Research Question 1 Analyses for SANE Training Program Data

Steven J. Pierce[1]

[1]Michigan State University, Center for Statistical Training and Consulting, https://cstat.msu.edu

Table of contents

1 Purpose

This file is part of a research compendium (Pierce, 2026) associated with a study about a sexual assault nurse examiner training program (Dontje & Campbell, 07/01/2021–06/30/2025). It contains results from analyses that address RQ1.

2 Research Question

- **RQ1.** What are the eligibility and enrollment rates for this program, and do eligible vs. enrolled participants differ with respect to their background, motivations, potential barriers, and emotional readiness for this kind of work?

3 Setup

This section documents some setup tasks that are useful to the statistician on the team. Most readers of this document will probably want skip directly to Section 3.4.

3.1 Define Global Options

Global R chunk options are defined in the YAML header but local chunk options will over-ride global options. We can temporarily disable an individual chunk by inserting `#| eval: false` on a line at the top of the chunk. The method for creating a `cfsize` option that controls font size in code chunks and their text output is based on an answer to a question posted on stackoverflow.com.

````
```{r}
#| label: global-options

# Create a custom chunk hook/option for controlling font size in chunk & output.
def.chunk.hook  <- knitr::knit_hooks$get("chunk")
knitr::knit_hooks$set(chunk = function(x, options) {
  x <- def.chunk.hook(x, options)
  ifelse(options$cfsize != "normalsize",
         paste0("\n \\", options$cfsize,"\n\n", x, "\n\n \\normalsize"),
         x)
  })
```
````

3.2 Load Packages

R packages usually add new functions to the base R software, allowing you to do more things. Here, we load the specific R packages required for this script to work.

````
```{r}
#| label: load-packages
library(devtools)       # for session_info()
```
````

```
Loading required package: usethis
```

```{r}
#| label: load-packages
library(here)           # for here(), i_am(), makes code more portable.
```

```
here() starts at P:/Consulting/Cases_1600-1799/C1788/SANETPA/scripts
```

```{r}
#| label: load-packages
library(rmarkdown)      # for pandoc_version()
library(knitr)          # for kable()
library(dplyr)          # for %>%, filter(), group_by(), mutate(), rename(), etc.
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```{r}
#| label: load-packages
library(broom)          # for broom()
library(effectsize)     # for hedges_g()
library(tidyverse)      # for map(), map_dfr(), map_chr(), rowid_to_column(),
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v forcats   1.0.1     v readr     2.1.6
v ggplot2   4.0.2     v stringr   1.6.0
v lubridate 1.9.5     v tibble    3.3.1
v purrr     1.2.1     v tidyr     1.3.2
```

```
-- 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
```

```{r}
#| label: load-packages
                        # unnest(), etc.
library(haven)          # for as_factor()
library(janitor)        # for adorn_totals()
```

```
Attaching package: 'janitor'

The following objects are masked from 'package:stats':

    chisq.test, fisher.test
```

```{r}
#| label: load-packages
options(kableExtra.latex.load_packages = FALSE)
library(kableExtra)     # for kable_styling(), add_header_above(),
```

```
Attaching package: 'kableExtra'

The following object is masked from 'package:dplyr':

    group_rows
```

```{r}
#| label: load-packages
                          # column_spec(), row_spec() etc.
library(piercer)          # for dc_summary(), file_details(), git_report(),
                          # which_latex()
library(psych)            # for alpha()
```

```
Attaching package: 'psych'

The following objects are masked from 'package:ggplot2':

    %+%, alpha

The following object is masked from 'package:effectsize':

    phi
```

```{r}
#| label: load-packages
library(quarto)           # for quarto_version()
library(SANETPA)          # for version info
```

3.3 Declare Path

This next chunk declares the path to this script relative to the project-level root directory. If the file is not in the right location under the project root you'll get a warning message. This helps ensure relative paths are all working as expected. The chunk below uses the `SourceDir` and `SourceFile` parameters set in the YAML header.

```{r}
#| label: declare-path

# Declare path to this script relative to the project root directory.
here::i_am(path = paste0(params$SourceDir, params$SourceFile))
```

```
here() starts at P:/Consulting/Cases_1600-1799/C1788/SANETPA
```

3.4 Read Data

We start by reading in the datafile, which contains the following datasets that we need for RQ1 analyses:

- `Applicants`. This is a person-level data file containing one row for every applicant, regardless of eligibility status.,
- `Eligible_Applicants`. This is a person-level file containing one row for every person who both applied and was eligible for the program. It omits those who were ineligible.

```{r}
#| label: load-data
#| eval: true

# Store path to data file.
DataFile <- here("data/Imported_SANETP_Data.RData")

load(file = DataFile)
```

Table 1 shows meta-data about the data file we just loaded and Table 2 shows the sizes of the datasets it contains.

```{r}
#| label: tbl-imported-data-file
#| tbl-cap: "Meta-Data About the Data File Loaded"

file_details(DataFile) %>%
  kable(, format = "latex", booktabs = TRUE,
        col.names = c("File Name", "Size", "Last Modified")) %>%
  kable_styling()
```

