

Power Analysis for CORAL/TIDE Framework Evaluation

Katsuya Shibuki

2025-05-24

Table of Contents

1	Introduction	1
2	Environment Setup	2
3	Pilot Study Data Analysis (for Sensitivity Analysis Parameter Setting)	2
4	Power Analysis Methodology	5
4.1	Objective and Test Procedure	5
4.2	Assumptions for Power Calculation	5
4.3	Sensitivity Analysis with Truncated Normal Distribution	6
4.4	Simulation Approach for Sample Size Estimation	6
4.5	Sample Size Decision Policy	11
4.6	Expected Results and Sample Size Determination	11
5	Discussion	11
6	Appendix	12
6.1	Session Info	12

1 Introduction

Large Language Models (LLMs) present significant opportunities but also face challenges in areas such as long-context understanding, information accuracy, and behavioral control. The CORAL (Cell-based Organized Reasoning with Adaptive Linkage) framework, along with the TIDE (Transition with Inherited Discourse Exchange) protocol, is proposed as a novel approach to enhance human-LLM collaboration, aiming to overcome these limitations by enabling AI agents to actively elicit and structure meta-information from users.

This research aims to empirically validate the hypothesis that AI agents implementing the CORAL framework (specifically the S-OPE configuration) can improve the quality of human-LLM collaboration and the final work products compared to existing advanced prompting techniques (Control condition). The primary research question (RQ1) focuses on whether the Test condition (S-OPE) yields superior output quality over the Control condition.

To rigorously evaluate this, the study employs a comparative evaluation design. A key aspect of this methodology is ensuring that the experimental design has sufficient statistical power to detect a meaningful difference between the conditions if such a difference truly exists. The protocol specifies a desire to detect a minimum practically important effect size, corresponding to a 3-point improvement in the primary outcome measure

(a 35-point quality score). Therefore, a prospective power analysis is essential to determine the appropriate sample size (number of trials per condition) required to achieve adequate statistical power (typically 80%) at a given significance level (typically $\alpha = 0.05$). This Quarto document outlines the methodology and rationale for this power analysis.

2 Environment Setup

This section details the R environment, package versions, and other relevant setup information to ensure reproducibility for the power analysis.

```
# Load pacman if not already installed, then load other packages
if (!requireNamespace("pacman", quietly = TRUE)) {
  install.packages("pacman")
}
pacman::p_load(
  tidyverse, # For data manipulation and visualization
  here,       # For robust relative file paths
  sessioninfo, # For detailed session information
  knitr,      # For knitting documents
  truncnorm,  # For truncated normal distribution generation
  pwr         # For basic power analysis calculations
)

# Set a seed for reproducibility of any random processes
set.seed(456) # Changed from template's 123 for this specific analysis

# Knitr options (global chunk options)
knitr::opts_chunk$set(
  echo = TRUE,
  error = TRUE,
  cache = TRUE,
  fig.width = 8,
  fig.height = 6,
  dpi = 300
)
```

