809145	0.1084253
TDDS_p_z	0.03664188	0.2590547
SOI_a_r_z	0.24555603	0.4757293
sexFemale	-0.08302237	0.3537507

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(SOI_a_r_z, CRS_z, TDDS_p_z, sex)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H3_s_mediate_results <- mediate(model.m = H3_M1_s,
                                model.y = H3_full_s,
                                treat = "TDDS_p_z",
                                mediator = "SOI_a_r_z",
                                covariates = "sex",
                                boot = TRUE, # Bootstrap the CI for the indirect effect
                                sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results for the mediation analysis
summary(H3_s_mediate_results)
```

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

(Inference Conditional on the Covariate Values Specified in `covariates')

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.0692	0.0266	0.12	<2e-16 ***
ADE	0.1493	0.0431	0.27	0.008 **
Total Effect	0.2185	0.1091	0.35	<2e-16 ***
Prop. Mediated	0.3167	0.1362	0.62	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- **Multiple Regression Results:**

- Based on the bootstrapped CIs not containing zero, both pathogen disgust and restricted sociosexual attitudes, but not sex, are significantly positively related to religiosity (with the dummy coding of sex being 1 = Female).

- **Mediation Analysis Results:**

- The ACME indicates that there is a significant indirect effect of pathogen disgust on religiosity through a monogamous mating strategy ( $B = .069$ ,  $BCI = [.027, .12]$ ).
  - Given this result, we, again, have evidence for the weak version of Hypothesis 3. That is, the relationship between pathogen disgust and religiosity is partially mediated by a monogamous mating strategy, even while controlling for sex.
- The proportion of the total effect mediated (the ratio of the indirect effect to the direct effect) is .32. That is, 32% of the covariance between pathogen disgust and religiosity seems to be able to be accounted for by restricted sociosexual attitudes. Our inferential criterion for concluding full mediation was that this ratio would be at least .80, with the total effect greater than a  $B = .2$ . Even the upper bound of the bootstrapped confidence interval for this proportion mediated does not cross this threshold ( $BCI = [.136, .62]$ ).
  - Given this result, we do not have evidence for the strong version of Hypothesis 3. That is, the relationship between pathogen disgust and religiosity is not fully mediated by a monogamous mating strategy.

## Hypothesis 1: Religious Importance

### Preliminary Simple Linear Regression Models

Now we will retest each hypothesis with religious importance as the outcome variable, as a check for robustness.

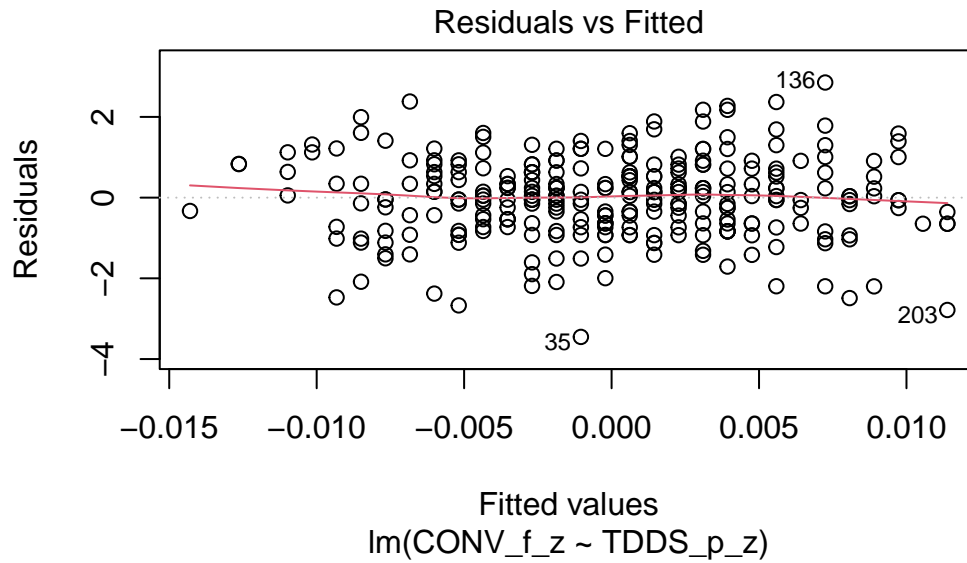
As mentioned above, to test Hypothesis 1, we must first establish that pathogen disgust affects adherence to traditional practices (i.e., `CONV_f_z`) and that adherence to traditional practices influences religious importance.

```
# Fitting the models
H1_M1_RI <- lm(CONV_f_z ~ TDDS_p_z, data = data)
H1_M2_RI <- lm(religious_importance_z ~ CONV_f_z, data = data)
```

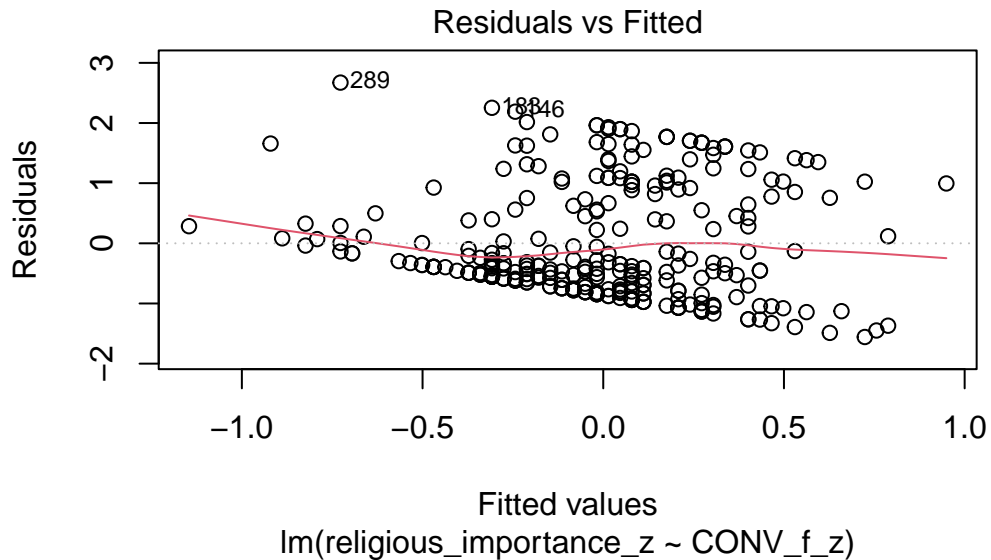
## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values  
plot(H1_M1_RI, which = 1) # For model 1
```



```
plot(H1_M2_RI, which = 1) # For model 2
```

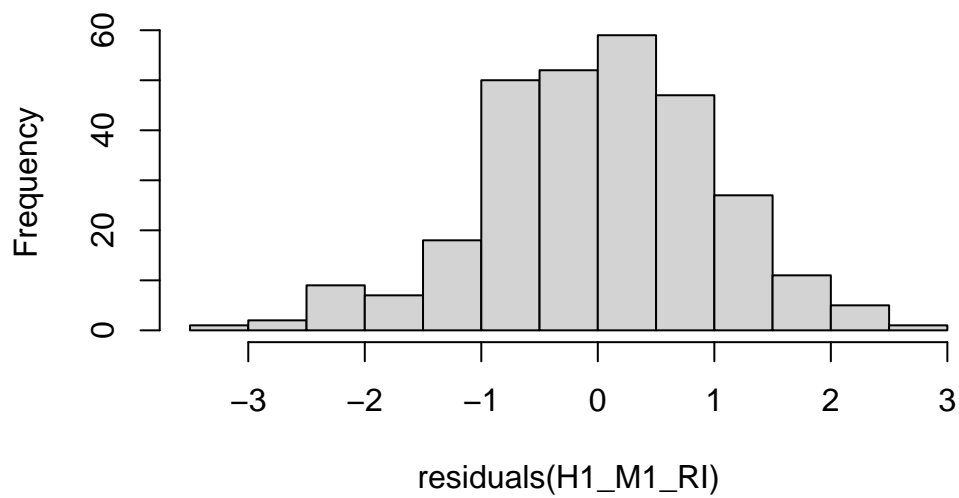


- The residuals vs. fitted plot does not stand out as violating the assumption of linearity or heteroscedasticity for the first model, but there is a trend towards more negative residual values at the higher fitted values of the model for model 2. This is problematic for parametric inference.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

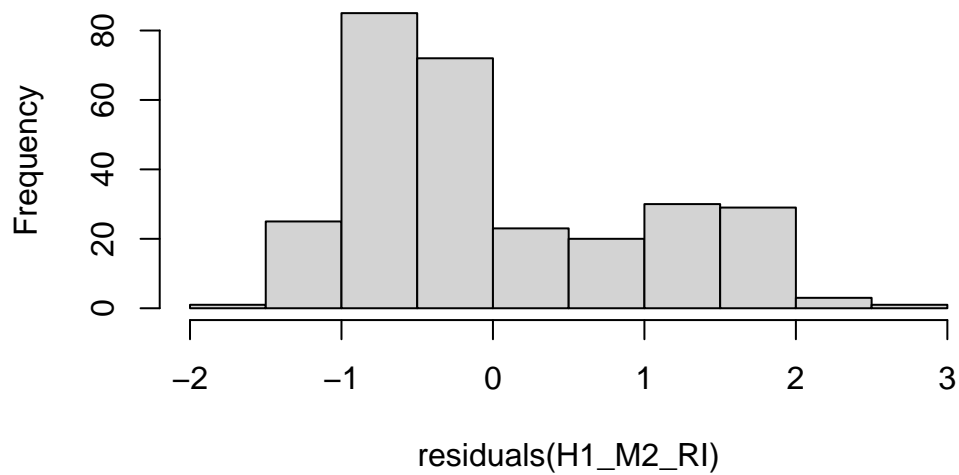
```
# Histogram for normality of residuals
hist(residuals(H1_M1_RI)) # For model 1
```

**Histogram of residuals(H1\_M1\_RI)**



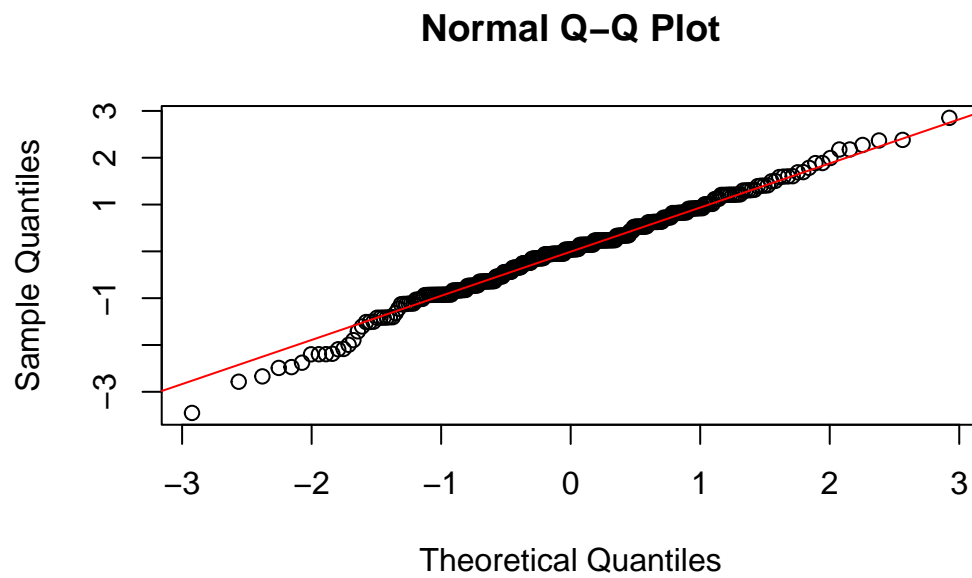
```
hist(residuals(H1_M2_RI)) # For model 2
```

**Histogram of residuals(H1\_M2\_RI)**



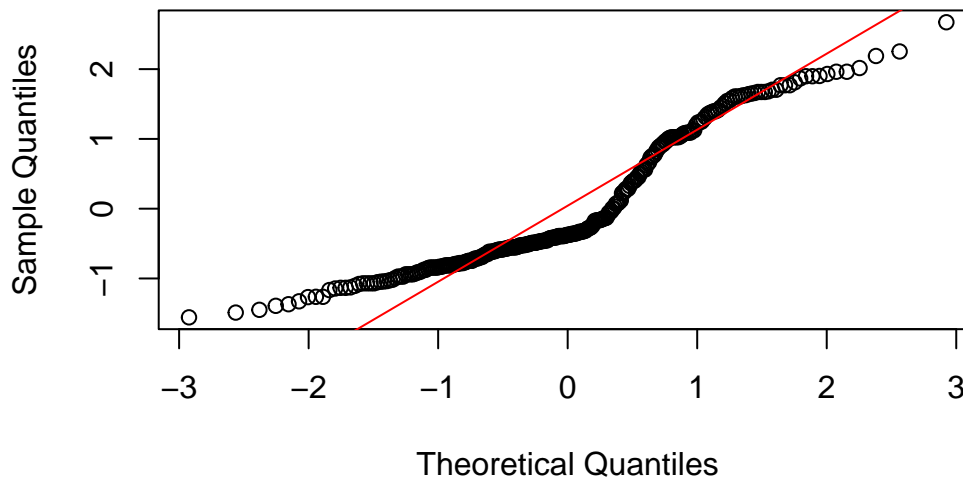
```
# QQ-plot for normality of residuals of Model 1  
qqnorm(residuals(H1_M1_RI))
```

```
qqline(residuals(H1_M1_RI), col = "red")
```



```
# QQ-plot for normality of residuals of Model 2  
qqnorm(residuals(H1_M2_RI))  
qqline(residuals(H1_M2_RI), col = "red")
```

## Normal Q-Q Plot



- **Model 1:**
  - The histogram and QQ-plot suggest that the residuals are normally distributed for Model 1.
- **Model 2:**
  - The histogram and QQ-plot of the residuals of Model 2 indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 SD.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictor for Model 2.

3. **Multicollinearity:** Not applicable because each model only has one predictor.