Table 1: Meta-Data About the Data File Loaded

File Name	Size	Last Modified
Imported_SANETP_Data.RData	181K	2026-02-14 12:56:34

```{r}
#| label: tbl-datasets
#| tbl-cap: "Sizes of the Datasets Used for RQ1"

data.frame(Dataset = c("Applicants", "Eligible_Applicants"),
           N_Rows = c(nrow(Applicants), nrow(Eligible_Applicants)),
           N_Cols = c(ncol(Applicants), ncol(Eligible_Applicants))) %>%
  kable(, format = "latex", booktabs = TRUE,
        col.names = c("Dataset", "N Rows", "N Columns")) %>%
  kable_styling()
```

Table 2: Sizes of the Datasets Used for RQ1

Dataset	N Rows	N Columns
Applicants	497	154
Eligible_Applicants	327	154

3.5 Update Data

Below, we just convert some variables to factors to improve output formatting later.

```{r}
#| label: update-Eligible-Applicants

Eligible_Applicants <- Eligible_Applicants %>%
  mutate(Education = as_factor(Education),
         Education = fct_relevel(Education, c("Diploma", "Associates degree")),
         License = as_factor(License),
         Setting = as_factor(Setting),
         Employ_Status = as_factor(Employ_Status),
         Employ_Nurse = as_factor(Employ_Nurse),
         Prior_SANE = as_factor(Prior_SANE))
```

We will need to create tables that have both overall summaries for all eligible applicants regardless of `Enrolled` value, and breakdowns for `Enrolled = No` and `Enrolled = Yes`. Creating a copy of `Eligible_Applicants` called `Overall` where I've recoded to set `Enrolled = Both` for all cases and made that a factor with 3 levels (`Both`, `No`, and `Yes`) will allow me to bind `Overall` and `Eligible_Applicants` together then use the same code that already worked to get the breakdowns for `Enrolled = No` and `Enrolled = Yes` to also get the overall summary.

```{r}
#| label: create-Overall

Overall <- Eligible_Applicants %>%
  mutate(Enrolled = "Overall",
         Enrolled = factor(Enrolled, levels = c("Overall", "No", "Yes")))
```

4 Methods

This section describes the data sources, measures, and statistical methods used to answer RQ1.

4.1 Data Sources

We rely on two different datasets for answering RQ1. The `Applicants` dataset has one row for every unique applicant regardless of eligibility or enrollment status ($N = 497$). Meanwhile, the `Eligible_Applicants` dataset contains the subset of applicants who met all eligibility criteria for participating in the program ($N = 327$). Some, but not all, eligible applicants enrolled in the training program.

4.2 Measures

The sections below describe the measures used to answer RQ1.

4.2.1 Eligibility Status

We measured applicants' eligibility status with a binary variable (`Eligible`, coded 0 = *No*, 1 = *Yes*). This was the case selection variable for creating the `Eligible_Applicants` dataset from the `Applicants` dataset and the outcome variable for assessing eligibility rate.

4.2.2 Enrollment Status

Both the eligible and enrolled applicants met the program eligibility criteria. The two groups are actually defined by the applicant's enrollment status, which is a binary variable (`Enrolled`, coded 0 = *No*, 1 = *Yes*). We can therefore also refer to these groups as unenrolled versus enrolled applicants for a bit more clarity. Enrollment status is the outcome variable for estimating the enrollment rate, but becomes the independent variable when we compare unenrolled and enrolled groups of applicants.

4.2.3 Background Characteristics

We measured applicants' education via a categorical variable (`Education`) with four levels: *Diploma*, *Associates degree*, *Bachelors degree*, and *Graduate degree*. Their licensure or certification for practice was recorded in a categorical variable (`License`) with three levels: *Certified Nurse Midwife (CNM)*, *Nurse Practitioner (NP)*, and *Registered Nurse (RN)*.

The primary setting where applicants practiced nursing (`Setting`) was measured by a categorical variable with three levels: *Urban*, *Rural/Tribal*, and *Suburban*. As noted in the data import script output, *Tribal* was initially a separate category but later combined with *Rural* because it was too small to be analyzed on its own and those categories represent settings most similar in size. This glosses over important cultural differences between rural and tribal settings, but it is better than omitting the applicants from tribal settings.

Applicants' current employment status was measured by a categorical variable (`Employ_Status`) with three levels: *Full time*, *Part time*, and *Unemployed*. They were also asked whether they were currently employed as a nurse: `Employ_Nurse` was coded *No*, *Yes*, or *NA* (for missing data that were not available). We also treated that as a categorical variable with three levels. Applicants reported how many years they had been been practicing as a nurse since completing their initial nursing program, with `Nurse_Years` recorded as a continuous variable.

Finally, applicants were asked whether they had attended any previous SANE training, with `Prior_SANE` recorded as a three level categorical variable coded *No training*, *Completed didactic training*, and *Completed didactic and clinical skills training*.

4.2.4 Motivations

The two motivation measures are binary (coded $0 =$ *No*, $1 =$ *Yes*) and include whether the applicant was motivated to seek the training by (a) a need for SANE services in their community or organization (`Motivation_NeedSANE`), and (b) a personal connection to sexual assault (e.g., someone they know is a survivor) (`Motivation_PersonalConn`). We refer to the proportions for these outcomes as endorsement rates because they reflect how often applicants endorsed having each of these motivations.