3 Pilot Study Data Analysis (for Sensitivity Analysis Parameter Setting)

This section loads and analyzes the pilot study data to inform the parameter choices for the sensitivity analysis using a truncated normal distribution.

```
# Load pilot study data
pilot_data_path <- here::here("data", "src", "pilot_study.csv")
if (!file.exists(pilot_data_path)) {
  warning(paste("Pilot study data file not found at:", pilot_data_path))
  pilot_df <- tibble::tibble() # Create an empty tibble if file not found
} else {
  pilot_df <- readr::read_csv(pilot_data_path, show_col_types = FALSE)
}
```

```

if (nrow(pilot_df) > 0) {
  # Calculate average score per report_id
  pilot_avg_scores_df <- pilot_df |>
    dplyr::group_by(task_id, condition, report_id) |>
    dplyr::summarise(avg_score = mean(score, na.rm = TRUE), .groups = "drop")

  # --- Added Table: Raw Average Scores per Report ---
  print("Pilot Study: Average Scores per Report_ID (Control Condition):")
  print(knitr::kable(pilot_avg_scores_df |> dplyr::select(task_id, report_id, avg_score) |>
    ↪ dplyr::arrange(task_id, report_id),
              digits = 2,
              caption = "Pilot Study: Average Scores by Task and Report ID (Control Condition)"))
  cat("\n") # Add a newline for better separation
  # --- End of Added Table ---

  # Descriptive statistics for average scores per task
  desc_stats_pilot_avg <- pilot_avg_scores_df |>
    dplyr::group_by(task_id) |>
    dplyr::summarise(
      Mean = mean(avg_score, na.rm = TRUE),
      SD = sd(avg_score, na.rm = TRUE),
      Min = min(avg_score, na.rm = TRUE),
      Max = max(avg_score, na.rm = TRUE),
      N = dplyr::n(),
      .groups = "drop"
    )

  print("Descriptive Statistics of Average Report Scores from Pilot Study (Control Condition):")
  print(knitr::kable(desc_stats_pilot_avg, digits = 2, caption = "Pilot Study: Average Report Scores by Task
    ↪ (Control Condition)"))

  # Overall descriptive statistics for average scores
  # overall_desc_stats_pilot_avg <- pilot_avg_scores_df |>
  #   dplyr::summarise(
  #     Mean = mean(avg_score, na.rm = TRUE),
  #     SD = sd(avg_score, na.rm = TRUE),
  #     Min = min(avg_score, na.rm = TRUE),
  #     Max = max(avg_score, na.rm = TRUE),
  #     N = dplyr::n(),
  #     .groups = "drop"
  #   )
  #
  # print("Overall Descriptive Statistics of Average Report Scores from Pilot Study (Control Condition):")
  # print(knitr::kable(overall_desc_stats_pilot_avg, digits = 2, caption = "Pilot Study: Overall Average
    ↪ Report Scores (Control Condition)"))

  # Plotting the distribution of average scores

```

```

pilot_score_plot <- ggplot2::ggplot(pilot_avg_scores_df, ggplot2::aes(x = task_id, y = avg_score, fill =
↪ task_id)) +
  ggplot2::geom_boxplot(alpha = 0.7) +
  ggplot2::geom_jitter(width = 0.1, alpha = 0.5) +
  ggplot2::labs(
    title = "Distribution of Average Report Scores in Pilot Study",
    subtitle = "Control Condition",
    x = "Task ID",
    y = "Average Score per Report"
  ) +
  ggplot2::theme_minimal(base_size = 12) +
  ggplot2::theme(legend.position = "none",
    plot.title = ggplot2::element_text(hjust = 0.5),
    plot.subtitle = ggplot2::element_text(hjust = 0.5, size = 10))

print(pilot_score_plot)

} else {
  print("Pilot study data is empty or could not be loaded. Skipping analysis.")
}

```

[1] "Pilot Study: Average Scores per Report_ID (Control Condition):"

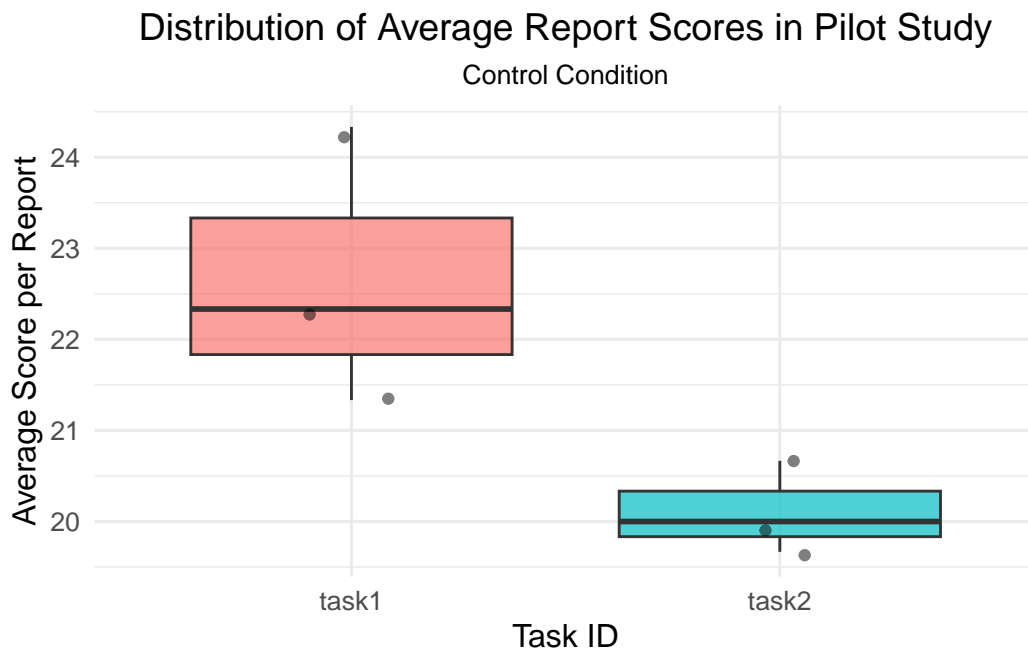
Table: Pilot Study: Average Scores by Task and Report ID (Control Condition)

