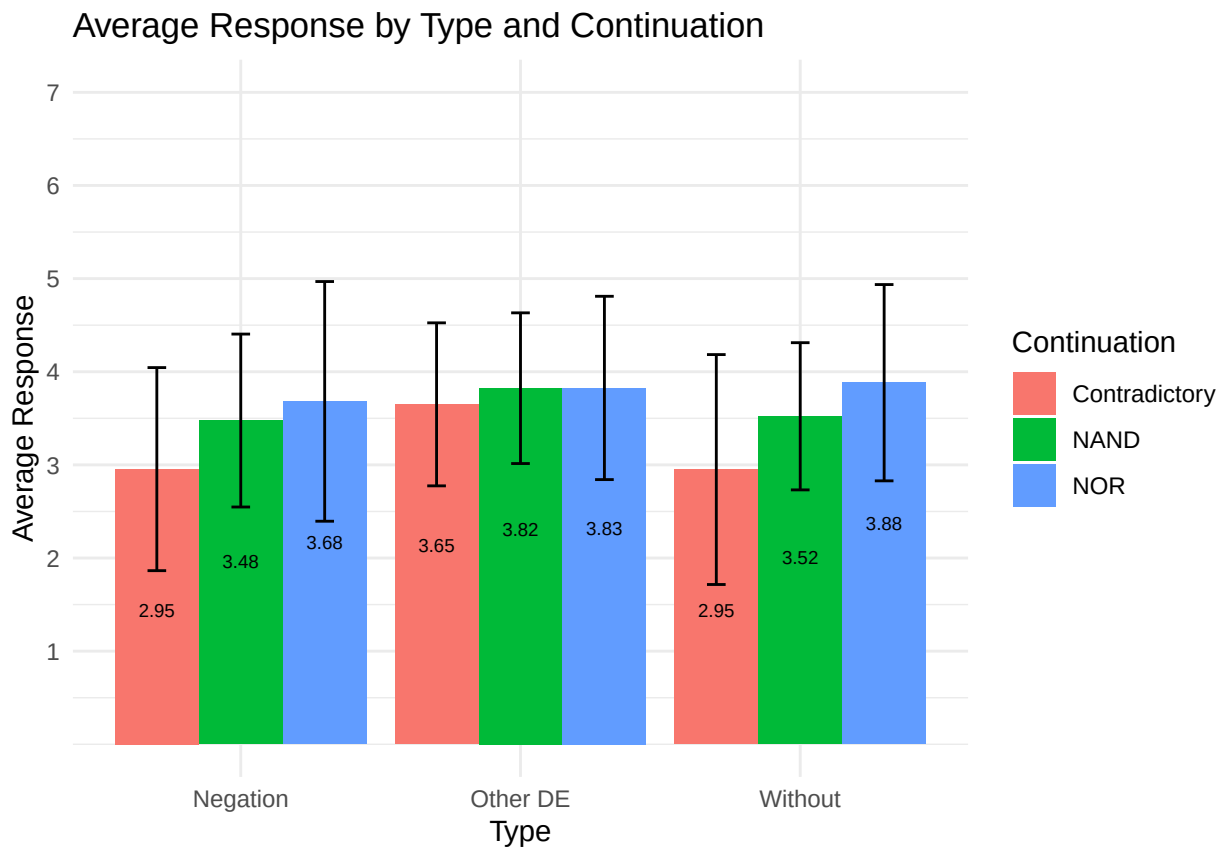
LLM Logic Scale Task Plots

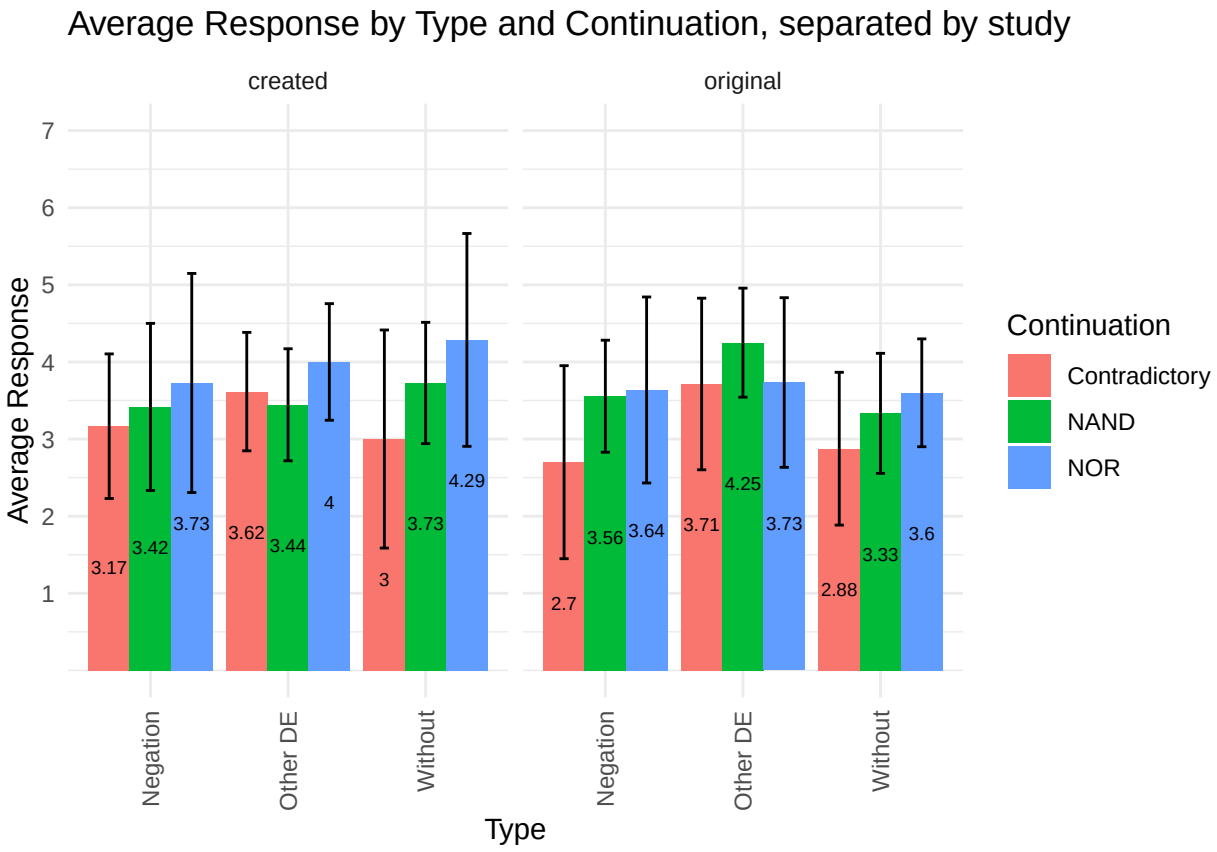
Carson Chiem

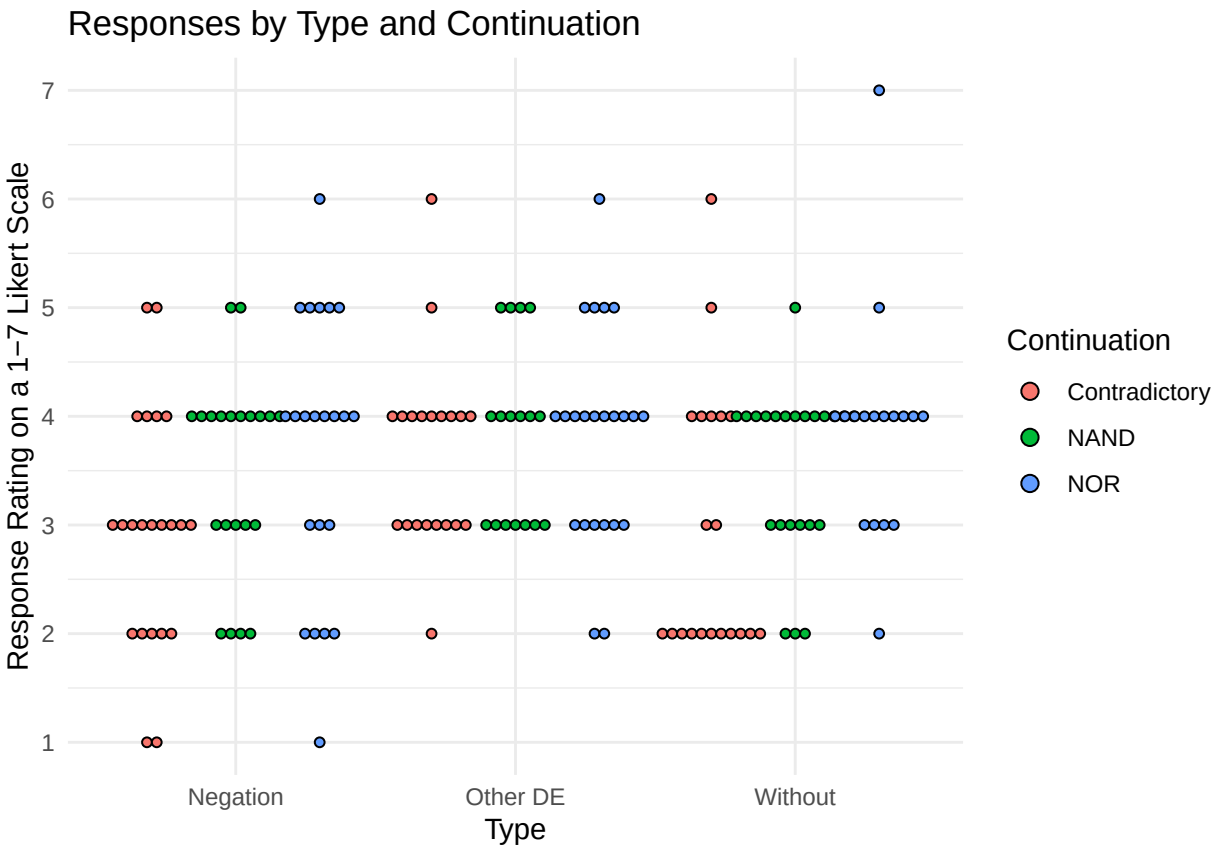
2024-01-22

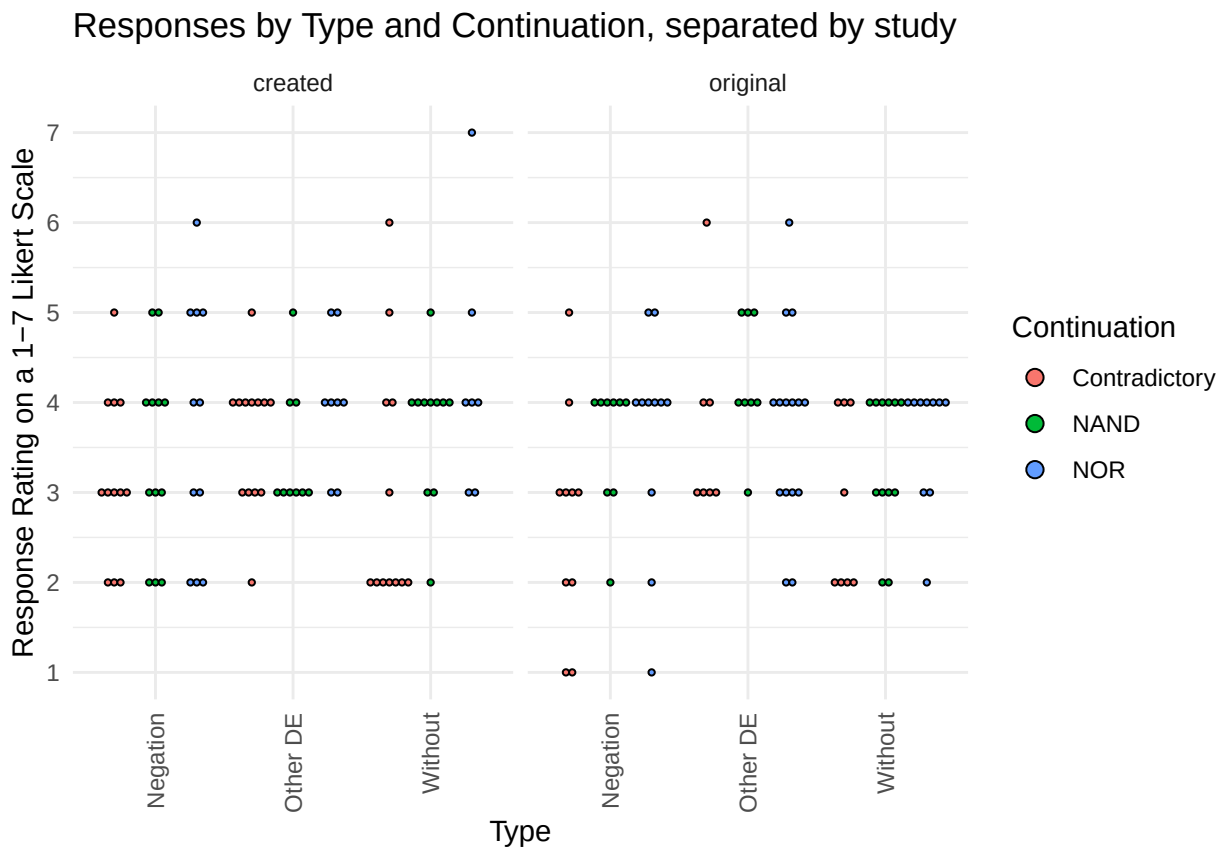
```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.4.4     v tibble    3.2.1
v lubridate 1.9.3     v tidyr     1.3.0
v purrr     1.0.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
Rows: 655 Columns: 7
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (5): Prompt, Continuation, Condition, Study, Type
dbl (2): Original order, Response

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.