4.2.5 Potential Barriers

The potential barriers outcome variables are single-item measures of barriers to participation due to family obligations (`Barrier_FO`) and work responsibilities (`Barrier_WR`) as competing demands on applicants' time. While these are technically ordinal items following a 5-point Likert-response format, that should be enough categories to treat them as continuous variables for these analyses.

4.2.6 Emotional Readiness

The emotional readiness measures are all continuous scores measured by ProQOL subscales for burnout (`ProQOL_BO`), compassion satisfaction (`ProQOL_CS`), and secondary traumatic stress (`ProQOL_STS`) (Stamm, 2010).

4.3 Planned Statistical Analyses

This section describes the statistical analysis methods used to answer RQ1.

4.3.1 Eligibility and Enrollment Rates

Estimating the eligibility and enrollment rates was simple. We estimated the means of relevant binary variables (`Eligible` and `Enrolled`). Starting from a dataset with the correct set of applicants ensures that we use the correct denominator for each rate (the mean of a binary variable is the proportion of observations that have a value of 1). The eligibility rate used data about all applicants, whereas the enrollment rate used only data from eligible applicants. We used Wilson score confidence intervals to quantify the uncertainty in both rates (Newcombe, 2012; Wilson, 1927).

4.3.2 Comparisons of Eligible vs Enrolled Applicants

The method used to compare groups of applicants depended on the type of outcome measure involved. We used an extension of Fisher's exact test to examine whether the distributions of nominal outcomes with 3 or more levels differed between unenrolled versus enrolled applicants. The null hypothesis is that that 2 categorical variables representing rows and columns of a 2-way contingency table are independent. This test does not produce a test statistic other than a p-value. In the contingency tables, rows represent values of the outcome, columns represent the unenrolled versus enrolled participants. We report column percentages to emphasize the conditional distribution of the outcome within each enrollment status group.

We estimated the the proportions of eligible applicants in each group who have a 1 (*Yes*) on each binary outcome, along with corresponding Wilson score confidence intervals (Newcombe, 2012; Wilson, 1927). Then, we estimated the difference in proportions ($p_1 - p_2$) to obtain an unstandardized effect size measure, conducted a score test with continuity correction to test for a difference between independent proportions (i.e., $H_0 : p_1 - p_2 = 0$ or $H_0 : p_1 = p_2$) and obtained a corresponding score confidence interval for the difference in proportions (Newcombe, 1998, 2012).

For continuous outcomes, we estimated the mean and corresponding confidence interval for each group, then conducted Welch's t-tests to compare the means using a standard null hypothesis ($H_0 : \mu_1 - \mu_2 = 0$ or $H_0 : \mu_1 = \mu_2$). Welch's t-test does not assume equal variances between the groups and its degrees of freedom are a function of the group sizes and standard deviations (Delacre et al., 2017). The t-tests are supported by point estimates and confidence intervals for two effect size measures: the raw mean difference ($\bar{x_1} - \bar{x_2}$, an unstandardized measure in the units of the original variables) and Hedges' g_s^*. The latter measure is a bias-corrected, standardized mean difference in units of standard deviations that is better than Cohen's d when variances are unequal, so it pairs well with Welch's t-test (Delacre et al., 2021).

5 Results

5.1 Eligibility and Enrollment Rates

The eligibility and enrollment rates are summarized in Table 3.

````
```{r}
#| label: tbl-Rates
#| tbl-cap: Eligibility and Enrollment Rates

FN <- paste("CI, Wilson score confidence interval for a proportion.")

bind_rows(dc_summary(data = Applicants, vars = "Eligible"),
          dc_summary(data = Eligible_Applicants, vars = "Enrolled")) %>%
  bind_cols(Among = c("All applicants", "Eligible applicants")) %>%
  mutate(No = as.character(No),
         Yes = as.character(Yes),
         Total = as.character(Total)) %>%
  relocate(Variable, Among) %>%
  kable(format = "latex", booktabs = TRUE, digits = 2, align = "llrrrrrr",
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2,"Binary Variable (N)" = 3, " ",
                              "95% CI" = 2)) %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```
````

Table 3: Eligibility and Enrollment Rates

		Binary Variable (N)				95% CI	
Variable	Among	No	Yes	Total	Rate	LL	UL
Eligible	All applicants	170	327	497	0.66	0.62	0.70
Enrolled	Eligible applicants	73	254	327	0.78	0.73	0.82

Note: CI, Wilson score confidence interval for a proportion.

5.2 Comparisons of Eligible vs Enrolled Applicants

This section is divided into subsections by which group of variables (background, motivations, potential barriers, and emotional readiness) is being summarized. All of these comparisons used the `Eligible_Applicants` dataset.

> Tip
>
> When inserting these results into the manuscript, it will be more space-efficient to combine tables for outcomes that have the same level of measurement (nominal variables with 3 or more levels, binary, or continuous). The similar table structure across such outcomes should make that easy.

5.2.1 Background Characteristics

Table 4 shows the contingency table comparing unenrolled and enrolled applicants with respect to education, while Table 5 compares their licensing, Table 6 compares the settings where they practice, Table 7 compares their employment status, Table 8 compares whether they were employed as nurses. Table 9 shows the mean scores for years of employment as nurses separately for unenrolled and enrolled groups of eligible applicants, while Table 10 shows the t-tests for mean differences between the two groups, along with estimates and confidence intervals for the effect size. Finally, Table 11 shows the contingency table comparing unenrolled and enrolled applicants with respect to prior SANE training.