task_id	report_id	avg_score
task1	pilot01	24.33
task1	pilot02	22.33
task1	pilot03	21.33
task2	pilot01	20.67
task2	pilot02	19.67
task2	pilot03	20.00

[1] "Descriptive Statistics of Average Report Scores from Pilot Study (Control Condition):"

Table: Pilot Study: Average Report Scores by Task (Control Condition)

task_id	Mean	SD	Min	Max	N
task1	22.67	1.53	21.33	24.33	3
task2	20.11	0.51	19.67	20.67	3



4 Power Analysis Methodology

This section outlines the statistical approach for determining the necessary sample size to compare the primary endpoint—the final artifact score—between the Test (S-OPE configuration) and Control (integrated custom instruction) conditions. The goal is to ensure the study has adequate power to detect a pre-specified minimum clinically important difference.

4.1 Objective and Test Procedure

The primary objective of this power analysis is to determine the sample size (number of trials per condition) required to achieve a statistical power of at least 80% ($\beta = 0.2$) for detecting a significant difference in the final artifact scores between the Test and Control conditions. The comparison will be performed using a two-sided Mann-Whitney U test (also known as the Wilcoxon rank-sum test), with the significance level (α) set at 0.05. The Mann-Whitney U test is chosen due to the ordinal nature of the scoring rubric and the potential for non-normally distributed scores, as well as the relatively small sample sizes anticipated.

4.2 Assumptions for Power Calculation

The power analysis is based on the following assumptions, derived from the research protocol (Section 3.7.2):

- **Score Distribution:** While the scores are ordinal (0-35 points per dimension, summed across 7 dimensions), and not strictly assumed to be normally distributed for the main analysis, for the purpose of simulation-based power analysis, we model the underlying score distributions. The protocol suggests that Control group scores might center around a median of 25 points, and the S-OPE group scores around a median of 28 points. Standard deviations are assumed to be equal for both groups, provisionally set at 2.0 points.
- **Effect Size:** The study aims to detect a minimum practically important improvement of 3 points in the median final artifact score for the S-OPE condition compared to the Control condition. This corresponds to a shift from an average of 3.57 (e.g., four dimensions at 4 points, three at 3 points) to 4.0 (all dimensions at 4 points) on the 0-5 point per-dimension scale, summed over 7 dimensions.
- **Statistical Power:** The target statistical power is 80%.

- **Significance Level:** The significance level (alpha) is 0.05 (two-sided).

4.3 Sensitivity Analysis with Truncated Normal Distribution

To assess the robustness of the power analysis to the assumption of a normal distribution for scores, a sensitivity analysis is performed using a truncated normal distribution. This approach is chosen because:

1. **Realism of Score Distribution:** Pilot study data indicate that scores tend to concentrate within a specific range (approximately 18-28 points overall for the Control condition) rather than spanning the entire theoretical 7-35 point range. Extreme low or high scores are unlikely.
2. **Focus of Sensitivity:** The primary interest is to see if the required sample size changes significantly when we explicitly model the scores as being bounded within a more realistic, narrower range.