## Summarizing the Models

### Model 1

```
# Create a scatterplot with the regression line to visualize
plot(data$TDDS_p_z, data$CONV_f_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Pathogen Disgust",
```

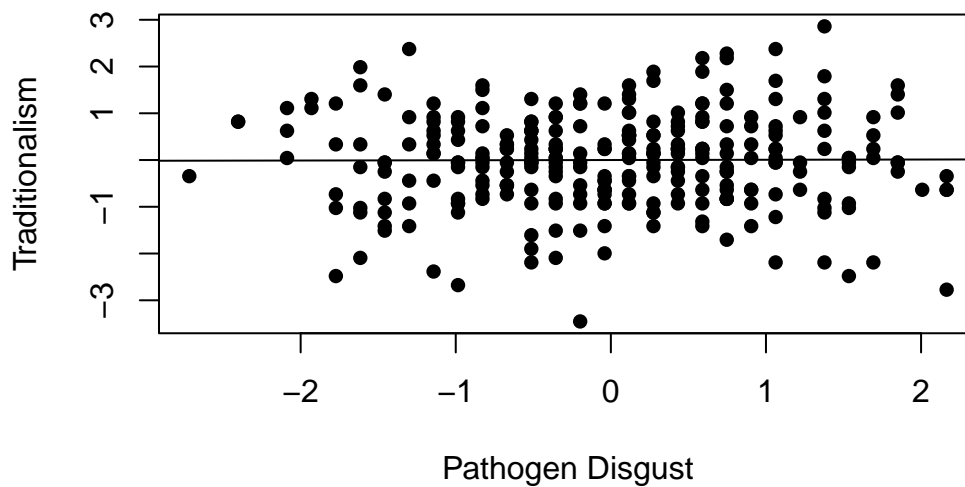


```

ylab = "Traditionalism",
pch = 16, # For solid circle points
col = "black")
# Add regression line
abline(H1_M1_RI, col = "black", lwd = 1)

```

## Scatter Plot with Linear Regression Line



```

# Summarizing the model
summary(H1_M1_RI)

```

Call:

```
lm(formula = CONV_f_z ~ TDDS_p_z, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.4517	-0.6429	0.0402	0.6287	2.8524

Coefficients:

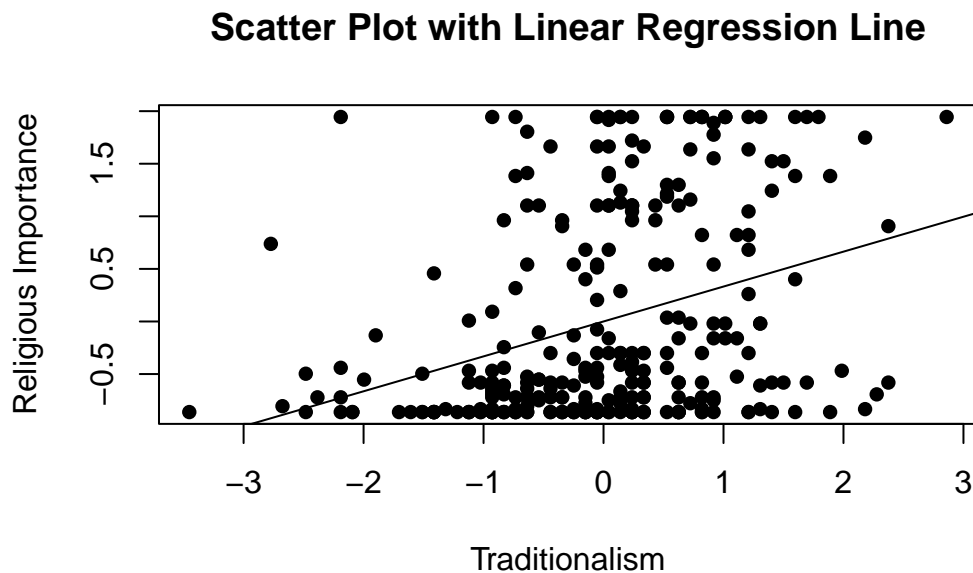
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-7.860e-17	5.893e-02	0.000	1.000
TDDS_p_z	5.260e-03	5.903e-02	0.089	0.929

Residual standard error: 1.002 on 287 degrees of freedom  
Multiple R-squared: 2.766e-05, Adjusted R-squared: -0.003457  
F-statistic: 0.00794 on 1 and 287 DF, p-value: 0.9291

- We can see from the Scatter Plot that there is no relationship between pathogen disgust and traditionalism to speak of ( $B = .005$ ,  $t(287) = .089$ ,  $p = .929$ ), just as with the correlation above.
  - Although this precludes the possibility that adherence to tradition mediates the relationship between disgust and religious importance, we will continue with the analysis for completions sake.

## Model 2

```
# Create a scatterplot with the regression line to visualize
plot(data$CONV_f_z, data$religious_importance_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Traditionalism",
     ylab = "Religious Importance",
     pch = 16, # For solid circle points
     col = "black")
# Add regression line
abline(H1_M2_RI, col = "black", lwd = 1)
```



```
# Adding the variables to the broader environment so the Boot function will run
CONV_f_z <- as.numeric(data$CONV_f_z)
religious_importance_z <- as.numeric(data$religious_importance_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H1_M2_RI <- Boot(H1_M2_RI, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H1_M2_RI) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original  bootBias  bootSE  bootMed
(Intercept) 2.3298e-17 4.6614e-06 0.055910 -0.0027938
CONV_f_z     3.3199e-01 1.4028e-03 0.055438 0.3338146
```

```
confint(boot_H1_M2_RI) # Getting the BCIs
```

Bootstrap bca confidence intervals

```
              2.5 %    97.5 %
(Intercept) -0.1002662 0.1200648
CONV_f_z     0.2204035 0.4370547
```

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(CONV_f_z, religious_importance_z)
```

- We can see from the Scatter Plot that there is a positive relationship between traditionalism and religious importance.
- Based on the estimates, the standardized coefficient for traditionalism is around .33.
  - In addition, this relationship is significant, as the BCIs do not contain zero (upper = .44, lower = .22).

## Full Mediation Model

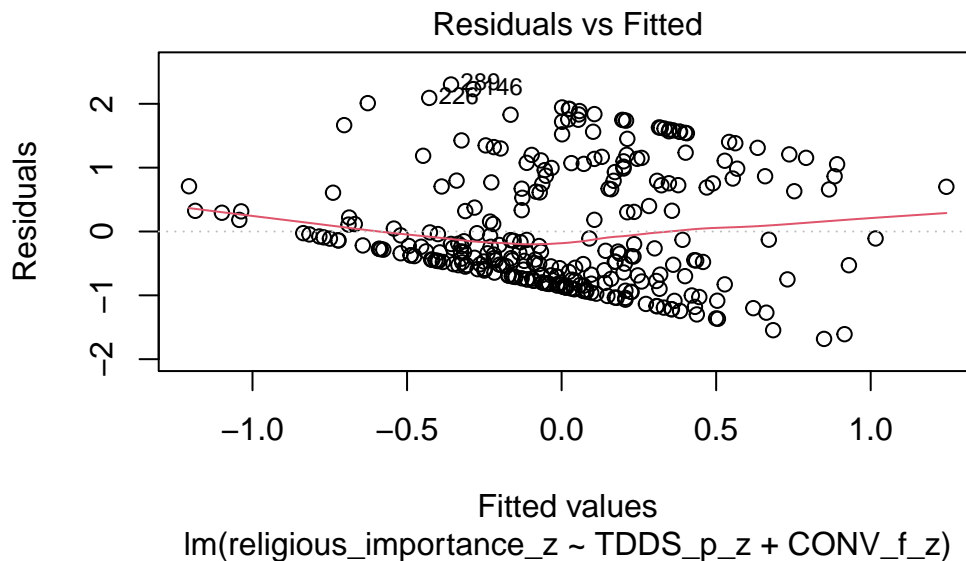
The fact that pathogen disgust does not predict traditionalism indicates that traditionalism could not mediate the relationship between pathogen disgust and religious importance, but I will run the analysis that we planned anyway just to have the results.

```
# Fitting the full mediation model
H1_full_RI <- lm(religious_importance_z ~ TDDS_p_z + CONV_f_z, data = data)
```

## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

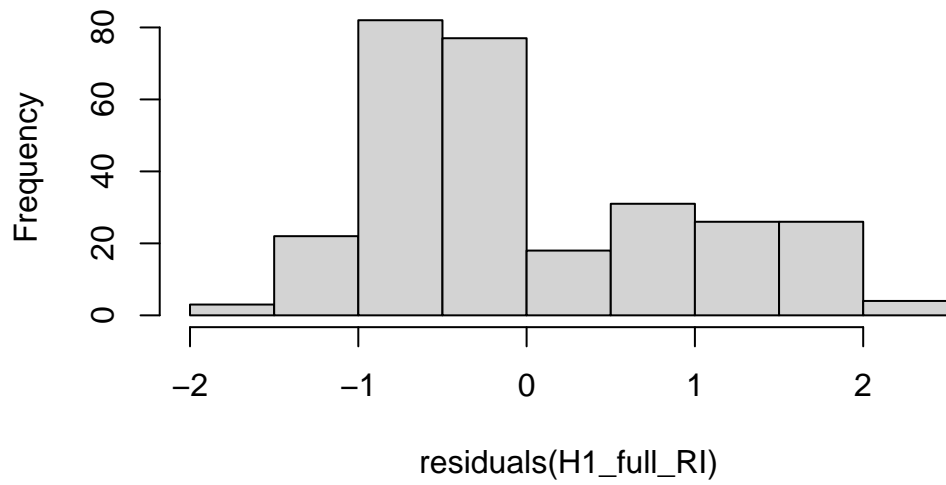
```
# Plotting the residual vs. fitted values
plot(H1_full_RI, which = 1)
```



- There is a trend towards more negative residual values at the higher fitted values of the model. This is problematic for parametric inference.
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

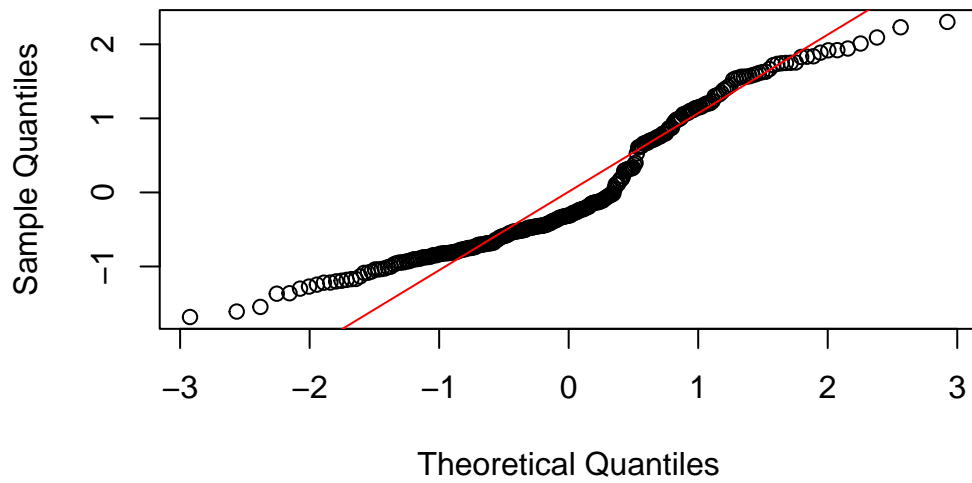
```
# Histogram for normality of residuals
hist(residuals(H1_full_RI))
```

**Histogram of residuals(H1\_full\_RI)**



```
# QQ-plot for normality of residuals of the full model
qqnorm(residuals(H1_full_RI))
qqline(residuals(H1_full_RI), col = "red")
```

**Normal Q-Q Plot**



- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just below 1 SD.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictors for the model.

### 3. Multicollinearity: Variance Inflation Factor (VIF) from `car` package

```
# Calculating VIF
vif(H1_full_RI)
```

```
TDDS_p_z CONV_f_z
1.000028 1.000028
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
CONV_f_z <- as.numeric(data$CONV_f_z)
religious_importance_z <- as.numeric(data$religious_importance_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)
```

```
# Setting the seed
set.seed(1042557)
```

```
# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H1_full_RI <- Boot(H1_full_RI, R = 1000) # 1,000 resamples
```

```
# Summarizing the estimates and CIs
summary(boot_H1_full_RI) # Getting the coefficients
```

Number of bootstrap replications R = 1000

	original	bootBias	bootSE	bootMed
(Intercept)	1.4280e-17	-0.00064525	0.054844	-0.0038586
TDDS_p_z	2.1682e-01	0.00030940	0.057438	0.2166807
CONV_f_z	3.3085e-01	0.00205738	0.051646	0.3331901

```
confint(boot_H1_full_RI) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.09907331	0.1172830
TDDS_p_z	0.10874782	0.3291841
CONV_f_z	0.22029722	0.4256569

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(CONV_f_z, religious_importance_z, TDDS_p_z)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H1_RI_mediate_results <- mediate(model.m = H1_M1_RI,
  model.y = H1_full_RI,
  treat = "TDDS_p_z",
  mediator = "CONV_f_z",
  boot = TRUE, # Bootstrap the CI for the indirect effect
  sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results for the mediation analysis
summary(H1_RI_mediate_results)
```

## Causal Mediation Analysis

### Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.00174	-0.03773	0.04	0.96
ADE	0.21682	0.10884	0.33	<2e-16 ***
Total Effect	0.21856	0.11198	0.34	<2e-16 ***
Prop. Mediated	0.00796	-0.22993	0.21	0.96

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- Multiple Regression Results:

- Based on the bootstrapped CIs not containing zero, both pathogen disgust and traditionalism are significantly positively related to religious importance.
- The coefficients for pathogen disgust (.22) and traditionalism (.33) are similar to the model with the CRS.

- **Mediation Analysis Results:**

- The ACME indicates that there is no significant indirect effect of pathogen disgust on religious importance through traditionalism ( $B = .002$ ,  $BCI = [-.038, .04]$ ).
- This result is inconsistent with both the weak and strong version version of Hypothesis 1. That is, there is no evidence here that the relationship between pathogen disgust and religious importance is mediated by traditionalism.