````
```{r}
#| label: tbl-Education
#| tbl-cap: Contingency Table for Education by Enrollment Status Among All
#|          Eligible Applicants
#| message: false

CT <- Eligible_Applicants %>%
  xtabs(~ Education + Enrolled, addNA = TRUE, data = .)

FN <- paste0("Two-sided Fisher's exact test for independence, p = ",
             display_num(fisher.test(x = CT)$p.value), ".")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled)) %>%
  group_by(Enrolled, Education) %>%
  summarize(N = n()) %>%
  pivot_wider(id_cols = Education, names_from = Enrolled, values_from = N) %>%
  mutate(Overall = No + Yes,
         Overall.p = 100*Overall/sum(Overall),
         No.p = 100*No/sum(No),
         Yes.p = 100*Yes/sum(Yes)) %>%
  select(Education, Overall, Overall.p, No, No.p, Yes, Yes.p) %>%
  adorn_totals(where = "row") %>%
  kable(format = "latex", booktabs = TRUE, digits = 1,
        col.names = c("Education", "N", "%", "N", "%", "N", "%")) %>%
  kable_styling() %>%
````
````

```
  add_header_above(header = c(" ", "Overall" = 2, "No" = 2, "Yes" = 2)) %>%
  add_header_above(header = c(" " = 3, "Enrolled" = 4)) %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```

Table 4: Contingency Table for Education by Enrollment Status Among All Eligible Applicants

			Enrolled			
	Overall		No		Yes	
Education	N	%	N	%	N	%
Diploma	2	0.6	1	1.4	1	0.4
Associates degree	79	24.2	16	21.9	63	24.8
Bachelors degree	175	53.5	39	53.4	136	53.5
Graduate degree	71	21.7	17	23.3	54	21.3
Total	327	100.0	73	100.0	254	100.0

Note: Two-sided Fisher's exact test for independence, p = 0.660.

```
```{r}
#| label: tbl-License
#| tbl-cap: Contingency Table for License by Enrollment Status Among All
#|          Eligible Applicants
#| message: false

CT <- Eligible_Applicants %>%
  xtabs(~ License + Enrolled, addNA = FALSE, data = .)

FN <- paste0("Two-sided Fisher's exact test for independence, p = ",
             display_num(fisher.test(x = CT)$p.value), ".")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled)) %>%
  group_by(Enrolled, License) %>%
  summarize(N = n()) %>%
  pivot_wider(id_cols = License, names_from = Enrolled, values_from = N) %>%
  mutate(Overall = No + Yes,
         Overall.p = 100*Overall/sum(Overall),
         No.p = 100*No/sum(No),
         Yes.p = 100*Yes/sum(Yes)) %>%
  select(License, Overall, Overall.p, No, No.p, Yes, Yes.p) %>%
  adorn_totals(where = "row") %>%
  kable(format = "latex", booktabs = TRUE, digits = 1,
        col.names = c("License", "N", "%", "N", "%", "N", "%")) %>%
  kable_styling() %>%
  add_header_above(header = c(" ", "Overall" = 2, "No" = 2, "Yes" = 2)) %>%
  add_header_above(header = c(" " = 3, "Enrolled" = 4)) %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```
```

```
```{r}
#| label: tbl-Setting
#| tbl-cap: Contingency Table for Practice Setting by Enrollment Status Among
#|          All Eligible Applicants
#| message: false

CT <- Eligible_Applicants %>%
  xtabs(~ Setting + Enrolled, addNA = FALSE, data = .)

FN <- paste0("Two-sided Fisher's exact test for independence, p = ",
             display_num(fisher.test(x = CT)$p.value), ".")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled)) %>%
  group_by(Enrolled, Setting) %>%
  summarize(N = n()) %>%
  pivot_wider(id_cols = Setting, names_from = Enrolled, values_from = N) %>%
  mutate(Overall = No + Yes,
         Overall.p = 100*Overall/sum(Overall),
```
```

Table 5: Contingency Table for License by Enrollment Status Among All Eligible Applicants

			Enrolled			
	Overall		No		Yes	
License	N	%	N	%	N	%
Certified Nurse Midwife (CNM)	6	1.8	1	1.4	5	2.0
Nurse Practitioner (NP)	33	10.1	11	15.1	22	8.7
Registered Nurse (RN)	288	88.1	61	83.6	227	89.4
Total	327	100.0	73	100.0	254	100.0

Note: Two-sided Fisher's exact test for independence, p = 0.266.