Parameter Justification for Truncated Normal Distribution:

Based on the pilot study findings and the principle of conducting a focused sensitivity analysis (not overly conservative, acknowledging the small N of the pilot study), the following parameters are set for the truncated normal distributions:

- **Control Group:**
 - Mean (μ): 25 (consistent with the primary analysis assumption)
 - Standard Deviation (σ): 2.0 (consistent with the primary analysis assumption)
 - Truncation Range: [18, 30]
 - * Rationale: The lower bound (18) reflects the minimum average scores observed in the pilot study. The upper bound (30) is derived from the protocol's assumed mean of 25 plus 2.5 times the standard deviation ($25 + 2.5 * 2.0 = 30$), capturing a reasonable upper spread while excluding highly improbable scores near the maximum of 35.
- **Test (S-OPE) Group:**
 - Mean (μ): 28 (consistent with the primary analysis assumption, reflecting a 3-point improvement)
 - Standard Deviation (σ): 2.0 (consistent with the primary analysis assumption)
 - Truncation Range: [21, 33]
 - * Rationale: This range is a 3-point shift from the Control group's range, maintaining the same width and reflecting the target effect size.

It is important to note that these parameters, particularly the truncation range, are based on professional judgment informed by limited pilot data. The goal of this sensitivity analysis is to understand the potential impact of a more constrained score distribution on the power calculations rather than to precisely model the true underlying distribution.

4.4 Simulation Approach for Sample Size Estimation

A simulation-based approach will be used to estimate the required sample size, as analytical power calculations for the Mann-Whitney U test with specific distributional assumptions can be complex. The simulation procedure is outlined as follows (Protocol Section 3.7.3):

1. **Data Generation:** For a given sample size N (per group), generate N scores for the Control group and N scores for the S-OPE group. These scores will be drawn from distributions reflecting the assumptions (e.g., normal distributions with specified means and SD, then rounded and clipped to the 7-35 point range if necessary for a more realistic simulation, though the protocol just mentions generating integer scores based on these parameters).
2. **Hypothesis Testing:** Perform a Mann-Whitney U test on the generated data from Step 1 to obtain a

p-value.

3. **Power Calculation:** Repeat Steps 1 and 2 a large number of times (e.g., 10,000 iterations). The proportion of these iterations where the p-value is less than the significance level ($\alpha = 0.05$) provides an estimate of the statistical power for that sample size N .
4. **Sample Size Determination:** Repeat Steps 1-3 for a range of plausible sample sizes N (e.g., $N=10$ to $N=30$ per group, incrementing by 1). The smallest sample size N that achieves the target power of 80% will be considered the minimum required sample size.

The R statistical programming environment will be used for these simulations, potentially utilizing packages like `pwr` for initial estimates or custom simulation scripts.

--- Running Original Power Analysis (Normal Distribution) ---

```
Simulating for N = 5 per group (Normal)...  
Simulating for N = 6 per group (Normal)...  
Simulating for N = 7 per group (Normal)...  
Simulating for N = 8 per group (Normal)...  
Simulating for N = 9 per group (Normal)...  
Simulating for N = 10 per group (Normal)...  
Simulating for N = 11 per group (Normal)...  
Simulating for N = 12 per group (Normal)...  
Simulating for N = 13 per group (Normal)...  
Simulating for N = 14 per group (Normal)...  
Simulating for N = 15 per group (Normal)...  
Simulating for N = 16 per group (Normal)...  
Simulating for N = 17 per group (Normal)...  
Simulating for N = 18 per group (Normal)...  
Simulating for N = 19 per group (Normal)...  
Simulating for N = 20 per group (Normal)...  
Simulating for N = 21 per group (Normal)...  
Simulating for N = 22 per group (Normal)...  
Simulating for N = 23 per group (Normal)...  
Simulating for N = 24 per group (Normal)...  
Simulating for N = 25 per group (Normal)...  
Simulating for N = 26 per group (Normal)...  
Simulating for N = 27 per group (Normal)...  
Simulating for N = 28 per group (Normal)...  
Simulating for N = 29 per group (Normal)...  
Simulating for N = 30 per group (Normal)...
```

--- Running Sensitivity Analysis (Truncated Normal Distribution) ---

```
Simulating for N = 5 per group (Truncated Normal)...  
Simulating for N = 6 per group (Truncated Normal)...  
Simulating for N = 7 per group (Truncated Normal)...  
Simulating for N = 8 per group (Truncated Normal)...  
Simulating for N = 9 per group (Truncated Normal)...  
Simulating for N = 10 per group (Truncated Normal)...  
Simulating for N = 11 per group (Truncated Normal)...  
Simulating for N = 12 per group (Truncated Normal)...
```