## Hypothesis 2: Religious Importance

### Preliminary Simple Linear Regression Models

As mentioned above, to test Hypothesis 2, we must first establish that pathogen disgust affects out-group avoidance (i.e., GENE\_p\_z) and that out-group avoidance influences religious importance.

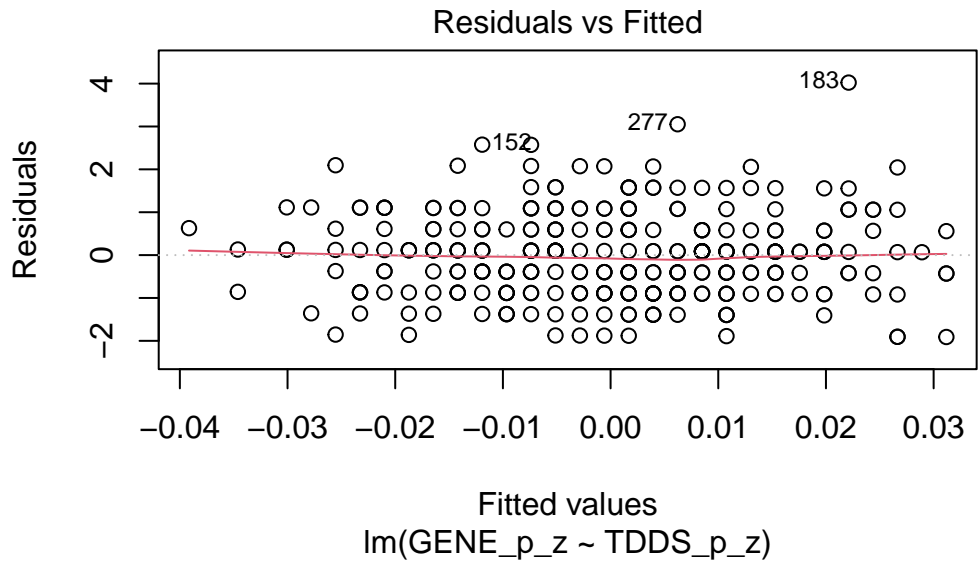
```
# Fitting the models
H2_M1_RI <- lm(GENE_p_z ~ TDDS_p_z, data = data)
H2_M2_RI <- lm(religious_importance_z ~ GENE_p_z, data = data)
```

### Assumptions

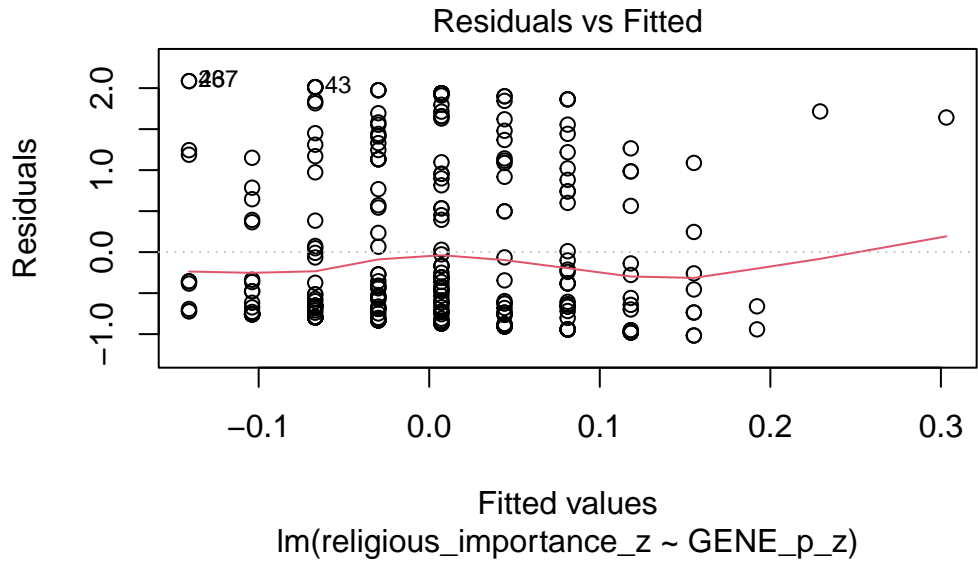
1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H2_M1_RI, which = 1) # For model 1
```





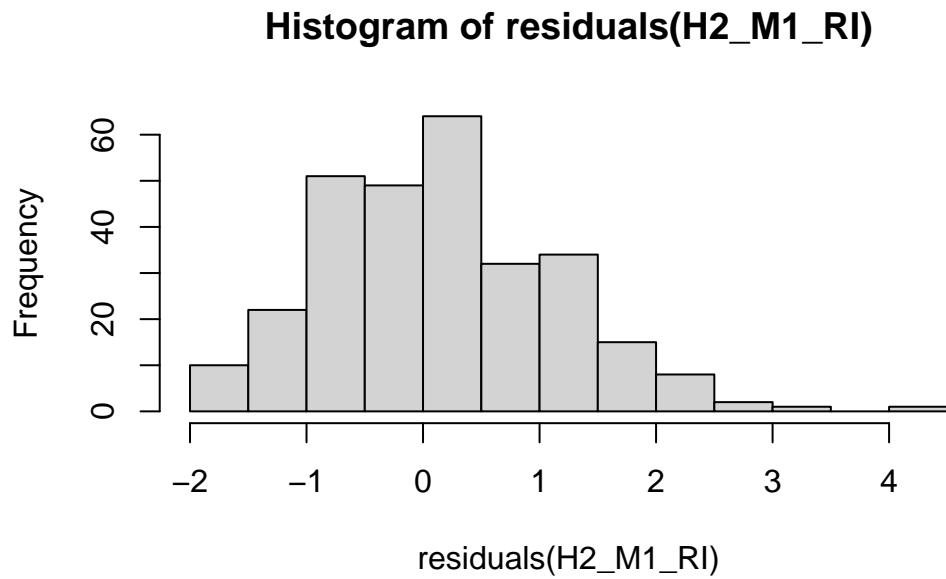
```
plot(H2_M2_RI, which = 1) # For model 2
```



- The residuals vs. fitted plot does not stand out as violating the assumption of linearity or heteroscedasticity for the either model.

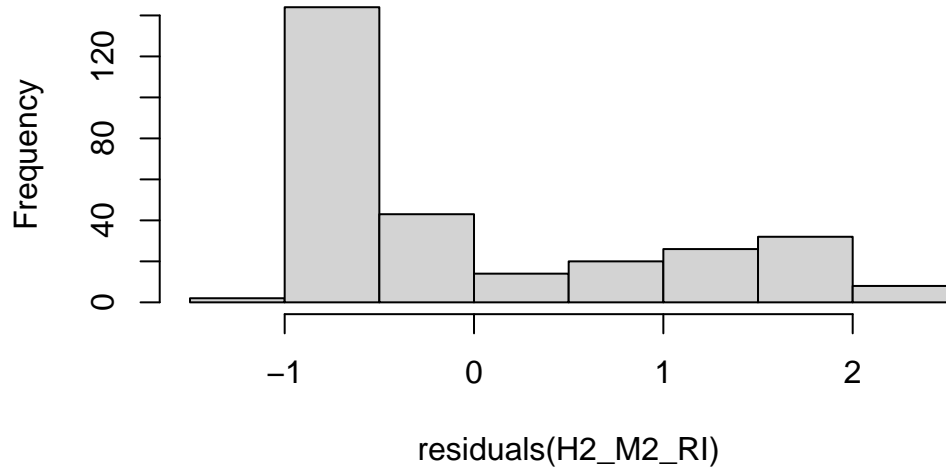
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

```
# Histogram for normality of residuals  
hist(residuals(H2_M1_RI)) # For model 1
```



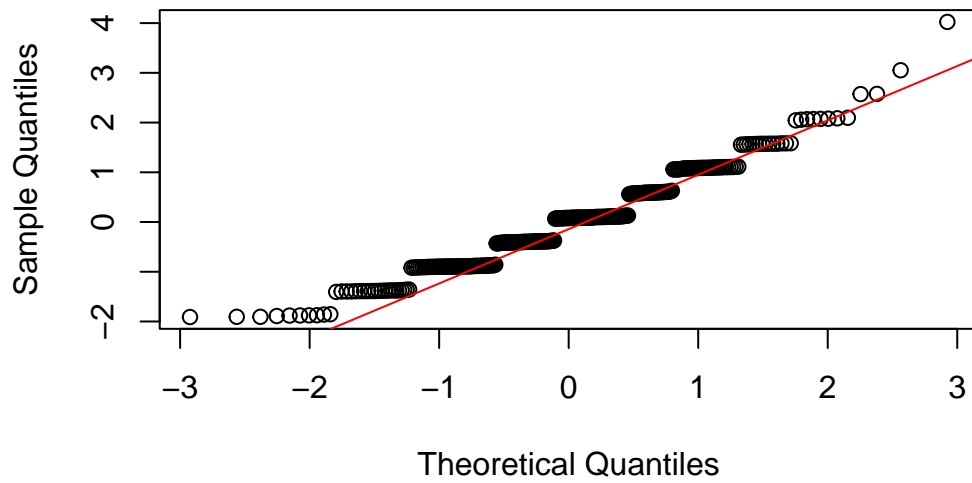
```
hist(residuals(H2_M2_RI)) # For model 2
```

**Histogram of residuals(H2\_M2\_RI)**

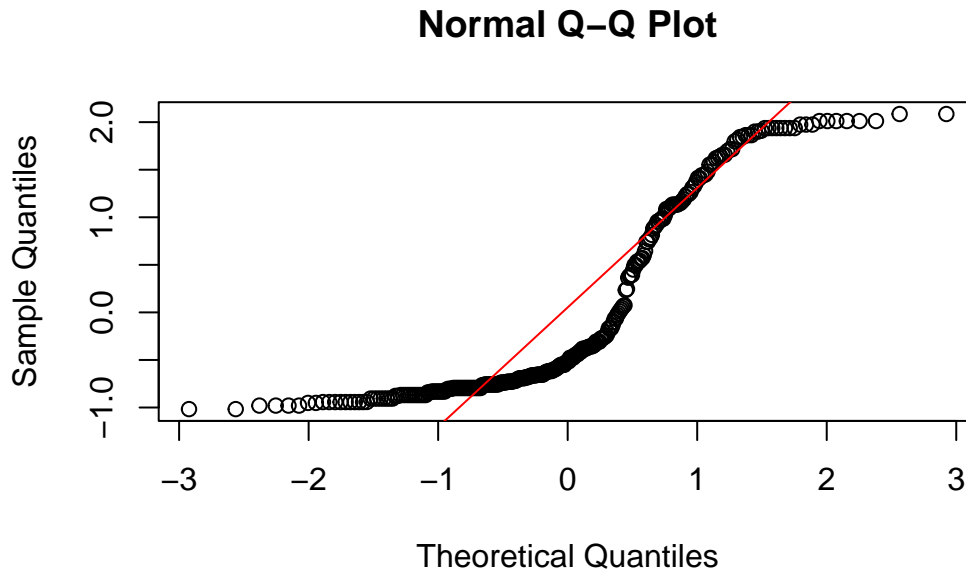


```
# QQ-plot for normality of residuals of Model 1  
qqnorm(residuals(H2_M1_RI))  
qqline(residuals(H2_M1_RI), col = "red")
```

**Normal Q-Q Plot**



```
# QQ-plot for normality of residuals of Model 2
qqnorm(residuals(H2_M2_RI))
qqline(residuals(H2_M2_RI), col = "red")
```



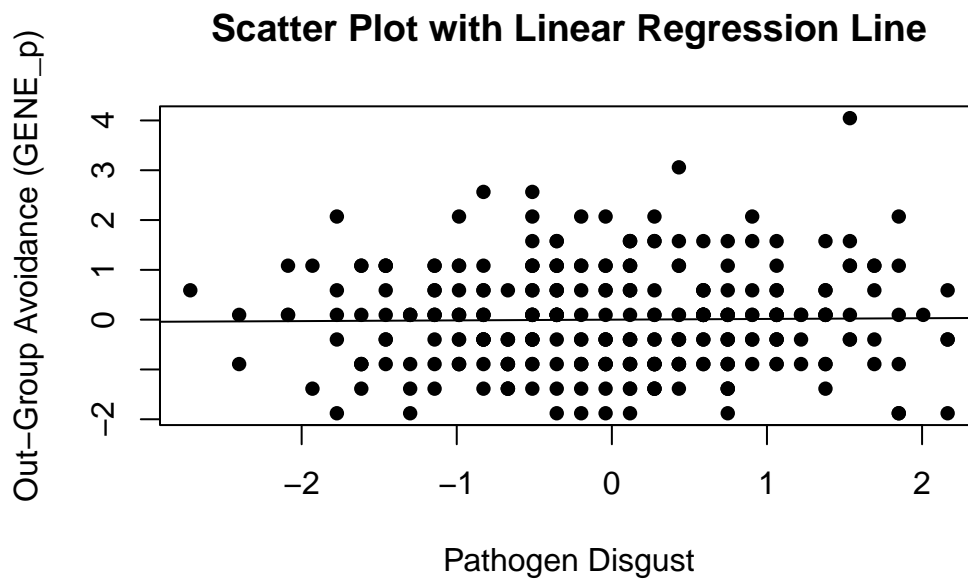
- **Model 1:**
  - The histogram and QQ-plot suggest that the residuals are relatively normally distributed for Model 1.
- **Model 2:**
  - The histogram and QQ-plot of the residuals of Model 2 indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 or 1.5 SD.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictor for Model 2.

3. **Multicollinearity:** Not applicable because each model only has one predictor.

### Summarizing the Models

## Model 1

```
# Create a scatterplot with the regression line to visualize
plot(data$TDDS_p_z, data$GENE_p_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Pathogen Disgust",
     ylab = "Out-Group Avoidance (GENE_p)",
     pch = 16, # For solid circle points
     col = "black")
# Add regression line
abline(H2_M1_RI, col = "black", lwd = 1)
```



```
# Summarizing the model
summary(H2_M1_RI)
```

Call:

```
lm(formula = GENE_p_z ~ TDDS_p_z, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.9107	-0.8800	0.0804	0.5946	4.0239

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.593e-16	5.892e-02	0.000	1.000
TDDS_p_z	1.440e-02	5.902e-02	0.244	0.807

Residual standard error: 1.002 on 287 degrees of freedom

Multiple R-squared: 0.0002075, Adjusted R-squared: -0.003276

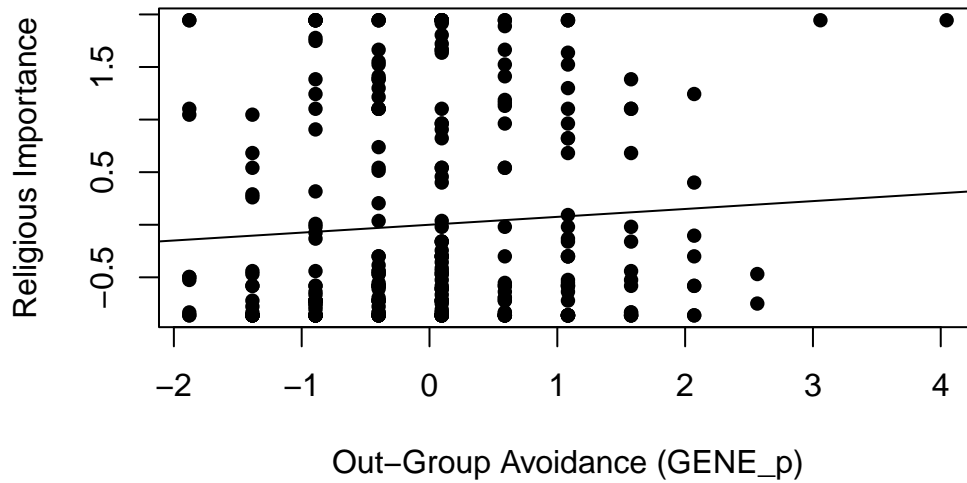
F-statistic: 0.05956 on 1 and 287 DF, p-value: 0.8074

- We can see from the Scatter Plot that there is no relationship between pathogen disgust and out-group avoidance to speak of ( $B = .014$ ,  $t(287) = .244$ ,  $p = .807$ )—same as the model above.
  - Although this precludes the possibility that out-group avoidance mediates the relationship between disgust and religious importance, we will continue with the analysis for completions sake.

## Model 2

```
# Create a scatterplot with the regression line to visualize
plot(data$GENE_p_z, data$religious_importance_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Out-Group Avoidance (GENE_p)",
     ylab = "Religious Importance",
     pch = 16, # For solid circle points
     col = "black")
# Add regression line
abline(H2_M2_RI, col = "black", lwd = 1)
```

## Scatter Plot with Linear Regression Line



```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_z <- as.numeric(data$GENE_p_z)
religious_importance_z <- as.numeric(data$religious_importance_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_M2_RI <- Boot(H2_M2_RI, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H2_M2_RI) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original    bootBias    bootSE    bootMed
(Intercept) -1.4709e-17  0.00069972 0.059421 0.00035152
GENE_p_z      7.4934e-02 -0.00254865 0.062068 0.07214367
```

```
confint(boot_H2_M2_RI) # Getting the BCIs
```

```
Bootstrap bca confidence intervals
```

```

                2.5 %    97.5 %
(Intercept) -0.11600555 0.1186471
GENE_p_z    -0.04012943 0.1972286

```

```

# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_z, religious_importance_z)

```

- We can see from the Scatter Plot that there is a positive trend in the relationship between out-group avoidance and religious importance.
- Based on the estimates, the standardized coefficient for traditionalism is around .07.
  - However, this relationship is not significant, as the BCI contains zero (upper = .20, lower = -.04).

## Full Mediation Model

The fact that pathogen disgust does not predict out-group avoidance and, in turn, out-group avoidance does not predict religious importance indicates that out-group avoidance could not mediate the relationship between pathogen disgust and religious importance, but I will run the analysis that we planned anyway just to have the results.

```

# Fitting the full mediation model
H2_full_RI <- lm(religious_importance_z ~ TDDS_p_z + GENE_p_z, data = data)

```

## Assumptions

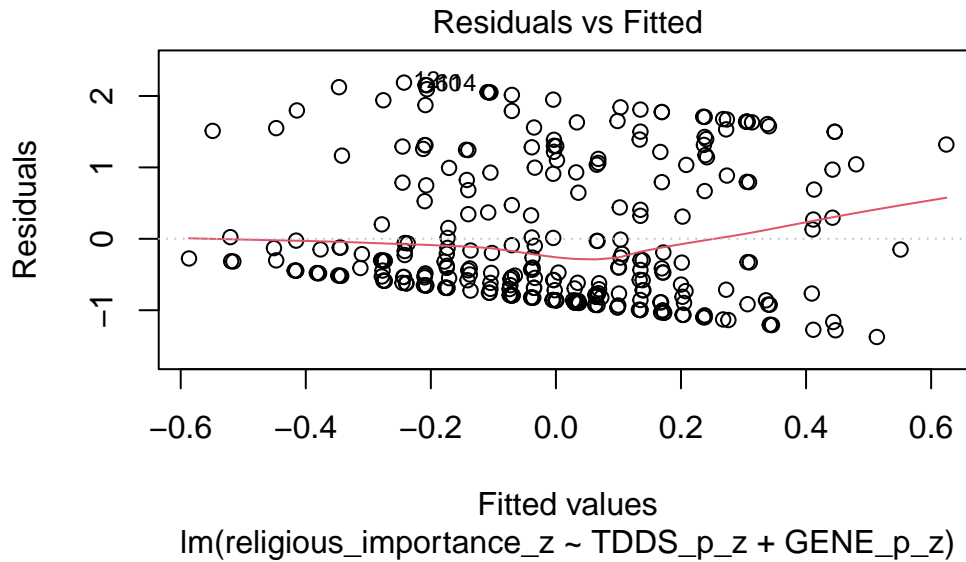
1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```

# Plotting the residual vs. fitted values
plot(H2_full_RI, which = 1)

```



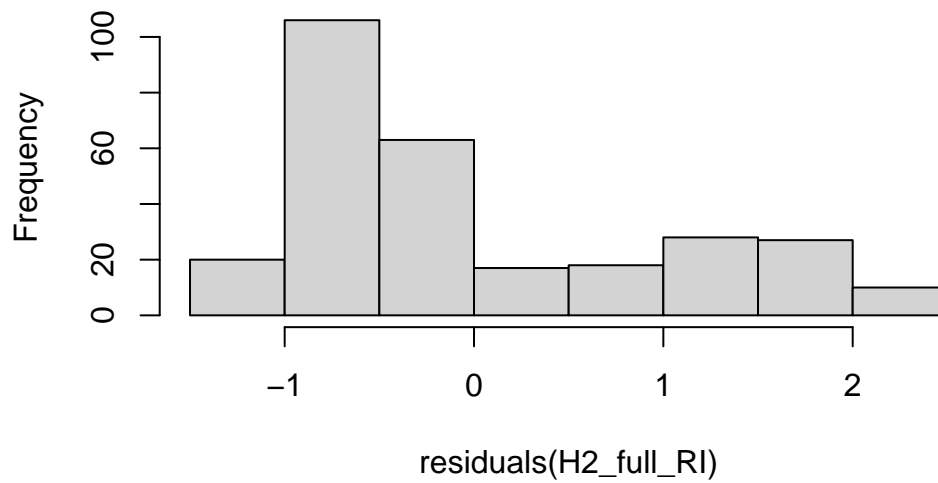


- There is a trend towards more negative residual values at the higher fitted values of the model. This may be problematic.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

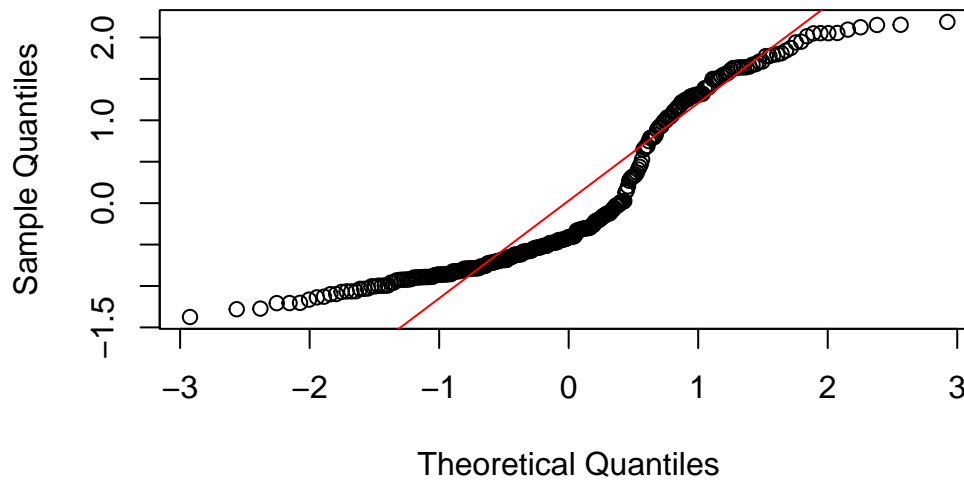
```
# Histogram for normality of residuals
hist(residuals(H2_full_RI))
```

**Histogram of residuals(H2\_full\_RI)**



```
# QQ-plot for normality of residuals of the full model
qqnorm(residuals(H2_full_RI))
qqline(residuals(H2_full_RI), col = "red")
```

**Normal Q-Q Plot**



- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 SD.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictors for the model.

### 3. Multicollinearity: Variance Inflation Factor (VIF) from `car` package

```
# Calculating VIF
vif(H2_full_RI)
```

```
TDDS_p_z GENE_p_z
1.000208 1.000208
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_z <- as.numeric(data$GENE_p_z)
religious_importance_z <- as.numeric(data$religious_importance_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)
```

```
# Setting the seed
set.seed(1042557)
```

```
# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_full_RI <- Boot(H2_full_RI, R = 1000) # 1,000 resamples
```

```
# Summarizing the estimates and CIs
summary(boot_H2_full_RI) # Getting the coefficients
```

Number of bootstrap replications R = 1000

	original	bootBias	bootSE	bootMed
(Intercept)	-2.3166e-17	5.6297e-05	0.058364	5.6535e-05
TDDS_p_z	2.1752e-01	2.5512e-04	0.059996	2.1741e-01
GENE_p_z	7.1801e-02	-2.2410e-03	0.059293	6.9262e-02

```
confint(boot_H2_full_RI) # Getting the BCIs
```

Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.11278531	0.1217364
TDDS_p_z	0.09532824	0.3302613
GENE_p_z	-0.03957865	0.1880959

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_z, religious_importance_z, TDDS_p_z)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H2_RI_mediate_results <- mediate(model.m = H2_M1_RI,
  model.y = H2_full_RI,
  treat = "TDDS_p_z",
  mediator = "GENE_p_z",
  boot = TRUE, # Bootstrap the CI for the indirect effect
  sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results of the mediation analysis
summary(H2_RI_mediate_results)
```

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.00103	-0.00902	0.01	0.83
ADE	0.21752	0.11006	0.34	<2e-16 ***
Total Effect	0.21856	0.11198	0.34	<2e-16 ***
Prop. Mediated	0.00473	-0.04487	0.07	0.83

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- Multiple Regression Results:

- Based on the bootstrapped CIs, pathogen disgust, but not out-group avoidance, is significantly positively related to religious importance.
- The coefficients for pathogen disgust (.22) and out-group avoidance (.07) are virtually identical in the model with CRS.

- **Mediation Analysis Results:**

- The ACME indicates that there is no significant indirect effect of pathogen disgust on religious importance through out-group avoidance ( $B = .001$ ,  $BCI = [-.009, .01]$ ).
- This result is inconsistent with both the weak and strong version version of Hypothesis 2. That is, there is no evidence here that the relationship between pathogen disgust and religious importance is mediated by out-group avoidance.

### Hypothesis 3: Religious Importance

#### Preliminary Simple Linear Regression Models

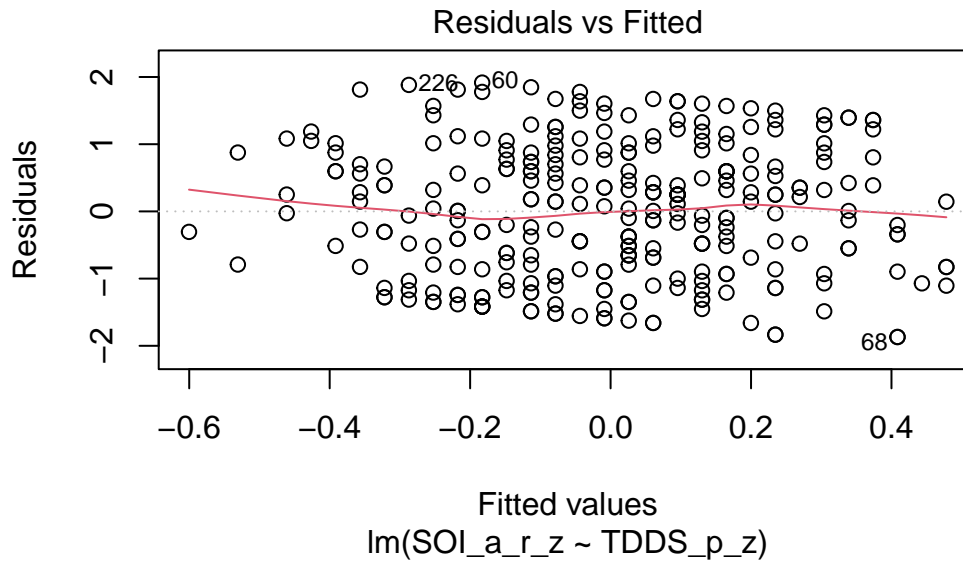
As mentioned above, to test Hypothesis 3, we must first establish that pathogen disgust affects one's mating strategy (i.e., make it more monogamous; as measured by more restricted sociosexuality; SOI\_a\_r\_z) and that a monogamous mating strategy influences religious importance.

```
# Fitting the models
H3_M1_RI <- lm(SOI_a_r_z ~ TDDS_p_z, data = data)
H3_M2_RI <- lm(religious_importance_z ~ SOI_a_r_z, data = data)
```