```
        No.p = 100*No/sum(No),
        Yes.p = 100*Yes/sum(Yes)) %>%
  select(Setting, Overall, Overall.p, No, No.p, Yes, Yes.p) %>%
  adorn_totals(where = "row") %>%
  kable(format = "latex", booktabs = TRUE, digits = 1,
        col.names = c("Setting", "N", "%", "N", "%", "N", "%")) %>%
  kable_styling() %>%
  add_header_above(header = c(" ", "Overall" = 2, "No" = 2, "Yes" = 2)) %>%
  add_header_above(header = c(" " = 3, "Enrolled" = 4)) %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```

Table 6: Contingency Table for Practice Setting by Enrollment Status Among All Eligible Applicants

			Enrolled			
	Overall		No		Yes	
Setting	N	%	N	%	N	%
Urban	116	35.5	27	37	89	35.0
Rural/Tribal	107	32.7	27	37	80	31.5
Suburban	104	31.8	19	26	85	33.5
Total	327	100.0	73	100	254	100.0

Note:
Two-sided Fisher's exact test for independence, p = 0.450.

```{r}
#| label: tbl-Employ-Status
#| tbl-cap: Contingency Table for Employment Status by Enrollment Status Among
#|          All Eligible Applicants
#| message: false

CT <- Eligible_Applicants %>%
  xtabs(~ Employ_Status + Enrolled, addNA = FALSE, data = .)

FN <- paste0("Two-sided Fisher's exact test for independence, p = ",
             display_num(fisher.test(x = CT)$p.value), ".")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled)) %>%
  group_by(Enrolled, Employ_Status) %>%
  summarize(N = n()) %>%
  pivot_wider(id_cols = Employ_Status, names_from = Enrolled, values_from = N) %>%
  mutate(Overall = No + Yes,
         Overall.p = 100*Overall/sum(Overall),
         No.p = 100*No/sum(No),
```

```
         Yes.p = 100*Yes/sum(Yes)) %>%
  select(Employ_Status, Overall, Overall.p, No, No.p, Yes, Yes.p) %>%
  adorn_totals(where = "row") %>%
  kable(format = "latex", booktabs = TRUE, digits = 1,
        col.names = c("Employment Status", "N", "%", "N", "%", "N", "%")) %>%
  kable_styling() %>%
  add_header_above(header = c(" ", "Overall" = 2, "No" = 2, "Yes" = 2)) %>%
  add_header_above(header = c(" " = 3, "Enrolled" = 4)) %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```

Table 7: Contingency Table for Employment Status by Enrollment Status Among All Eligible Applicants

			Enrolled			
	Overall		No		Yes	
Employment Status	N	%	N	%	N	%
Full time	255	78.0	57	78.1	198	78.0
Part time	56	17.1	12	16.4	44	17.3
Unemployed	16	4.9	4	5.5	12	4.7
Total	327	100.0	73	100.0	254	100.0

Note: Two-sided Fisher's exact test for independence, p = 0.934.

```{r}
#| label: tbl-Employ-Nurse
#| tbl-cap: Contingency Table for Employed as Nurse by Enrollment Status Among
#|          All Eligible Applicants
#| message: false

CT <- Eligible_Applicants %>%
  xtabs(~ Employ_Nurse + Enrolled, addNA = TRUE, data = .)

FN <- paste0("Two-sided Fisher's exact test for independence, p = ",
             display_num(fisher.test(x = CT)$p.value), ". ",
             "NA, not available (missing data).")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled)) %>%
  group_by(Enrolled, Employ_Nurse) %>%
  summarize(N = n()) %>%
  pivot_wider(id_cols = Employ_Nurse, names_from = Enrolled, values_from = N) %>%
  mutate(Overall = No + Yes,
         Overall.p = 100*Overall/sum(Overall),
         No.p = 100*No/sum(No),
         Yes.p = 100*Yes/sum(Yes)) %>%
  select(Employ_Nurse, Overall, Overall.p, No, No.p, Yes, Yes.p) %>%
  adorn_totals(where = "row") %>%
  kable(format = "latex", booktabs = TRUE, digits = 1,
        col.names = c("Employed as Nurse", "N", "%", "N", "%", "N", "%")) %>%
  kable_styling() %>%
  add_header_above(header = c(" ", "Overall" = 2, "No" = 2, "Yes" = 2)) %>%
  add_header_above(header = c(" " = 3, "Enrolled" = 4)) %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```

```{r}
#| label: tbl-Nurse-Years-means
#| tbl-cap: Nurse Years Means by Enrollment Status Among All Eligible Applicants
#| message: false

FN <- paste("N, number of eligible applicants;",
            "N_m, number of missing values;",
            "N_o, number of observed values.")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled),
         Enrolled = fct_expand(Enrolled, "Overall", after = 0)) %>%
  bind_rows(., Overall) %>%
  select(ID, Enrolled, Nurse_Years) %>%
  pivot_longer(cols = c(Nurse_Years), names_to = "Variable",
```

Table 8: Contingency Table for Employed as Nurse by Enrollment Status Among All Eligible Applicants

			Enrolled			
	Overall		No		Yes	
Employed as Nurse	N	%	N	%	N	%
No	10	3.1	2	2.7	8	3.1
Yes	306	93.6	68	93.2	238	93.7
NA	11	3.4	3	4.1	8	3.1
Total	327	100.0	73	100.0	254	100.0