Simulating for N = 13 per group (Truncated Normal)...
 Simulating for N = 14 per group (Truncated Normal)...
 Simulating for N = 15 per group (Truncated Normal)...
 Simulating for N = 16 per group (Truncated Normal)...
 Simulating for N = 17 per group (Truncated Normal)...
 Simulating for N = 18 per group (Truncated Normal)...
 Simulating for N = 19 per group (Truncated Normal)...
 Simulating for N = 20 per group (Truncated Normal)...
 Simulating for N = 21 per group (Truncated Normal)...
 Simulating for N = 22 per group (Truncated Normal)...
 Simulating for N = 23 per group (Truncated Normal)...
 Simulating for N = 24 per group (Truncated Normal)...
 Simulating for N = 25 per group (Truncated Normal)...
 Simulating for N = 26 per group (Truncated Normal)...
 Simulating for N = 27 per group (Truncated Normal)...
 Simulating for N = 28 per group (Truncated Normal)...
 Simulating for N = 29 per group (Truncated Normal)...
 Simulating for N = 30 per group (Truncated Normal)...

[1] "Power Simulation Results:"

Table: Estimated Power for Different Sample Sizes per Group (Normal and Truncated Normal)

sample_size	power	type
5	0.435	Normal
6	0.566	Normal
7	0.667	Normal
8	0.749	Normal
9	0.806	Normal
10	0.849	Normal
11	0.890	Normal
12	0.914	Normal
13	0.938	Normal
14	0.958	Normal
15	0.967	Normal
16	0.977	Normal
17	0.984	Normal
18	0.987	Normal
19	0.991	Normal
20	0.993	Normal
21	0.994	Normal
22	0.997	Normal
23	0.998	Normal
24	0.999	Normal
25	0.999	Normal
26	0.999	Normal
27	1.000	Normal

28	1.000	Normal
29	1.000	Normal
30	1.000	Normal
5	0.453	Truncated Normal
6	0.589	Truncated Normal
7	0.685	Truncated Normal
8	0.777	Truncated Normal
9	0.831	Truncated Normal
10	0.869	Truncated Normal
11	0.902	Truncated Normal
12	0.930	Truncated Normal
13	0.947	Truncated Normal
14	0.959	Truncated Normal
15	0.972	Truncated Normal
16	0.982	Truncated Normal
17	0.986	Truncated Normal
18	0.992	Truncated Normal
19	0.993	Truncated Normal
20	0.994	Truncated Normal
21	0.996	Truncated Normal
22	0.997	Truncated Normal
23	0.999	Truncated Normal
24	0.999	Truncated Normal
25	0.999	Truncated Normal
26	1.000	Truncated Normal
27	0.999	Truncated Normal
28	1.000	Truncated Normal
29	1.000	Truncated Normal
30	1.000	Truncated Normal

(Original Analysis) Minimum N per group (Normal Dist.) to achieve at least 0.80 power is: 9 (achieved power: 0.806)