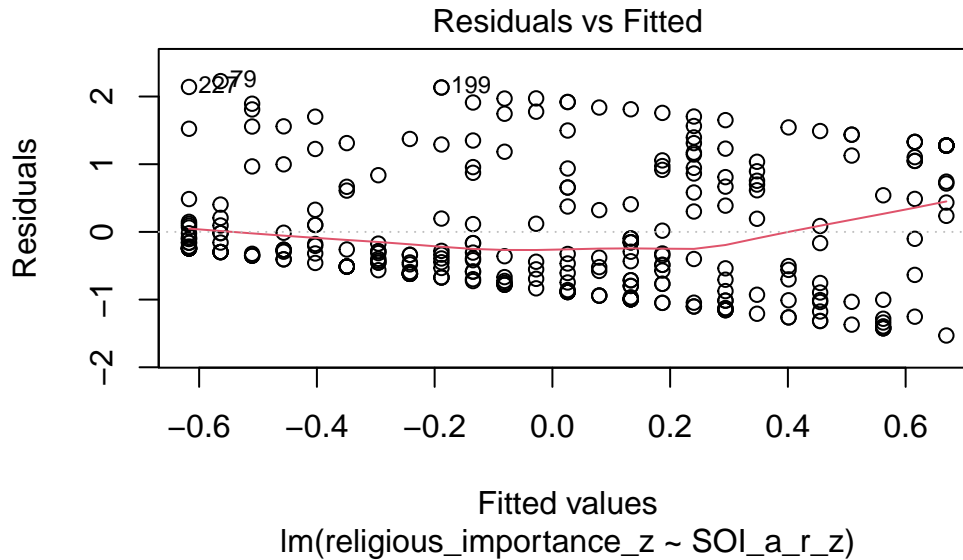
#### Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H3_M1_RI, which = 1) # For model 1
```



```
plot(H3_M2_RI, which = 1) # For model 2
```

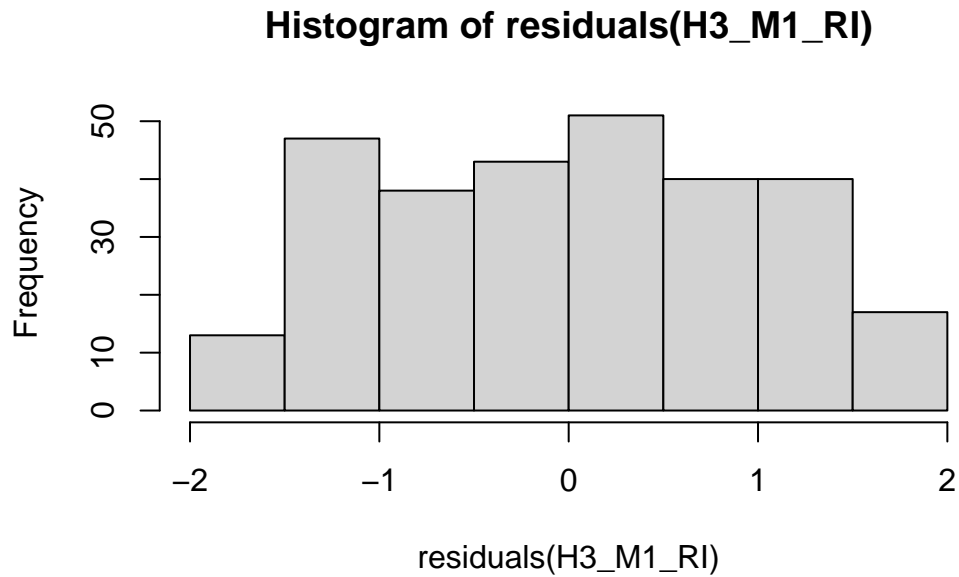


- The residuals vs. fitted plot does not stand out as violating the assumption of linearity or heteroscedasticity for model 1, but model two has a pretty decent trend

toward more negative residual values at higher fitted values.

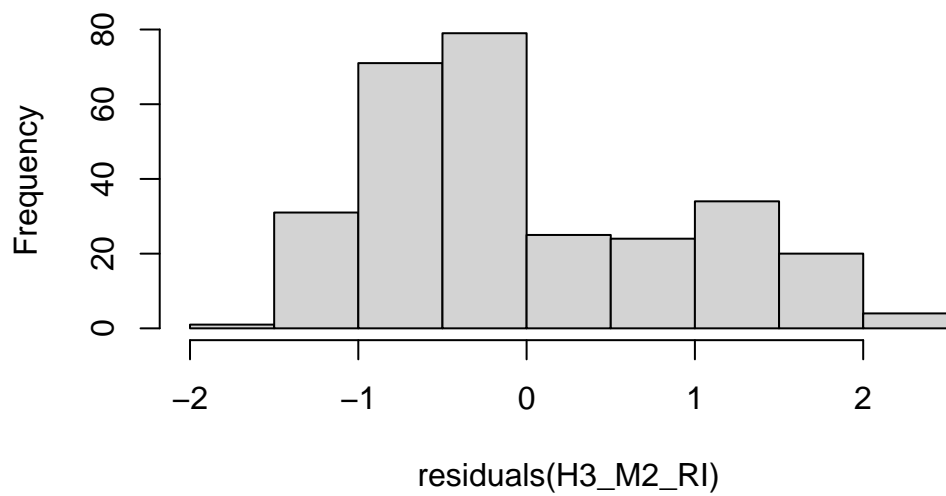
2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

```
# Histogram for normality of residuals  
hist(residuals(H3_M1_RI)) # For model 1
```



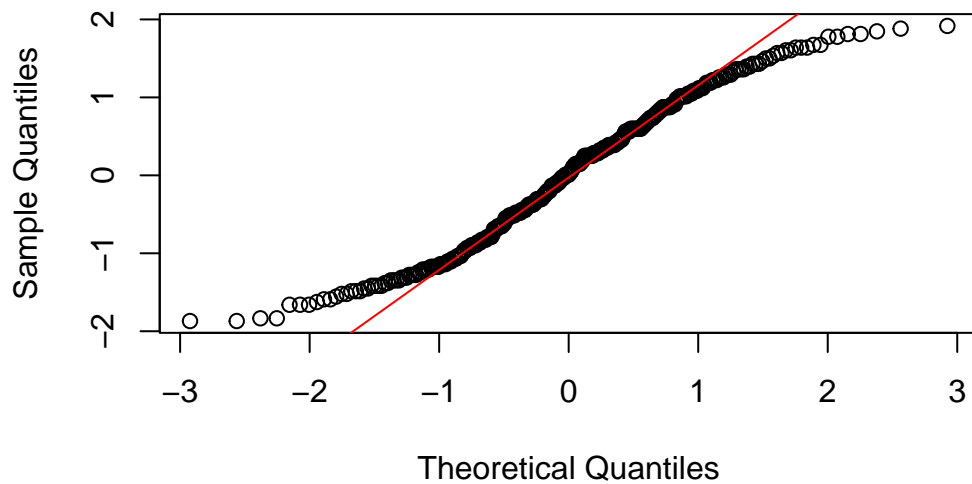
```
hist(residuals(H3_M2_RI)) # For model 2
```

**Histogram of residuals(H3\_M2\_RI)**



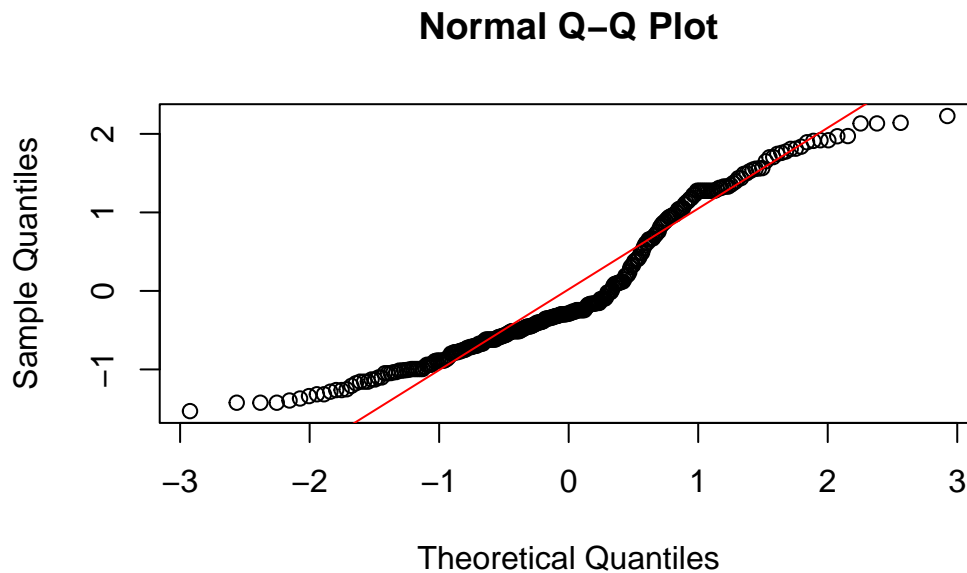
```
# QQ-plot for normality of residuals of Model 1  
qqnorm(residuals(H3_M1_RI))  
qqline(residuals(H3_M1_RI), col = "red")
```

**Normal Q-Q Plot**





```
# QQ-plot for normality of residuals of Model 2
qqnorm(residuals(H3_M2_RI))
qqline(residuals(H3_M2_RI), col = "red")
```



- **Both Models:**

- The QQ-plots for both models do not indicate that the residuals are normally distributed.
- Because of this failure of this assumption, we will conduct a bootstrapped test of the predictor for both models.

3. **Multicollinearity:** Not applicable because each model only has one predictor.

## Summarizing the Models

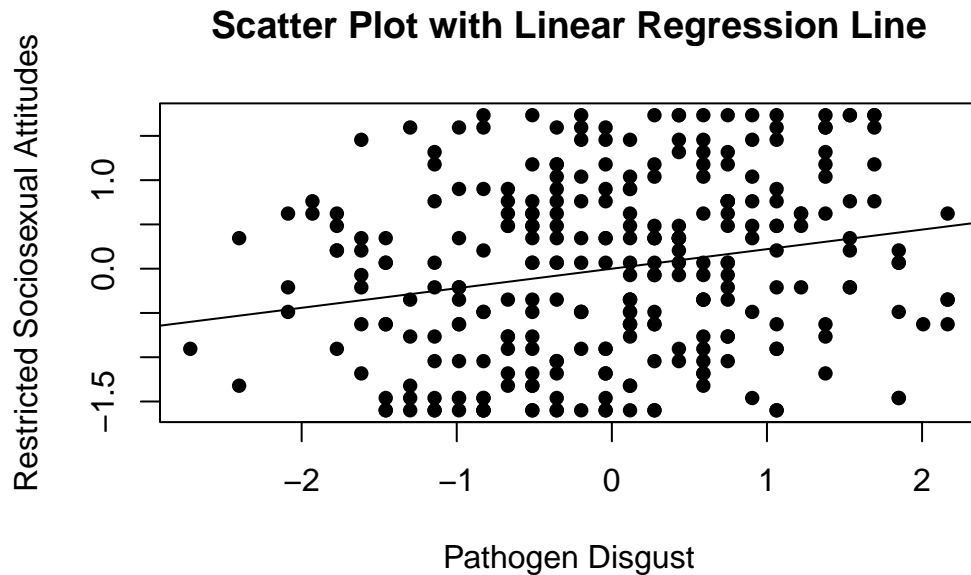
### Model 1

```
# Create a scatterplot with the regression line to visualize
plot(data$TDDS_p_z, data$SOI_a_r_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Pathogen Disgust",
     ylab = "Restricted Sociosexual Attitudes",
```

```

    pch = 16, # For solid circle points
    col = "black")
# Add regression line
abline(H3_M1_RI, col = "black", lwd = 1)

```



```

# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_M1_RI <- Boot(H3_M1_RI, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H3_M1_RI) # Getting the coefficients

```

```

Number of bootstrap replications R = 1000
      original bootBias  bootSE  bootMed
(Intercept) -1.1616e-17 0.0012152 0.059909 0.0021975

```

```
TDDS_p_z      2.2087e-01 0.0015469 0.059225 0.2236528
```

```
confint(boot_H3_M1_RI) # Getting the BCIs
```

Bootstrap bca confidence intervals

```
          2.5 %    97.5 %  
(Intercept) -0.1256230 0.1076692  
TDDS_p_z      0.1039745 0.3299802
```

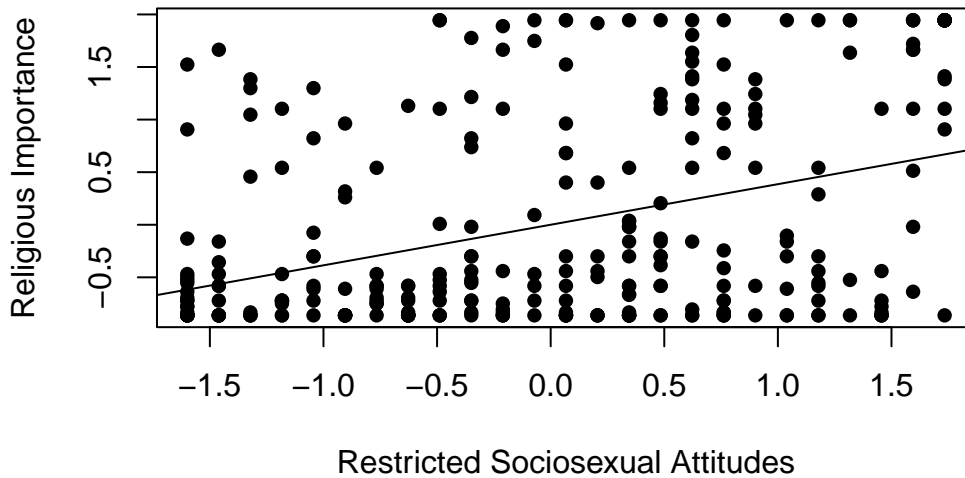
```
# Removing the vectors in the environment that the Boot package required outside of the data  
rm(SOI_a_r_z, TDDS_p_z)
```

- We can see from the Scatter Plot that there is a positive relationship between pathogen disgust and restricted sociosexual attitudes.
- Based on the estimates, the standardized coefficient for pathogen disgust is around .22.
  - In addition, this relationship is significant, as the BCI excludes zero (upper = .33, lower = .10).

## Model 2