Note: Two-sided Fisher's exact test for independence, p = 0.914. NA, not available (missing data).

```
               values_to = "Value") %>%
  group_by(Variable, Enrolled) %>%
  summarize(N = n(),
            N_m = sum(is.na(Value)),
            N_o = sum(!is.na(Value)),
            Mean = mean(Value, na.rm = TRUE),
            SD = sd(Value, na.rm = TRUE),
            SE = SD/sqrt(N_o),
            LL = Mean - qt(0.975, df = N_o - 1)*SE,
            UL = Mean + qt(0.975, df = N_o - 1)*SE) %>%
  mutate(N = as.character(N),
         N_m = as.character(N_m),
         N_o = as.character(N_o)) %>%
  kable(format = "latex", booktabs = TRUE, digit = 2, align = "llrrrrrrrr",
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2, "Sample Size" = 3, " " = 3,
                              "95% CI" = 2)) %>%
  collapse_rows(columns = 1:2, valign = "top", latex_hline = "major",
                row_group_label_position = "first") %>%
  footnote(general = FN, footnote_as_chunk = TRUE, threeparttable = TRUE)
```

Table 9: Nurse Years Means by Enrollment Status Among All Eligible Applicants

		Sample Size						95% CI	
Variable	Enrolled	N	N_m	N_o	Mean	SD	SE	LL	UL
Nurse_Years	Overall	327	0	327	1.39	2.36	0.13	1.13	1.65
	No	73	0	73	1.39	1.22	0.14	1.11	1.68
	Yes	254	0	254	1.39	2.60	0.16	1.07	1.71

Note: N, number of eligible applicants; N_m, number of missing values; N_o, number of observed values.

```
```{r}
#| label: tbl-Nurse-Years-ttests
#| tbl-cap: T-Tests Comparing Nurse Years Measures Between Unenrolled
#|          and Enrolled Groups of Eligible Applicants
#| message: false

FN <- paste("Difference, difference in means (Unenrolled - Enrolled);",
            "LL, lower limit; UL, upper limit.")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled)) %>%
  select(Enrolled, Nurse_Years) %>%
  pivot_longer(cols = c(Nurse_Years), names_to = "Variable",
               values_to = "Value") %>%
  nest(.by = Variable) %>%
  mutate(t.test = map(data, ~tidy(t.test(Value ~ Enrolled, data = .x))),
```
```

```
       Hedges_G = map(data, ~tibble(hedges_g(Value ~ Enrolled, data = .x,
                                            pooled_sd = FALSE)))) %>%
  unnest(t.test) %>%
  rename(Difference = estimate, t = statistic, p = p.value, df = parameter,
         LL = conf.low, UL = conf.high) %>%
  unnest(Hedges_G) %>%
  rename(g_LL = CI_low, g_UL = CI_high) %>%
  select(Variable, Difference, LL, UL, Hedges_g, g_LL, g_UL, t, df, p) %>%
  mutate(p = display_num(p)) %>%
  kable(format = "latex", booktabs = TRUE, digit = 2,
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2, "Diff. 95% CI" = 2,
                              " ", "g 95% CI" = 2,
                              "Welch 2-sample t-test" = 3)) %>%
  footnote(general = FN, footnote_as_chunk = TRUE, threeparttable = TRUE)
```

Table 10: T-Tests Comparing Nurse Years Measures Between Unenrolled and Enrolled Groups of Eligible Applicants

		Diff. 95% CI			g 95% CI		Welch 2-sample t-test		
Variable	Difference	LL	UL	Hedges_g	g_LL	g_UL	t	df	p
Nurse_Years	0.00	-0.43	0.43	0.00	-0.21	0.21	0.01	257.53	0.995

Note: Difference, difference in means (Unenrolled - Enrolled); LL, lower limit; UL, upper limit.

```{r}
#| label: tbl-Prior-SANE
#| tbl-cap: Contingency Table for Prior SANE Training by Enrollment Status Among
#|          All Eligible Applicants
#| message: false

CT <- Eligible_Applicants %>%
  xtabs(~ Prior_SANE + Enrolled, addNA = FALSE, data = .)

FN <- paste0("Two-sided Fisher's exact test for independence, p = ",
             display_num(fisher.test(x = CT)$p.value), ".")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled)) %>%
  group_by(Enrolled, Prior_SANE) %>%
  summarize(N = n()) %>%
  pivot_wider(id_cols = Prior_SANE, names_from = Enrolled, values_from = N) %>%
  mutate(Overall = No + Yes,
         Overall.p = 100*Overall/sum(Overall),
         No.p = 100*No/sum(No),
         Yes.p = 100*Yes/sum(Yes)) %>%
  select(Prior_SANE, Overall, Overall.p, No, No.p, Yes, Yes.p) %>%
  adorn_totals(where = "row") %>%
  kable(format = "latex", booktabs = TRUE, digits = 1,
        col.names = c("Prior SANE Training", "N", "%", "N", "%", "N", "%")) %>%
  kable_styling() %>%
  add_header_above(header = c(" ", "Overall" = 2, "No" = 2, "Yes" = 2)) %>%
  add_header_above(header = c(" " = 3, "Enrolled" = 4)) %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```

Table 11: Contingency Table for Prior SANE Training by Enrollment Status Among All Eligible Applicants

			Enrolled			
	Overall		No		Yes	
Prior SANE Training	N	%	N	%	N	%
No training	323	98.8	71	97.3	252	99.2
Completed didactic training	2	0.6	1	1.4	1	0.4
Completed didactic and clinical skills training	2	0.6	1	1.4	1	0.4
Total	327	100.0	73	100.0	254	100.0

Note: Two-sided Fisher's exact test for independence, $p = 0.217$.

5.2.2 Motivations

Table 12 shows the endorsement rates for each motivation measure separately for unenrolled and enrolled applicants, while Table 13 shows the corresponding tests for differences between the rates for the two groups on each measure, along with an estimate of each difference and its confidence interval.