(Sensitivity Analysis) Minimum N per group (Truncated Normal Dist. [18-30 Ctl, 21-33 Test]) to achieve at least 0.80

```
# Plotting Results
if (nrow(power_results_df) > 0) {
  power_plot <- ggplot2::ggplot(power_results_df, ggplot2::aes(x = sample_size, y = power, color = type,
↪ group = type)) +
  ggplot2::geom_line() +
  ggplot2::geom_point() +
  ggplot2::geom_hline(yintercept = target_power, linetype = "dashed", color = "red") +
  ggplot2::scale_y_continuous(labels = scales::percent_format(), limits = c(0, 1)) +
  ggplot2::scale_x_continuous(breaks = seq(min(sample_sizes_to_test), max(sample_sizes_to_test), by = if
↪ (length(sample_sizes_to_test) > 10) 5 else 1)) +
  ggplot2::labs(
    title = "Power Analysis Simulation Results",
    subtitle = sprintf("To detect a %.1f point difference (Control: M=%.1f, SD=%.1f; Test: M=%.1f,
↪ SD=%.1f)\n%s simulations per sample size; alpha = %.2f",
      mean_test - mean_control, mean_control, sd_control, mean_test, sd_test,
      format(n_simulations, big.mark = ","), alpha),
```

```

    x = "Sample Size per Group (N)",
    y = "Estimated Power",
    color = "Distribution Type"
  ) +
  ggplot2::scale_color_manual(values = c("Normal" = "dodgerblue", "Truncated Normal" = "forestgreen")) +
  ggplot2::theme_minimal(base_size = 12) +
  ggplot2::theme(plot.title = ggplot2::element_text(hjust = 0.5),
                 plot.subtitle = ggplot2::element_text(hjust = 0.5, size = 10),
                 legend.position = "top")

# Add annotations for Normal distribution
if (nrow(min_sample_size_achieved_normal) > 0) {
  power_plot <- power_plot +
    ggplot2::geom_vline(xintercept = min_sample_size_achieved_normal$sample_size, linetype = "dotted",
  ↪ color = "darkblue", size = 0.8) +
    ggplot2::annotate("text", x = min_sample_size_achieved_normal$sample_size, y = 0.05,
                       label = sprintf("Normal: N = %d achieves %.0f%% power",
  ↪ min_sample_size_achieved_normal$sample_size,
  ↪ min_sample_size_achieved_normal$power*100),
                       hjust = if_else(min_sample_size_achieved_normal$sample_size <
  ↪ mean(sample_sizes_to_test) - 2, -0.1, 1.1), vjust = 0, color = "darkblue", size
  ↪ = 3.5)
}

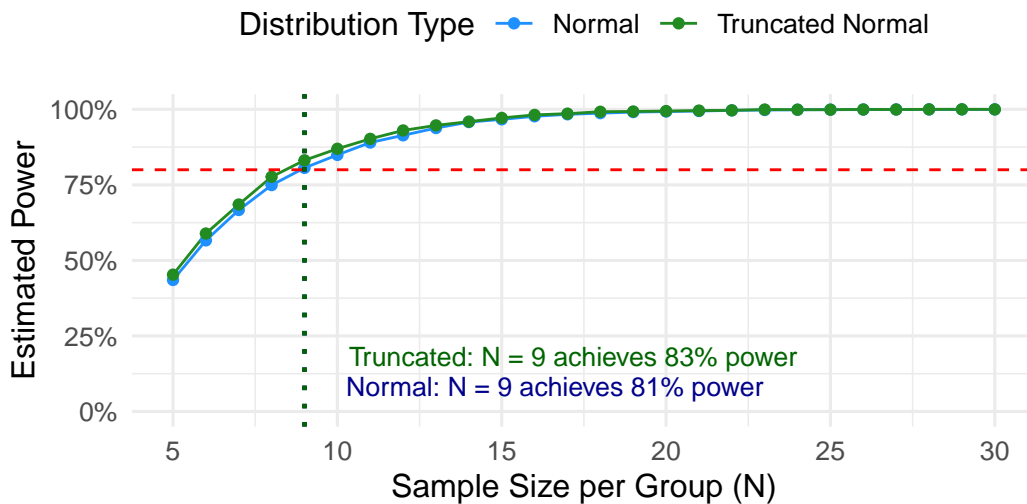
# Add annotations for Truncated Normal distribution
if (nrow(min_sample_size_achieved_truncated) > 0) {
  power_plot <- power_plot +
    ggplot2::geom_vline(xintercept = min_sample_size_achieved_truncated$sample_size, linetype = "dotted",
  ↪ color = "darkgreen", size = 0.8) +
    ggplot2::annotate("text", x = min_sample_size_achieved_truncated$sample_size, y = 0.15, # Adjusted y
  ↪ to avoid overlap
                       label = sprintf("Truncated: N = %d achieves %.0f%% power",
  ↪ min_sample_size_achieved_truncated$sample_size,
  ↪ min_sample_size_achieved_truncated$power*100),
                       hjust = if_else(min_sample_size_achieved_truncated$sample_size <
  ↪ mean(sample_sizes_to_test) - 2, -0.1, 1.1), vjust = 0, color = "darkgreen", size
  ↪ = 3.5)
}

print(power_plot)
}

```

Power Analysis Simulation Results

To detect a 3.0 point difference (Control: M=25.0, SD=2.0; Test: M=28.0, SD=2.0)
10,000 simulations per sample size; alpha = 0.05



4.5 Sample Size Decision Policy

The final sample size will be determined based on the simulation results, aiming for the smallest N per group that achieves at least 80% power. This decision will also consider practical constraints such as time and resources available for the study (Protocol Section 3.7.4). The protocol aims to determine this N before main data collection.