`summarise()` has grouped output by 'Continuation', 'Type'. You can override
using the `.groups` argument.
```




```
‘summarise()‘ has grouped output by ’Continuation’, ’Type’, ’Condition’. You
can override using the ‘.groups‘ argument.
```


Average Response by Type and Continuation, separated by study
created
original
Average Response
7
6
5
4
3
2
1
3.17
3.42
3.73
3.62
3.44
4
3
3.73
4.29
2.7
3.56
3.64
3.71
4.25
3.73
2.88
3.33
3.6
Negation
Other DE
Without
Type
Continuation
Contradictory
NAND
NOR


Responses by Type and Continuation
Response Rating on a 1–7 Likert Scale
7
6
5
4
3
2
1
Negation
Other DE
Without
Type
Continuation
Contradictory
NAND
NOR



R Appendix

```
knitr::opts_chunk$set(echo = FALSE, comment = NA)
# importing packages and data
library(tidyverse)
dataset = read_csv("C:/CS_programs/Python/LLM_Logic/data/output_responses.csv")
# dataset cleaning

# filtering out control trials
df <- dataset %>%
  filter(Condition != "Control")

#dataset$Continuation[dataset$Condition == "Control"] = "Control"
# bar plot with created and original study prompts combined
# finding average responses for each grouping of variables
avg_responses = df %>%
  group_by(Continuation, Type, Condition) %>%
  summarise(avg_response = mean(Response), sd = sd(Response, na.rm = TRUE))

# plotting
ggplot(avg_responses, aes(x = Type, y = avg_response, fill = Continuation)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.9)) +
  geom_errorbar(aes(ymin = avg_response - sd, ymax = avg_response + sd), position = position_dodge(0.9)
  geom_text(aes(label = round(avg_response, 2)), position = position_dodge(width = 0.9), vjust = 11, si
  labs(title = "Average Response by Type and Continuation",
       x = "Type",
       y = "Average Response",
```

```
       fill = "Continuation") +
  scale_y_continuous(limits = c(0, 7), breaks = seq(1, 7, 1), labels = seq(1, 7, 1)) +
  theme_minimal()
# bar plot with separate graphs for original and created prompts
# finding average responses for each grouping of variables including study
avg_responses = df %>%
  group_by(Continuation, Type, Condition, Study) %>%
  summarise(avg_response = mean(Response), sd = sd(Response, na.rm = TRUE))

# plotting
ggplot(avg_responses, aes(x = Type, y = avg_response, fill = Continuation)) +
  geom_bar(stat = "identity", position = "dodge") +
  facet_wrap(~ Study, drop = TRUE, scales = "free_x") +
  geom_errorbar(aes(ymin = avg_response - sd, ymax = avg_response + sd), na.rm = TRUE, position = positi
  geom_text(aes(label = round(avg_response, 2)), position = position_dodge(width = 0.9), vjust = 11, si
  labs(title = "Average Response by Type and Continuation, separated by study",
       x = "Type",
       y = "Average Response",
       fill = "Continuation") +
  theme_minimal() +
  scale_y_continuous(limits = c(0, 7), breaks = seq(1, 7, 1), labels = seq(1, 7, 1)) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
# Dot plot
# plotting
ggplot(df, aes(x = Type, y = Response, fill = Continuation)) +
  geom_dotplot(binaxis = "y", stackdir = "center", position = "dodge", dotsize=.9, binwidth = .1) +
  labs(title = "Responses by Type and Continuation",
       x = "Type",
       y = "Response Rating on a 1-7 Likert Scale",
       fill = "Continuation") +
  theme_minimal() +
  scale_y_continuous(limits = c(1, 7), breaks = seq(1, 7, 1), labels = seq(1, 7, 1))
# Dot plot
# plotting
ggplot(df, aes(x = Type, y = Response, fill = Continuation)) +
  geom_dotplot(binaxis = "y", stackdir = "center", position = "dodge", dotsize=.7, binwidth = .1) +
  facet_wrap(~ Study, drop = TRUE, scales = "free_x") +
  labs(title = "Responses by Type and Continuation, separated by study",
       x = "Type",
       y = "Response Rating on a 1-7 Likert Scale",
       fill = "Continuation") +
  theme_minimal() +
  scale_y_continuous(limits = c(1, 7), breaks = seq(1, 7, 1), labels = seq(1, 7, 1)) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```