```
```{r}
#| label: tbl-motivations-rates
#| tbl-cap: Endorsement Rates for Motivation Measures by Variable and
#|          Enrollment Status Among All Eligible Applicants
#| message: false

FN <- paste("CI, Wilson score confidence interval for a proportion.")

bind_rows(dc_summary(data = Eligible_Applicants,
          vars = c("Motivation_PersonalConn", "Motivation_NeedSANE")),
          dc_summary(data = Eligible_Applicants %>% filter(Enrolled == 0),
          vars = c("Motivation_PersonalConn", "Motivation_NeedSANE")),
          dc_summary(data = Eligible_Applicants %>% filter(Enrolled == 1),
          vars = c("Motivation_PersonalConn", "Motivation_NeedSANE"))) %>%
  bind_cols(Enrolled = c("Overall", "Overall", "No", "No", "Yes", "Yes")) %>%
  mutate(No = as.character(No),
         Yes = as.character(Yes),
         Total = as.character(Total),
         Enrolled = factor(Enrolled, levels = c("Overall", "No", "Yes"))) %>%
  relocate(Variable, Enrolled) %>%
  arrange(Variable, Enrolled) %>%
  kable(format = "latex", booktabs = TRUE, digits = 2, align = "llrrrrrr",
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2,"Motivation (N)" = 3, " ",
                              "95% CI" = 2)) %>%
  collapse_rows(columns = 1:2, valign = "top", latex_hline = "major",
                row_group_label_position = "first") %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```
```

```
```{r}
#| label: tbl-motivations-diff
#| tbl-cap: Tests for Difference in Endorsement Rates of Motivation Measures
#|          Between Unenrolled and Enrolled Groups of Eligible Applicants
#| message: false

FN <- paste("CI, score confidence interval for a difference in independent",
            "proportions (unenrolled - enrolled), with continuity correction.")

bind_rows(dc_summary(data = Eligible_Applicants %>% filter(Enrolled == 0),
          vars = c("Motivation_PersonalConn", "Motivation_NeedSANE")),
          dc_summary(data = Eligible_Applicants %>% filter(Enrolled == 1),
          vars = c("Motivation_PersonalConn", "Motivation_NeedSANE"))) %>%
  bind_cols(Enrolled = c("Unenrolled", "Unenrolled", "Enrolled", "Enrolled")) %>%
  arrange(Variable, desc(Enrolled)) %>%
  relocate(Variable, Enrolled) %>%
  select(Variable, Enrolled, Yes, Total, Rate) %>%
  pivot_wider(id_cols = Variable, names_from = Enrolled,
              values_from = c(Yes, Total, Rate)) %>%
```
```

Table 12: Endorsement Rates for Motivation Measures by Variable and Enrollment Status Among All Eligible Applicants

		Motivation (N)				95% CI	
Variable	Enrolled	No	Yes	Total	Rate	LL	UL
Motivation_NeedSANE	Overall	85	242	327	0.74	0.69	0.78
	No	26	47	73	0.64	0.53	0.74
	Yes	59	195	254	0.77	0.71	0.82
Motivation_PersonalConn	Overall	194	133	327	0.41	0.35	0.46
	No	51	22	73	0.30	0.21	0.41
	Yes	143	111	254	0.44	0.38	0.50

Note: CI, Wilson score confidence interval for a proportion.

```
  rowwise() %>%
  mutate(p.test = tidy(prop.test(x = c(Yes_Unenrolled, Yes_Enrolled),
                                 n = c(Total_Unenrolled, Total_Enrolled),
                                 conf.level = 0.95))) %>%
  unnest(p.test) %>%
  rename(Chisq = statistic, p = p.value, df = parameter, LL = conf.low,
         UL = conf.high) %>%
  mutate(Difference = estimate1 - estimate2,
         p = display_num(p)) %>%
  select(Variable, Difference, LL, UL, Chisq, df, p) %>%
  kable(format = "latex", booktabs = TRUE, digits = 2, align = "lrrrrr",
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2, "95% CI" = 2,
                              "Test for Diff. in Proportions" = 3)) %>%
  collapse_rows(columns = 1:2, valign = "top", latex_hline = "major",
                row_group_label_position = "first") %>%
  footnote(kable_input = ., general = FN, footnote_as_chunk = TRUE,
           threeparttable = TRUE)
```

Table 13: Tests for Difference in Endorsement Rates of Motivation Measures Between Unenrolled and Enrolled Groups of Eligible Applicants

		95% CI		Test for Diff. in Proportions		
Variable	Difference	LL	UL	Chisq	df	p
Motivation_NeedSANE	-0.12	-0.25	0.01	3.90	1.00	0.048
Motivation_PersonalConn	-0.14	-0.27	-0.01	3.78	1.00	0.052

Note: CI, score confidence interval for a difference in independent proportions (unenrolled - enrolled), with continuity correction.