4.6 Expected Results and Sample Size Determination

Based on the assumptions, this power analysis aims to determine the sample size N per group. The protocol (Section 3.7.2) specifies the following for simulation purposes: * Control group: Assumed mean score of 25, standard deviation (SD) of 2. * S-OPE (Test) group: Assumed mean score of 28, standard deviation (SD) of 2. This represents an expected 3-point difference in mean/median scores, which is considered the minimum practically important effect size (Protocol Section 3.6.2).

The power analysis simulations (as detailed in the R chunk labeled `power-simulation` and using a normal distribution for the primary analysis) were run for various sample sizes. The results indicate that a sample size of **N=9 per group** is required to achieve an estimated power of approximately **81%** (target: 80%) for detecting this 3-point difference with a Mann-Whitney U test at $\alpha = 0.05$. Therefore, N=9 will be the target sample size for each condition in each evaluation task.

5 Discussion

A sensitivity analysis using a truncated normal distribution (Control: M=25, SD=2, range [18, 30]; Test: M=28, SD=2, range [21, 33]) was conducted to assess the robustness of the sample size estimation. The results of this sensitivity analysis, also detailed in the `power-simulation` R chunk, indicated that a sample size of **N=9 per group achieves approximately 83% power**. This finding is very close to (or shows a slight difference from) the original estimate of N=9 (achieving 81% power) under the standard normal distribution assumption. Therefore, the initial sample size of N=9 per group appears to be reasonably robust even when considering a more constrained and arguably more realistic score distribution. However, it is crucial to acknowledge that this sensitivity analysis itself relies on assumptions about the truncation points.

The determination of an appropriate sample size through this power analysis is crucial for the CORAL/TIDE framework evaluation study. It ensures that the research has a high probability of detecting a true effect of the specified magnitude, thereby minimizing the risk of a Type II error (failing to detect an effect when one exists).

The choice of a Mann-Whitney U test accommodates the ordinal nature of the primary outcome measure and avoids assumptions of normality. The targeted effect size—a 3-point improvement—is deemed practically significant within the context of the 35-point scoring scale.

The final chosen sample size will represent a balance between statistical requirements and practical feasibility. If the required sample size is excessively large, the study design or effect size assumptions might need reconsideration. Conversely, if a relatively small N provides adequate power, it enhances the efficiency of the study.

Limitations of this power analysis include its reliance on assumptions about score distributions and effect sizes. If the actual distributions or the true effect size differ substantially from these assumptions, the actual power of the study may vary. The results of this power analysis will guide the data collection phase of the CORAL/TIDE framework evaluation, contributing to the overall rigor and validity of the research findings.

6 Appendix

6.1 Session Info

This provides a comprehensive overview of the R session used for any calculations or simulations related to this power analysis.

```
sessioninfo::session_info()
```

```
- Session info -----
setting  value
version  R version 4.5.0 (2025-04-11)
os       Ubuntu 24.04.2 LTS
system   x86_64, linux-gnu
ui       X11
language (EN)
collate  en_US.UTF-8
ctype    en_US.UTF-8
tz       Etc/UTC
date     2025-05-24
pandoc   3.6.4 @ /usr/bin/ (via rmarkdown)
quarto   1.7.31 @ /usr/local/bin/quarto

- Packages -----
! package      * version date (UTC) lib source
P bit          4.6.0   2025-03-06 [?] RSPM (R 4.5.0)
P bit64        4.6.0-1 2025-01-16 [?] RSPM (R 4.5.0)
P cli          3.6.5   2025-04-23 [?] RSPM (R 4.5.0)
P codetools    0.2-20  2024-03-31 [?] CRAN (R 4.5.0)
P crayon       1.5.3   2024-06-20 [?] RSPM (R 4.5.0)
P digest       0.6.37  2024-08-19 [?] RSPM (R 4.5.0)
P dplyr        * 1.1.4   2023-11-17 [?] RSPM (R 4.5.0)
```