```
# Create a scatterplot with the regression line to visualize  
plot(data$SOI_a_r_z, data$religious_importance_z,  
      main = "Scatter Plot with Linear Regression Line",  
      xlab = "Restricted Sociosexual Attitudes",  
      ylab = "Religious Importance",  
      pch = 16, # For solid circle points  
      col = "black")  
# Add regression line  
abline(H3_M2_RI, col = "black", lwd = 1)
```

## Scatter Plot with Linear Regression Line



```
# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
religious_importance_z <- as.numeric(data$religious_importance_z)
```

```
# Setting the seed
set.seed(1042557)
```

```
# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_M2_RI <- Boot(H3_M2_RI, R = 1000) # 1,000 resamples
```

```
# Summarizing the estimates and CIs
summary(boot_H3_M2_RI) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
```

	original	bootBias	bootSE	bootMed
(Intercept)	-1.7538e-18	-0.00046013	0.054763	-0.00017089
SOI_a_r_z	3.8592e-01	-0.00012983	0.056140	0.38640935

```
confint(boot_H3_M2_RI) # Getting the BCIs
```

```
Bootstrap bca confidence intervals
```

```

                2.5 %    97.5 %
(Intercept) -0.1039343 0.1170591
SOI_a_r_z    0.2740123 0.4924135

```

```

# Removing the vectors in the environment that the Boot package required outside of the data
rm(SOI_a_r_z, religious_importance_z)

```

- We can see from the Scatter Plot that there is a positive trend in the relationship between restricted sociosexual attitudes and religious importance.
- Based on the estimates, the standardized coefficient for restricted sociosexual attitudes is around .39.
  - In addition, this relationship is significant, as the BCI excludes zero (upper = .49, lower = .27).

## Full Mediation Model

The fact that pathogen disgust predicts a monogamous mating strategy and, in turn, a monogamous mating strategy predicts religious importance indicates that a monogamous mating strategy could potentially mediate the relationship between pathogen disgust and religious importance. Now we will run the mediation analysis.

```

# Fitting the full mediation model
H3_full_RI <- lm(religious_importance_z ~ TDDS_p_z + SOI_a_r_z, data = data)

```

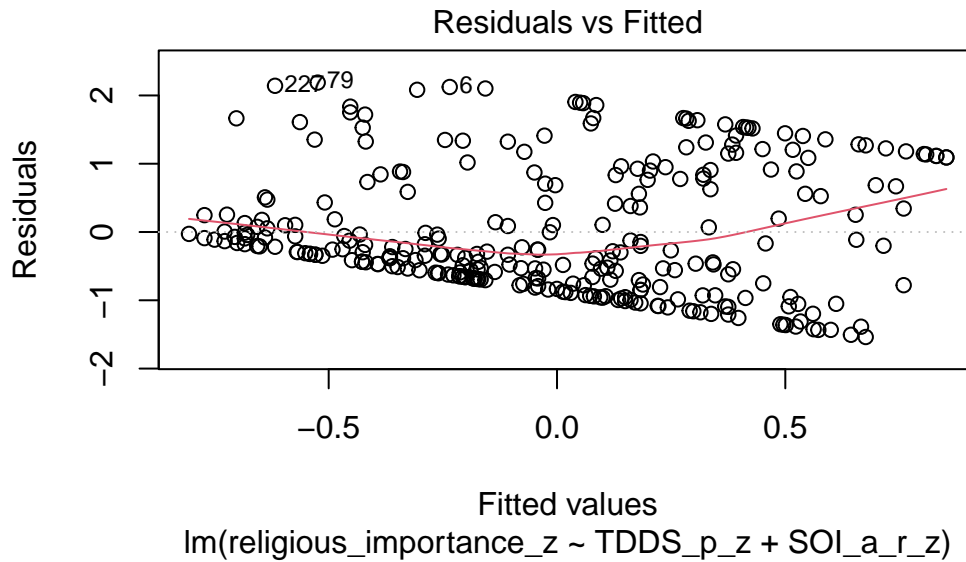
## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```

# Plotting the residual vs. fitted values
plot(H3_full_RI, which = 1)

```

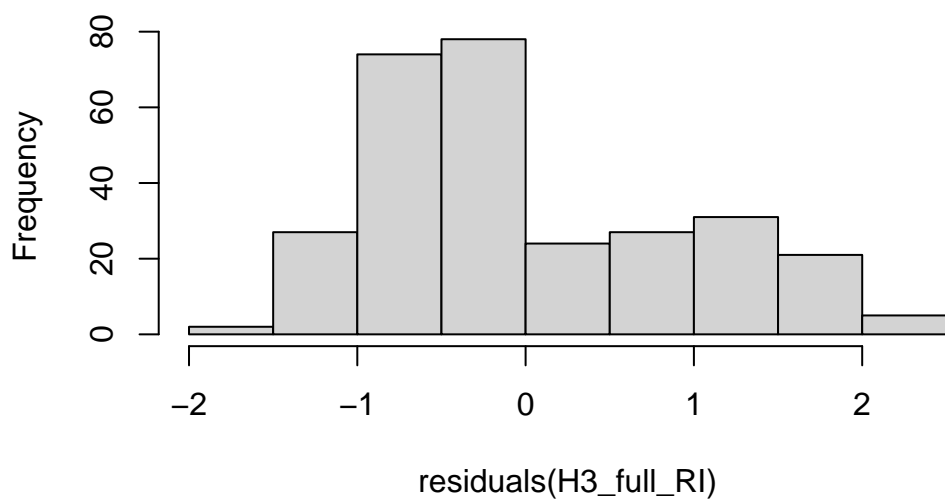


- There is a trend towards more negative residual values at the higher fitted values of the model. This may be problematic.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

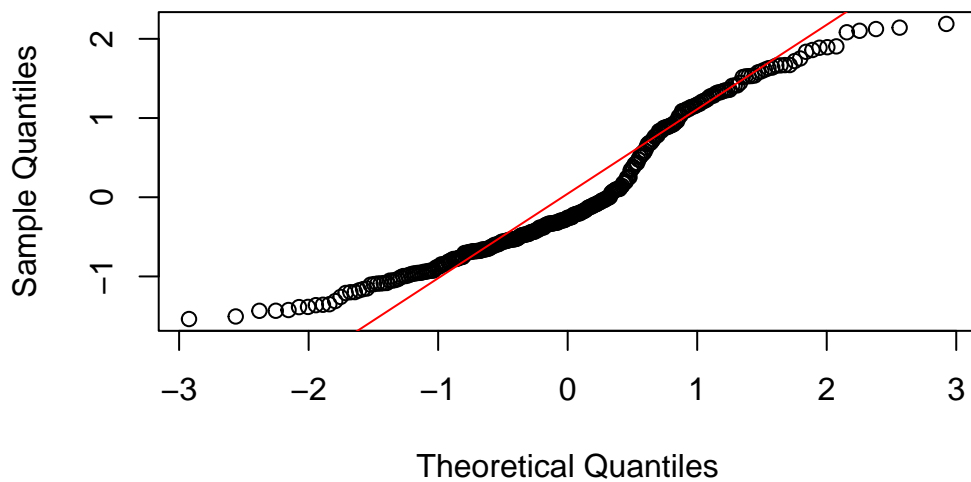
```
# Histogram for normality of residuals
hist(residuals(H3_full_RI))
```

**Histogram of residuals(H3\_full\_RI)**



```
# QQ-plot for normality of residuals of the full model
qqnorm(residuals(H3_full_RI))
qqline(residuals(H3_full_RI), col = "red")
```

**Normal Q-Q Plot**



- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 SD.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictors for the model.

### 3. Multicollinearity: Variance Inflation Factor (VIF) from `car` package

```
# Calculating VIF
vif(H3_full_RI)

TDDS_p_z SOI_a_r_z
1.051286  1.051286
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
SOI_a_r_z <- as.numeric(data$SOI_a_r_z)
religious_importance_z <- as.numeric(data$religious_importance_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H3_full_RI <- Boot(H3_full_RI, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H3_full_RI) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original    bootBias   bootSE   bootMed
(Intercept) -7.6035e-18 -0.00082361 0.054498 -0.0035079
TDDS_p_z     1.4015e-01  0.00033273 0.055713  0.1394299
SOI_a_r_z     3.5496e-01 -0.00063970 0.057809  0.3543830
```

```
confint(boot_H3_full_RI) # Getting the BCIs
```



Bootstrap bca confidence intervals

	2.5 %	97.5 %
(Intercept)	-0.09587216	0.1273944
TDDS_p_z	0.03656495	0.2534092
SOI_a_r_z	0.23746177	0.4667500

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(SOI_a_r_z, religious_importance_z, TDDS_p_z)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H3_RI_mediate_results <- mediate(model.m = H3_M1_RI,
  model.y = H3_full_RI,
  treat = "TDDS_p_z",
  mediator = "SOI_a_r_z",
  boot = TRUE, # Bootstrap the CI for the indirect effect
  sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results of the mediation analysis
summary(H3_RI_mediate_results)
```

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	0.0784	0.0368	0.13	<2e-16 ***
ADE	0.1402	0.0376	0.25	0.002 **
Total Effect	0.2186	0.1120	0.34	<2e-16 ***
Prop. Mediated	0.3587	0.1757	0.71	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- Multiple Regression Results:

- Based on the bootstrapped CIs, both pathogen disgust and restricted sociosexual attitudes are significantly positively related to religious importance, with the unique effect of restricted sociosexual attitudes being substantially larger ( $B = .355$ ,  $BCI = [.237, .467]$ ) than the unique effect of pathogen disgust ( $B = .140$ ,  $BCI = [.036, .253]$ ).

- **Mediation Analysis Results:**

- The ACME indicates that there is a significant indirect effect of pathogen disgust on religious importance through a monogamous mating strategy ( $B = .078$ ,  $BCI = [.037, .13]$ ).
  - \* Given this result, we have evidence for the weak version of Hypothesis 3. That is, the relationship between pathogen disgust and religiosity is partially mediated by a monogamous mating strategy.
- The proportion of the total effect mediated (the ratio of the indirect effect to the direct effect) is .36. That is, 36% of the covariance between pathogen disgust and religious importance seems to be able to be accounted for by restricted sociosexual attitudes. Our inferential criterion for concluding full mediation was that this ratio would be at least .80, with the total effect greater than a  $B = .2$ . Even the upper bound of the bootstrapped confidence interval for this proportion mediated does not cross this threshold ( $BCI = [.176, .71]$ ).
  - \* Given this result, we do not have evidence for the strong version of Hypothesis 3. That is, the relationship between pathogen disgust and religiosity is not fully mediated by a monogamous mating strategy.

## **Hypothesis 2: Removing GENE Preferences Item 1**

In the [Internal Reliability](#) analysis of the [Preferences Subscale of SFGENE-7](#), the drop alpha statistics indicated that removing Item 1 from the subscale would improve the internal reliability from an alpha of around .5 to around .8. Given that a lack of internal reliability may attenuate relationships between variables in correlational analyses, I will now conduct the same analysis for Hypothesis 2 but with Item 1 removed.

### **Preliminary Simple Linear Regression Models**

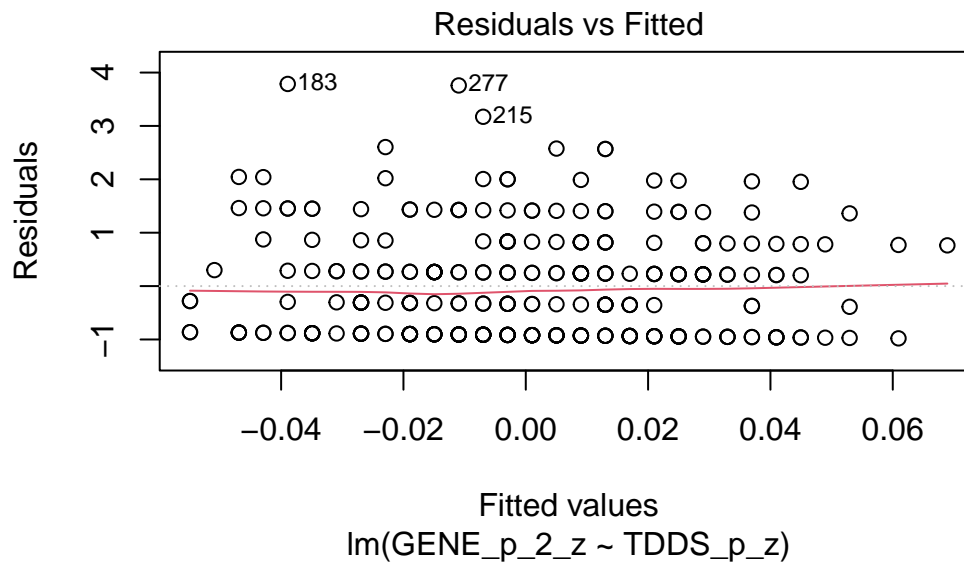
As mentioned above, to test Hypothesis 2, we must first establish that pathogen disgust affects out-group avoidance (i.e., `GENE_p_2_z`) and that out-group avoidance influences religiosity.

