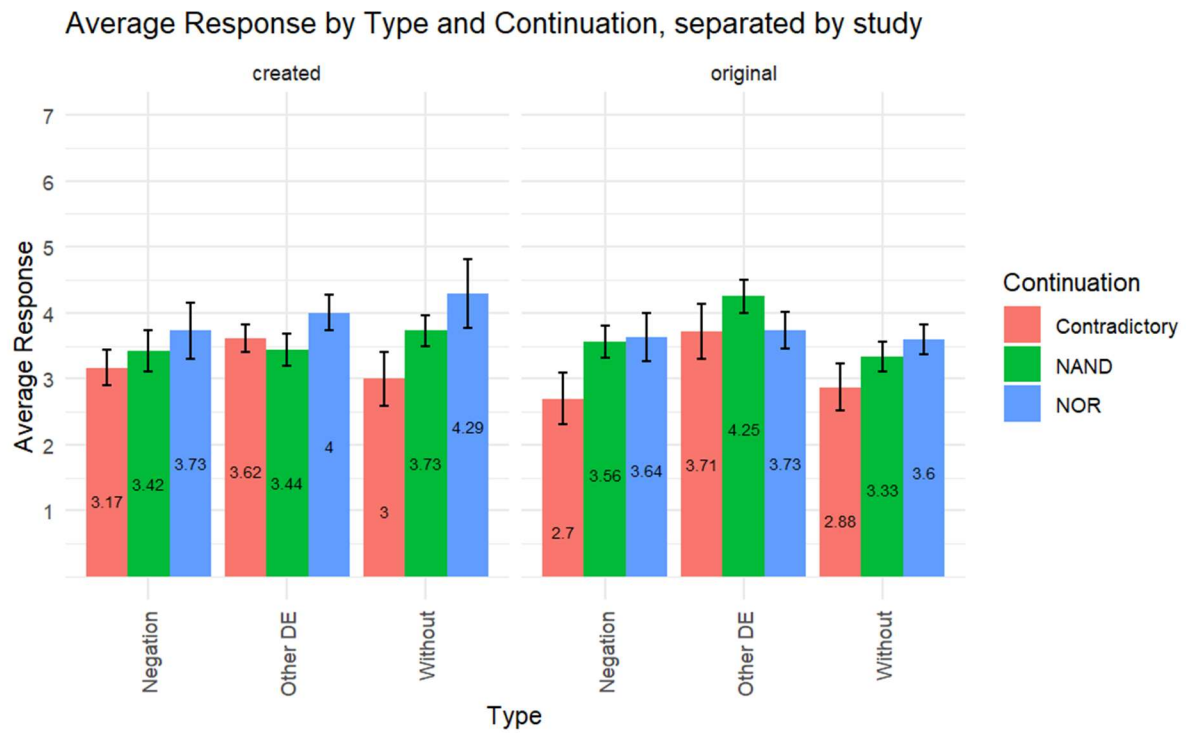
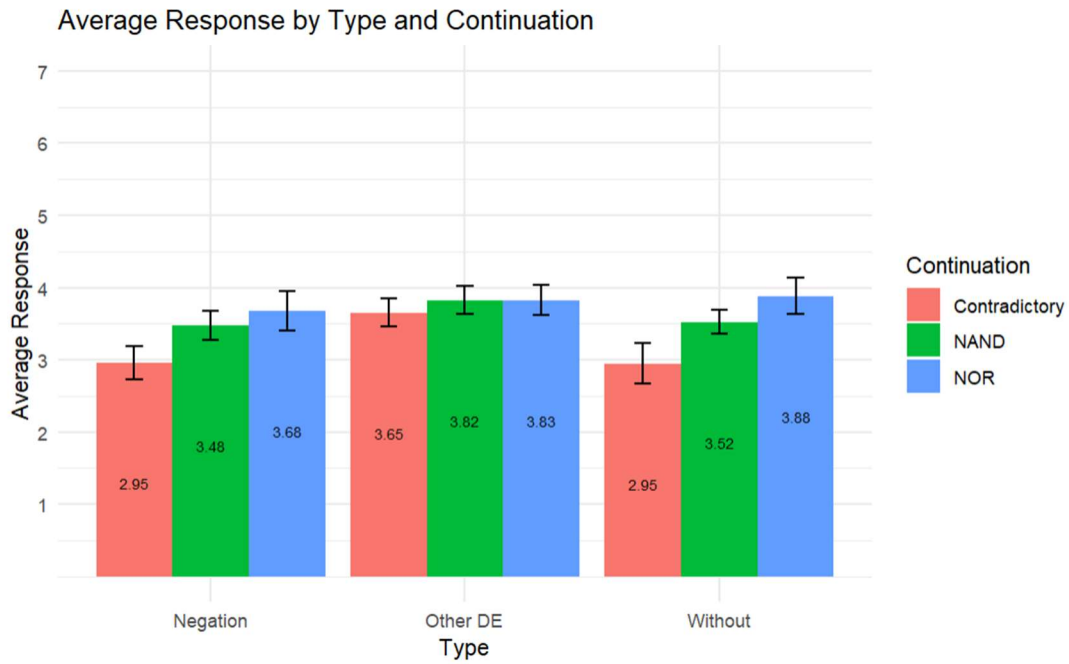
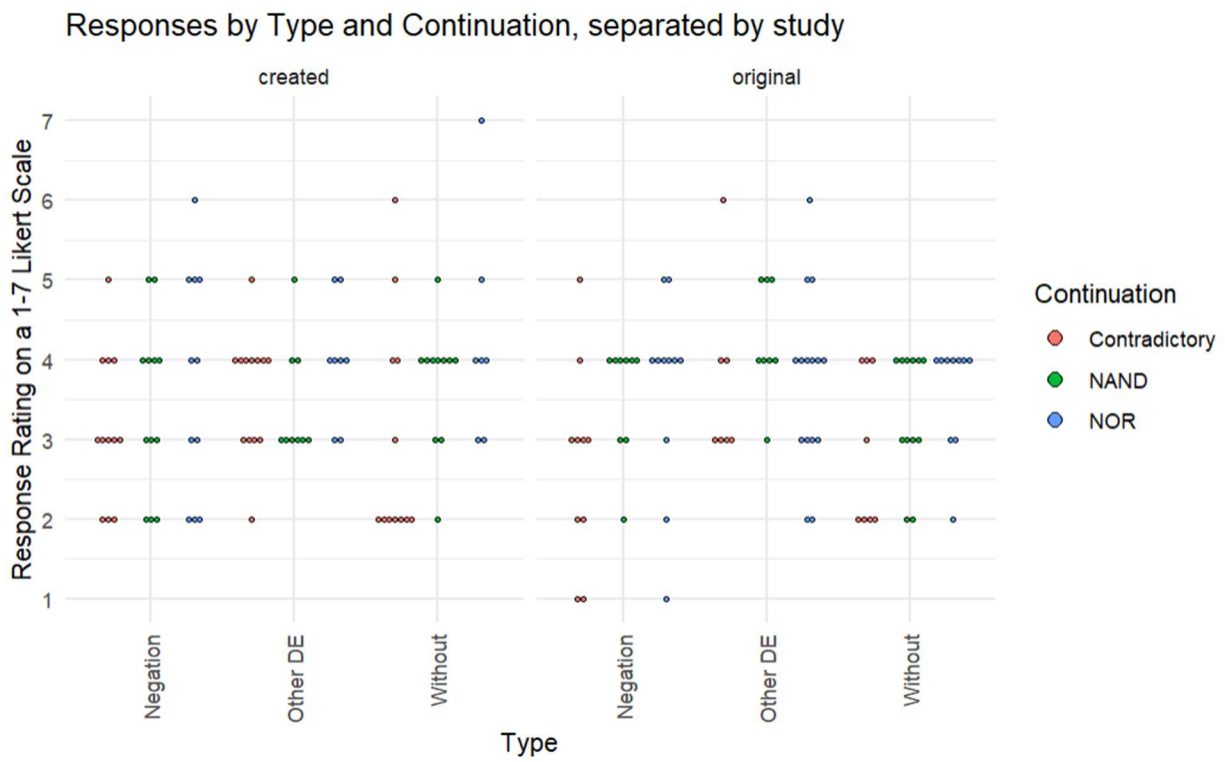
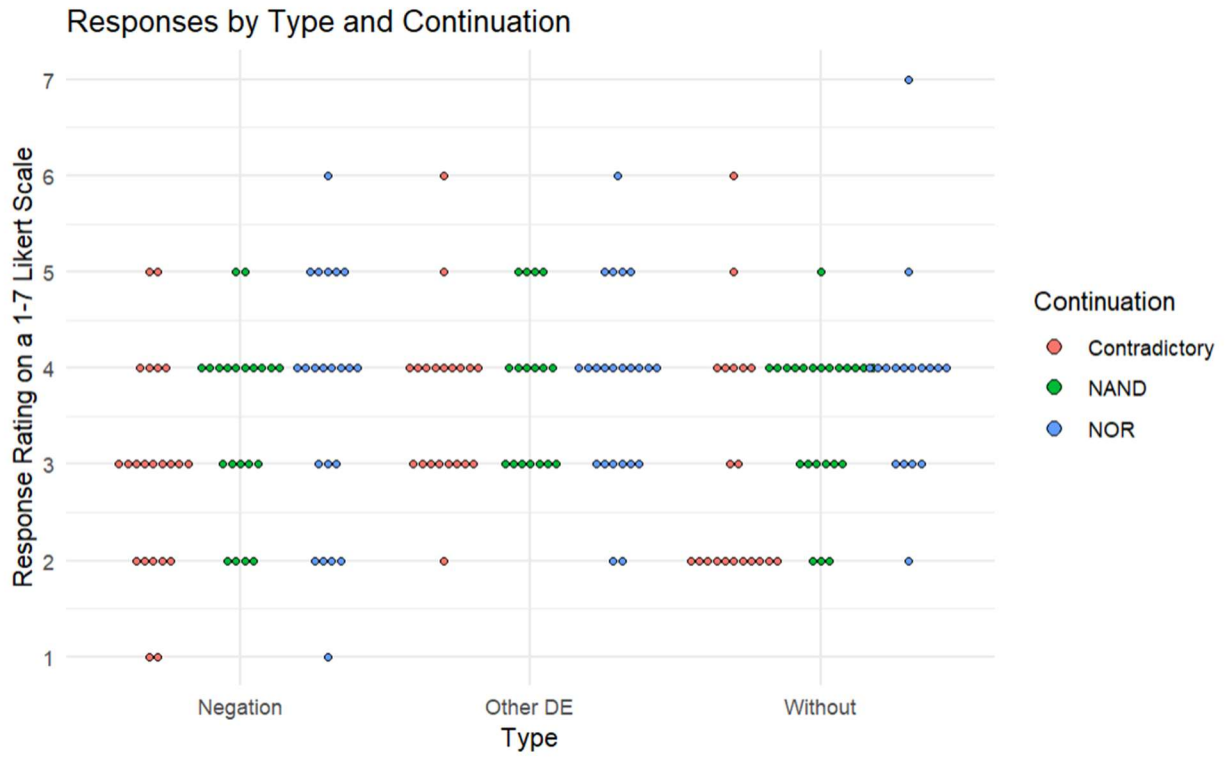


LLM Logic Scale Task Plots





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), se = sd(Response, na.rm =
TRUE)/sqrt(length(Response)))

# 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 - se, ymax = avg_response + se),
position = position_dodge(0.9), width = 0.2) +
  geom_text(aes(label = round(avg_response, 2)), position =
position_dodge(width = 0.9), vjust = 11, size = 2.5) +
  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), se = sd(Response, na.rm =
TRUE)/sqrt(length(Response)))

# 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 - se, ymax = avg_response + se),
na.rm = TRUE, position = position_dodge(0.9), width = 0.2) +
  geom_text(aes(label = round(avg_response, 2)), position =
position_dodge(width = 0.9), vjust = 11, size = 2.5) +
  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))

```