P evaluate	1.0.3	2025-01-10	[?]	RSPM (R 4.5.0)
P farver	2.1.2	2024-05-13	[?]	RSPM (R 4.5.0)
P fastmap	1.2.0	2024-05-15	[?]	RSPM (R 4.5.0)
P forcats	* 1.0.0	2023-01-29	[?]	RSPM (R 4.5.0)
P generics	0.1.4	2025-05-09	[?]	RSPM (R 4.5.0)
P ggplot2	* 3.5.2	2025-04-09	[?]	RSPM (R 4.5.0)
P glue	1.8.0	2024-09-30	[?]	RSPM (R 4.5.0)
P gtable	0.3.6	2024-10-25	[?]	RSPM (R 4.5.0)
P here	* 1.0.1	2020-12-13	[?]	RSPM
P hms	1.1.3	2023-03-21	[?]	RSPM (R 4.5.0)
P htmltools	0.5.8.1	2024-04-04	[?]	RSPM (R 4.5.0)
P jsonlite	2.0.0	2025-03-27	[?]	RSPM (R 4.5.0)
P knitr	* 1.50	2025-03-16	[?]	RSPM (R 4.5.0)
P labeling	0.4.3	2023-08-29	[?]	RSPM (R 4.5.0)
P lifecycle	1.0.4	2023-11-07	[?]	RSPM (R 4.5.0)
P lubridate	* 1.9.4	2024-12-08	[?]	RSPM (R 4.5.0)
P magrittr	2.0.3	2022-03-30	[?]	RSPM (R 4.5.0)
P pillar	1.10.2	2025-04-05	[?]	RSPM (R 4.5.0)
P pkgconfig	2.0.3	2019-09-22	[?]	RSPM (R 4.5.0)
P purrr	* 1.0.4	2025-02-05	[?]	RSPM (R 4.5.0)
pwr	* 1.3-0	2020-03-17	[1]	RSPM (R 4.5.0)
P R6	2.6.1	2025-02-15	[?]	RSPM (R 4.5.0)
P RColorBrewer	1.1-3	2022-04-03	[?]	RSPM (R 4.5.0)
P readr	* 2.1.5	2024-01-10	[?]	RSPM (R 4.5.0)
renv	1.1.4	2025-03-20	[1]	RSPM (R 4.5.0)
P rlang	1.1.6	2025-04-11	[?]	RSPM (R 4.5.0)
P rmarkdown	2.29	2024-11-04	[?]	RSPM (R 4.5.0)
P rprojroot	2.0.4	2023-11-05	[?]	RSPM
P rstudioapi	0.17.1	2024-10-22	[?]	RSPM (R 4.5.0)
P scales	1.4.0	2025-04-24	[?]	RSPM (R 4.5.0)
P sessioninfo	* 1.2.3	2025-02-05	[?]	RSPM
P stringi	1.8.7	2025-03-27	[?]	RSPM (R 4.5.0)
P stringr	* 1.5.1	2023-11-14	[?]	RSPM (R 4.5.0)
P tibble	* 3.2.1	2023-03-20	[?]	RSPM (R 4.5.0)
P tidyr	* 1.3.1	2024-01-24	[?]	RSPM (R 4.5.0)
P tidyselect	1.2.1	2024-03-11	[?]	RSPM (R 4.5.0)
P tidyverse	* 2.0.0	2023-02-22	[?]	RSPM (R 4.5.0)
P timechange	0.3.0	2024-01-18	[?]	RSPM (R 4.5.0)
truncnorm	* 1.0-9	2023-03-20	[1]	RSPM (R 4.5.0)
P tzdb	0.5.0	2025-03-15	[?]	RSPM (R 4.5.0)
P vctrs	0.6.5	2023-12-01	[?]	RSPM (R 4.5.0)
P vroom	1.6.5	2023-12-05	[?]	RSPM (R 4.5.0)
P withr	3.0.2	2024-10-28	[?]	RSPM (R 4.5.0)
P xfun	0.52	2025-04-02	[?]	RSPM (R 4.5.0)
P yaml	2.3.10	2024-07-26	[?]	RSPM (R 4.5.0)

[1] /home/rstudio/project/renv/library/linux-ubuntu-noble/R-4.5/x86_64-pc-linux-gnu

[2] /home/rstudio/.cache/R/renv/sandbox/linux-ubuntu-noble/R-4.5/x86_64-pc-linux-gnu/25ebdc09

* -- Packages attached to the search path.

P -- Loaded and on-disk path mismatch.