```
# Fitting the models
H2_M1_2 <- lm(GENE_p_2_z ~ TDDS_p_z, data = data)
H2_M2_2 <- lm(CRS_z ~ GENE_p_2_z, data = data)
```

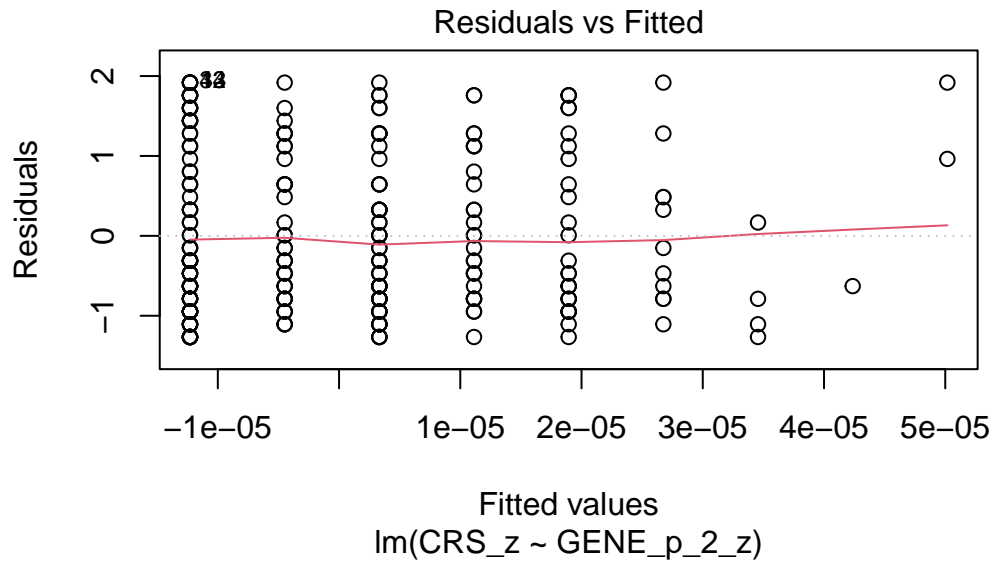
## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```
# Plotting the residual vs. fitted values
plot(H2_M1_2, which = 1) # For model 1
```



```
plot(H2_M2_2, which = 1) # For model 2
```

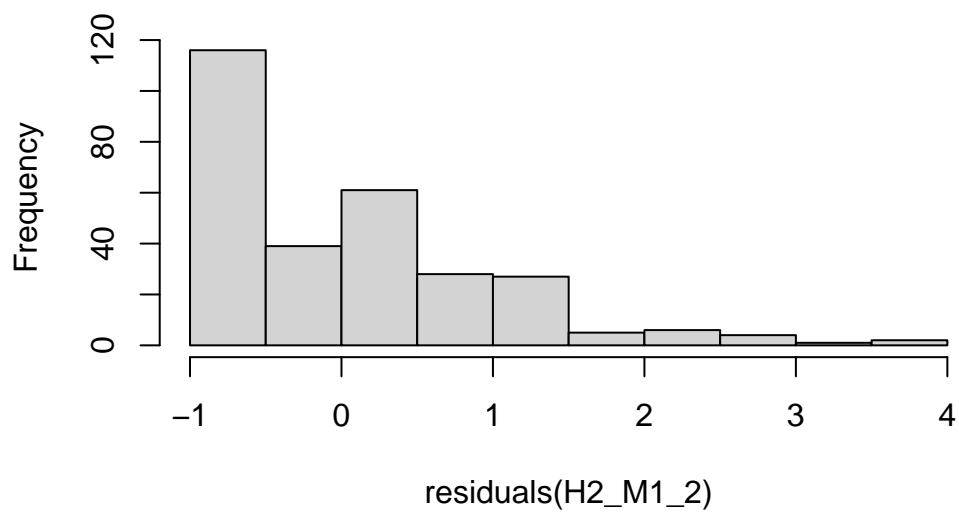


- The residuals vs. fitted plot does not stand out as violating the assumption of linearity or heteroscedasticity for the either model.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

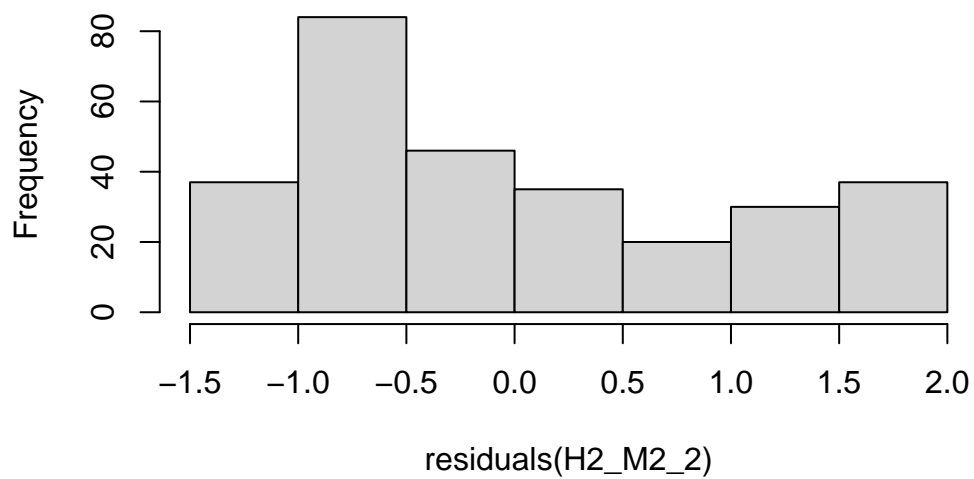
```
# Histogram for normality of residuals
hist(residuals(H2_M1_2)) # For model 1
```

**Histogram of residuals(H2\_M1\_2)**



```
hist(residuals(H2_M2_2)) # For model 2
```

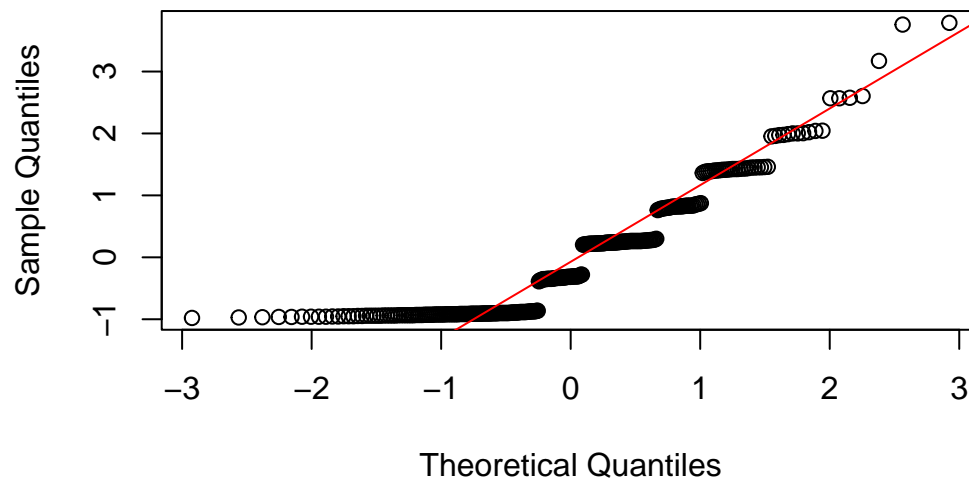
**Histogram of residuals(H2\_M2\_2)**



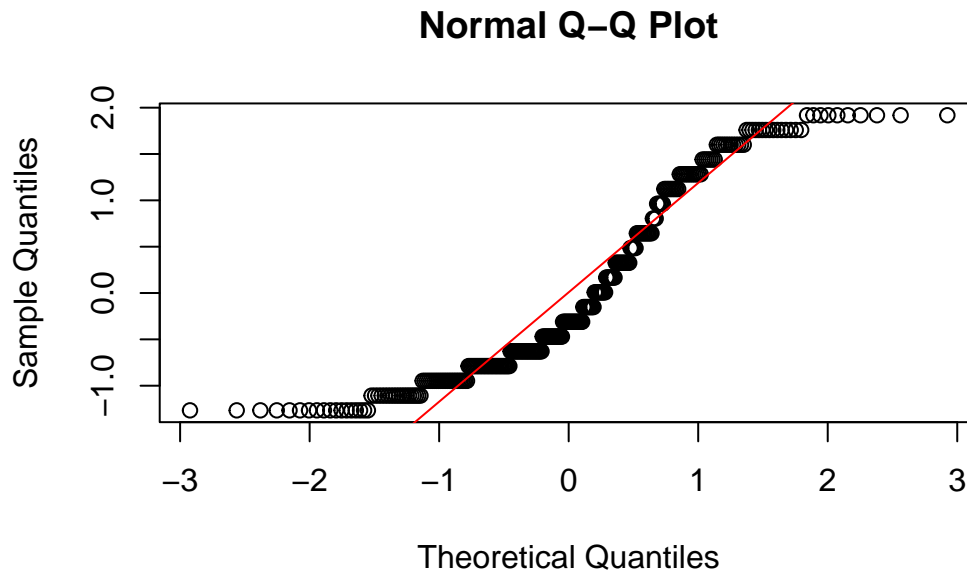
```
# QQ-plot for normality of residuals of Model 1  
qqnorm(residuals(H2_M1_2))
```

```
qqline(residuals(H2_M1_2), col = "red")
```

### Normal Q-Q Plot



```
# QQ-plot for normality of residuals of Model 2  
qqnorm(residuals(H2_M2_2))  
qqline(residuals(H2_M2_2), col = "red")
```



- **Both Models:**

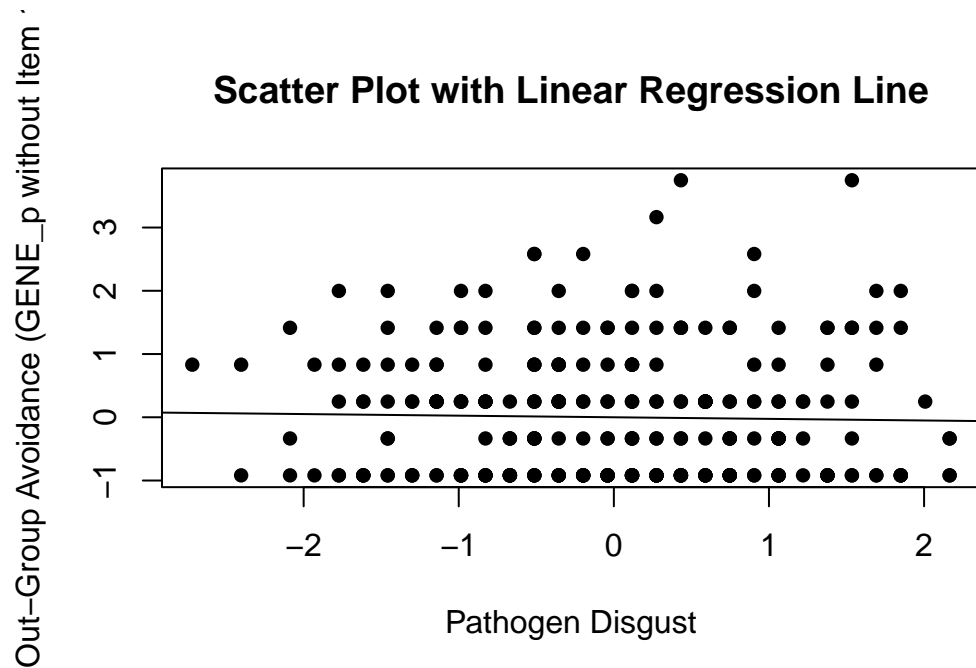
- The histograms and QQ-plots indicate that for both models the residuals are not normally distributed, although in different ways.
- We will need to bootstrapping for testing the predictors for these models.

3. **Multicollinearity:** Not applicable because each model only has one predictor.

## Summarizing the Models

### Model 1

```
# Create a scatterplot with the regression line to visualize
plot(data$TDDS_p_z, data$GENE_p_2_z,
     main = "Scatter Plot with Linear Regression Line",
     xlab = "Pathogen Disgust",
     ylab = "Out-Group Avoidance (GENE_p without Item 1)",
     pch = 16, # For solid circle points
     col = "black")
# Add regression line
abline(H2_M1_2, col = "black", lwd = 1)
```



```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_2_z <- as.numeric(data$GENE_p_2_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_M1_2 <- Boot(H2_M1_2, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H2_M1_2) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
      original  bootBias  bootSE  bootMed
(Intercept)  7.0301e-16 -0.0014508 0.058715 -0.0040643
TDDS_p_z     -2.5360e-02 -0.0020741 0.059085 -0.0284347
```

```
confint(boot_H2_M1_2) # Getting the BCIs
```

```
Bootstrap bca confidence intervals
```



	2.5 %	97.5 %
(Intercept)	-0.1071551	0.1267461
TDDS_p_z	-0.1333744	0.1058820

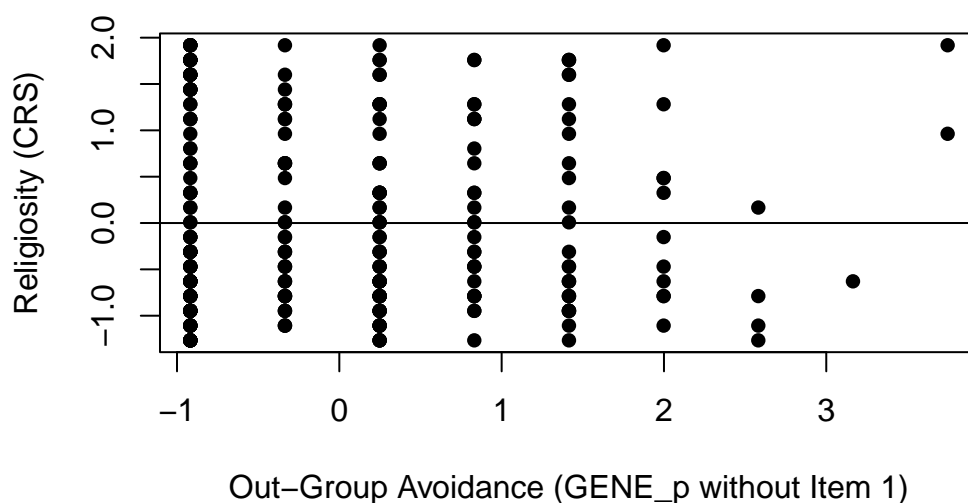
```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_2_z, TDDS_p_z)
```

- We can see from the Scatter Plot that there is no relationship between pathogen disgust and out-group avoidance to speak of ( $B = -.025$ ,  $BCI = [-.133, .106]$ )—just as with the analysis above.
  - Although this precludes the possibility that out-group avoidance mediates the relationship between disgust and religiosity, we will continue with the analysis for completions sake.

## Model 2

```
# Create a scatterplot with the regression line to visualize
plot(data$GENE_p_2_z, data$CRS_z,
      main = "Scatter Plot with Linear Regression Line",
      xlab = "Out-Group Avoidance (GENE_p without Item 1)",
      ylab = "Religiosity (CRS)",
      pch = 16, # For solid circle points
      col = "black")
# Add regression line
abline(H2_M2_2, col = "black", lwd = 1)
```

## Scatter Plot with Linear Regression Line



```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_2_z <- as.numeric(data$GENE_p_2_z)
CRS_z <- as.numeric(data$CRS_z)
```

```
# Setting the seed
set.seed(1042557)
```

```
# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_M2_2 <- Boot(H2_M2_2, R = 1000) # 1,000 resamples
```

```
# Summarizing the estimates and CIs
summary(boot_H2_M2_2) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original    bootBias    bootSE    bootMed
(Intercept) 2.6680e-17 -0.00011423 0.060017  0.0015515
GENE_p_2_z   1.3388e-05 -0.00097717 0.064575 -0.0015445
```

```
confint(boot_H2_M2_2) # Getting the BCIs
```

```
Bootstrap bca confidence intervals
```

```

                2.5 %    97.5 %
(Intercept) -0.1135425 0.1163811
GENE_p_2_z  -0.1190670 0.1228603

```

```

# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_2_z, CRS_z)

```

- We can see from the Scatter Plot that there is a positive trend in the relationship between out-group avoidance and religiosity.
- Based on the estimates, the standardized coefficient for out-group avoidance is around  $B = .00001$  ( $BCI = [-.119, .123]$ ).

## Full Mediation Model

The fact that pathogen disgust does not predict out-group avoidance and, in turn, out-group avoidance does not predict religiosity indicates that out-group avoidance could not mediate the relationship between pathogen disgust and religiosity, but I will run the analysis that we planned anyway just to have the results.

```

# Fitting the full mediation model
H2_2_full <- lm(CRS_z ~ TDDS_p_z + GENE_p_2_z, data = data)

```

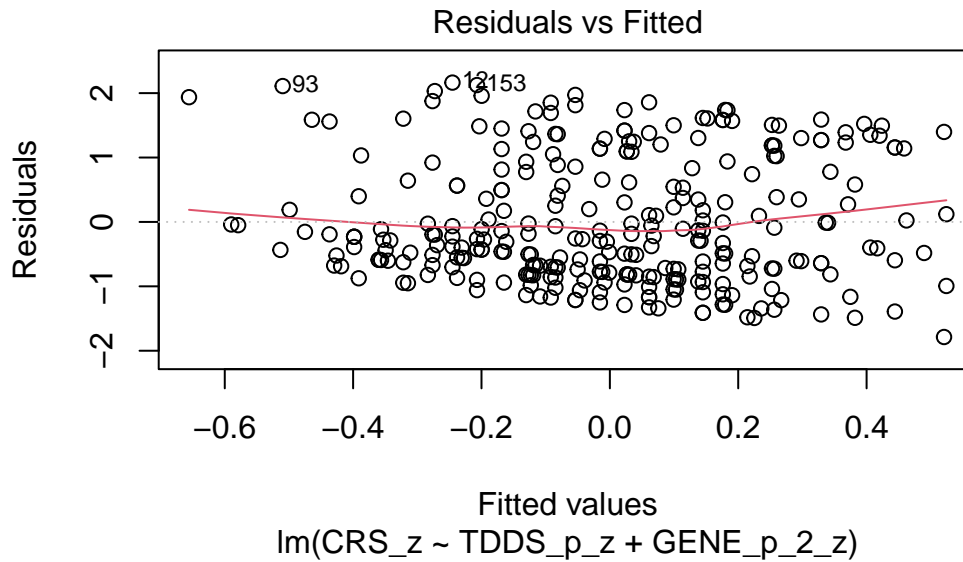
## Assumptions

1. **Linearity and Heteroscedasticity:** Assessed visually through plotting residual vs. fitted values

```

# Plotting the residual vs. fitted values
plot(H2_2_full, which = 1)

```

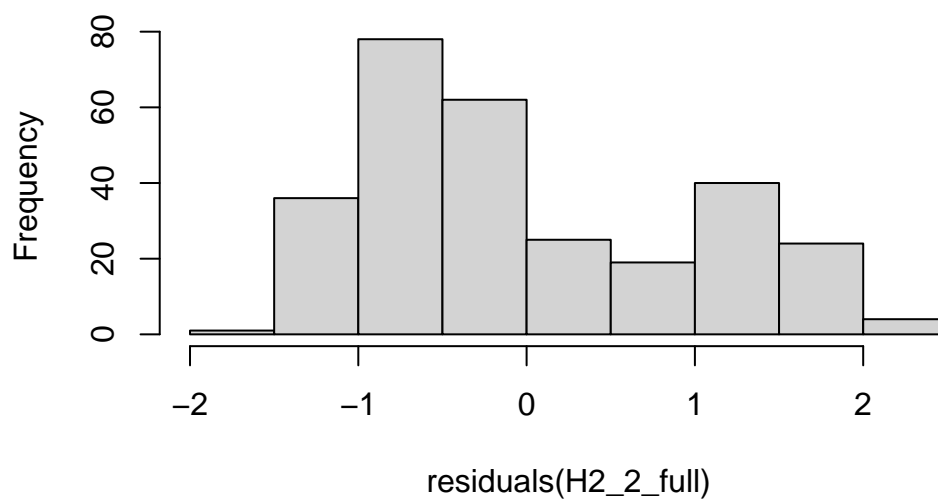


- There is a trend towards more negative residual values at the higher fitted values of the model. This may be problematic.

2. **Normality:** Histogram and QQ-plot (because Shapiro-Wilk test would be overpowered)

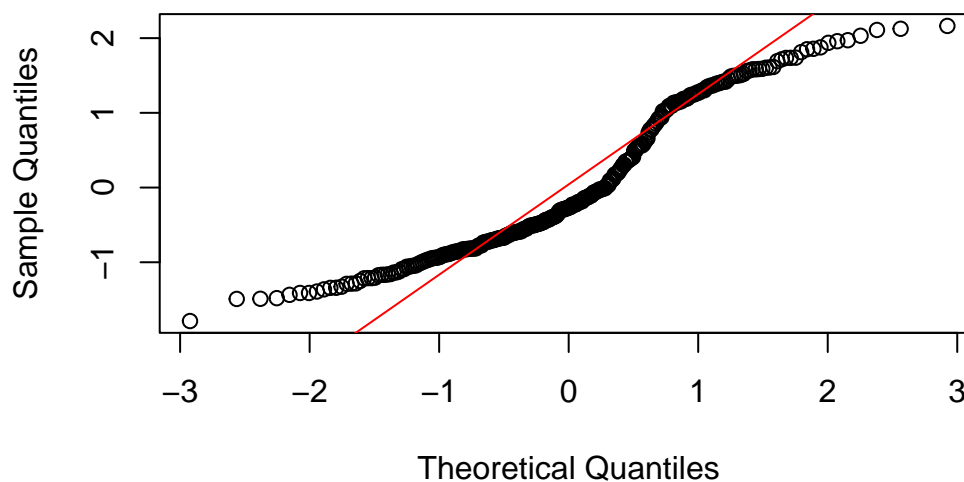
```
# Histogram for normality of residuals
hist(residuals(H2_2_full))
```

**Histogram of residuals(H2\_2\_full)**



```
# QQ-plot for normality of residuals of the full model
qqnorm(residuals(H2_2_full))
qqline(residuals(H2_2_full), col = "red")
```

**Normal Q-Q Plot**



- The histogram and QQ-plot of the residuals of the model indicate that the residuals are not normally distributed. That is, there is a strong tendency for higher frequencies of residuals at both just above -1 SD and just above 1 SD. The two clusters of similar errors in prediction may be due to sex differences in religiosity. In exploratory analyses, we will add **sex** as a control variable.
  - Because of this failure of this assumption, we will conduct a bootstrapped test of the predictors for the model.

### 3. Multicollinearity: Variance Inflation Factor (VIF) from **car** package

```
# Calculating VIF
vif(H2_2_full)
```

```
TDDS_p_z GENE_p_2_z
1.000644 1.000644
```

- The VIF is good here.

### Summarizing and Testing for Partial/Full Mediation

```
# Adding the variables to the broader environment so the Boot function will run
GENE_p_2_z <- as.numeric(data$GENE_p_2_z)
CRS_z <- as.numeric(data$CRS_z)
TDDS_p_z <- as.numeric(data$TDDS_p_z)

# Setting the seed
set.seed(1042557)

# Getting the bootstrapped estimates and CIs (using car package rather than boot for simplicity)
boot_H2_2_full <- Boot(H2_2_full, R = 1000) # 1,000 resamples

# Summarizing the estimates and CIs
summary(boot_H2_2_full) # Getting the coefficients
```

```
Number of bootstrap replications R = 1000
              original    bootBias  bootSE  bootMed
(Intercept) 1.2339e-17 -0.00077435 0.058551 -0.0013948
TDDS_p_z    2.4314e-01 -0.00072365 0.059883 0.2433389
GENE_p_2_z   6.1794e-03 -0.00060177 0.060073 0.0069054
```

```
confint(boot_H2_2_full) # Getting the BCIs
```

Bootstrap bca confidence intervals

```

                2.5 %    97.5 %
(Intercept) -0.1079394 0.1232773
TDDS_p_z     0.1198619 0.3565631
GENE_p_2_z   -0.1149113 0.1168755

```

```
# Removing the vectors in the environment that the Boot package required outside of the data
rm(GENE_p_2_z, CRS_z, TDDS_p_z)
```

```
# Using the mediate function from the mediation package to estimate the indirect effect with
H2_2_mediate_results <- mediate(model.m = H2_M1_2,
                                model.y = H2_2_full,
                                treat = "TDDS_p_z",
                                mediator = "GENE_p_2_z",
                                boot = TRUE, # Bootstrap the CI for the indirect effect
                                sims = 1000) # Do 1000 resamples
```

Running nonparametric bootstrap

```
# Printing the results of the mediation analysis
summary(H2_2_mediate_results)
```

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	-0.000157	-0.006528	0.01	0.99
ADE	0.243143	0.131919	0.36	<2e-16 ***
Total Effect	0.242987	0.133597	0.36	<2e-16 ***
Prop. Mediated	-0.000645	-0.031489	0.04	0.99

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 289

Simulations: 1000

- **Multiple Regression Results:**

- Based on the bootstrapped CIs, pathogen disgust, but not out-group avoidance, is significantly positively related to religiosity.
- The coefficients for pathogen disgust (.24) and out-group avoidance (.006) are similar to their values with Item 1 included.

- **Mediation Analysis Results:**

- The ACME indicates that there is no significant indirect effect of pathogen disgust on religiosity through out-group avoidance ( $B = -.0002$ ,  $BCI = [-.006, .01]$ ).
- This result is inconsistent with both the weak and strong version version of Hypothesis 2. That is, there is no evidence here that the relationship between pathogen disgust and religiosity is mediated by out-group avoidance.

## Saving Data

Now we will save the final version of the data frame (`data`) as a `data/final_data.csv`.

```
# Saving the data
write.csv(data, "./data/final_data.csv")
```

## References

- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Beribisky, N., Mara, C. A., & Cribbie, R. A. (2020). An equivalence testing approach for evaluating substantial mediation. *The Quantitative Methods for Psychology*, 16(5), 424–441. <https://doi.org/10.20982/tqmp.16.4.p424>
- Duckitt, J., Bizumic, B., Krauss, S. W., & Heled, E. (2010). A Tripartite Approach to Right-Wing Authoritarianism: The Authoritarianism-Conservatism-Traditionalism Model. *Political Psychology*, 31(5), 685–715. <https://doi.org/10.1111/j.1467-9221.2010.00781.x>
- Dunwoody, P. T., & Funke, F. (2016). The Aggression-Submission-Conventionalism Scale: Testing a New Three Factor Measure of Authoritarianism. *Journal of Social and Political Psychology*, 4(2), Article 2. <https://doi.org/10.5964/jspp.v4i2.168>



- Huber, S., & Huber, O. W. (2012). The Centrality of Religiosity Scale (CRS). *Religions*, 3(3), Article 3. <https://doi.org/10.3390/rel3030710>
- Neuliep, J. W. (2002). Assessing the Reliability and Validity of the Generalized Ethnocentrism Scale. *Journal of Intercultural Communication Research*, 31(4), 201. <https://twu.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=ufh&AN=10966365&site=eds-live&scope=site>
- Neuliep, J. W., & McCroskey, J. C. (1997). The development of a US and Generalized Ethnocentrism Scale. *Communication Research Reports*, 14(4), 385–398. <https://doi.org/10.1080/08824099709388682>
- Olatunji, B., Adams, T., Ciesielski, B., David, B., Sarawgi, S., & Broman-Fulks, J. (2012). The Three Domains of Disgust Scale. *Assessment*, 19, 205–225. <https://doi.org/10.1177/1073191111432881>
- Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships: Journal of Personality and Social Psychology. *Journal of Personality and Social Psychology*, 95(5), 1113–1135. <https://doi.org/10.1037/0022-3514.95.5.1113>
- Simpson, J. A., & Gangestad, S. W. (1991). Individual differences in sociosexuality: evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology*, 60(6), 870–883. <https://doi.org/10.1037//0022-3514.60.6.870>
- Tybur, J. M., Lieberman, D. L., & Griskevicius, V. G. (2009). Microbes, mating, and morality: Individual differences in three functional domains of disgust. *Journal of Personality and Social Psychology*, 96, 103–122. <https://doi.org/10.1037/a0015474>
- Yu, Z., Bali, P., Tsikandilakis, M., & Tong, E. M. W. (2022). “Look not at what is contrary to propriety”: A meta-analytic exploration of the association between religiosity and sensitivity to disgust. *British Journal of Social Psychology*, 61(1), 276–299. <https://doi.org/10.1111/bjso.12479>