Caution

Notice that both p-values in Table 13 are very close to the $\alpha = 0.05$ cutoff. There is also a subtle discrepancy between the p-values and the 95% confidence intervals for the difference in propoortions: normally 0 should fall *outside* the CI if $p < 0.05$ and *inside* the CI if $p \geq 0.05$, but we see the exact opposite here. We are using a base R function to do the Wilson score tests and get the CIs, so I am certain it has been extensively vetted by competent professional statisticians and programmers. We seem to be right at some boundary between significant vs. non-significant differences in these proportions because 0 is extremely close to the upper limit of both CIs. I'm inclined to say that the bulk of the evidence here suggests there are differences for both of these measures because nearly all values in the CIs are negative (suggesting that enrolled applicants endorse both motivations more often than unenrolled, eligible applicants).

5.2.3 Potential Barriers

Table 14 shows the mean scores for both barrier measures separately for unenrolled and enrolled groups of eligible applicants, while Table 15 shows the t-tests for mean differences between the two groups, along with estimates and confidence intervals for the effect size.

````
```{r}
#| label: tbl-barriers-means
#| tbl-cap: Potential Barriers Means by Variable and Enrollment Status Among All
#|          Eligible Applicants
#| message: false

FN <- paste("N, number of eligible applicants;",
            "N_m, number of missing values;",
            "N_o, number of observed values.")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled),
         Enrolled = fct_expand(Enrolled, "Overall", after = 0)) %>%
  bind_rows(., Overall) %>%
  select(ID, Enrolled, Barrier_FO, Barrier_WR) %>%
  pivot_longer(cols = c(Barrier_FO, Barrier_WR), names_to = "Variable",
               values_to = "Value") %>%
  group_by(Variable, Enrolled) %>%
  summarize(N = n(),
            N_m = sum(is.na(Value)),
            N_o = sum(!is.na(Value)),
            Mean = mean(Value, na.rm = TRUE),
            SD = sd(Value, na.rm = TRUE),
            SE = SD/sqrt(N_o),
            LL = Mean - qt(0.975, df = N_o - 1)*SE,
            UL = Mean + qt(0.975, df = N_o - 1)*SE) %>%
  mutate(N = as.character(N),
         N_m = as.character(N_m),
         N_o = as.character(N_o)) %>%
  kable(format = "latex", booktabs = TRUE, digit = 2, align = "llrrrrrrrr",
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2, "Sample Size" = 3, " " = 3,
                              "95% CI" = 2)) %>%
  collapse_rows(columns = 1:2, valign = "top", latex_hline = "major",
                row_group_label_position = "first") %>%
  footnote(general = FN, footnote_as_chunk = TRUE, threeparttable = TRUE)
```
````

````
```{r}
#| label: tbl-barriers-ttests
#| tbl-cap: T-Tests Comparing Potential Barriers Measures Between Unenrolled
#|          and Enrolled Groups of Eligible Applicants
#| message: false

FN <- paste("Difference, difference in means (Unenrolled - Enrolled);",
            "LL, lower limit; UL, upper limit.")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled)) %>%
  select(Enrolled, Barrier_FO, Barrier_WR) %>%
  pivot_longer(cols = c(Barrier_FO, Barrier_WR), names_to = "Variable",
               values_to = "Value") %>%
  nest(.by = Variable) %>%
  mutate(t.test = map(data, ~tidy(t.test(Value ~ Enrolled, data = .x))),
         Hedges_G = map(data, ~tibble(hedges_g(Value ~ Enrolled, data = .x,
                                              pooled_sd = FALSE)))) %>%
  unnest(t.test) %>%
  rename(Difference = estimate, t = statistic, p = p.value, df = parameter,
         LL = conf.low, UL = conf.high) %>%
  unnest(Hedges_G) %>%
  rename(g_LL = CI_low, g_UL = CI_high) %>%
  select(Variable, Difference, LL, UL, Hedges_g, g_LL, g_UL, t, df, p) %>%
  mutate(p = display_num(p)) %>%
````
````

Table 14: Potential Barriers Means by Variable and Enrollment Status Among All Eligible Applicants

		Sample Size						95% CI	
Variable	Enrolled	N	N_m	N_o	Mean	SD	SE	LL	UL
Barrier_FO	Overall	327	0	327	2.41	1.00	0.06	2.30	2.52
	No	73	0	73	2.63	1.10	0.13	2.37	2.89
	Yes	254	0	254	2.34	0.96	0.06	2.22	2.46
Barrier_WR	Overall	327	0	327	3.28	0.99	0.05	3.17	3.39
	No	73	0	73	3.44	1.01	0.12	3.20	3.67
	Yes	254	0	254	3.23	0.98	0.06	3.11	3.35

Note: N, number of eligible applicants; N_m, number of missing values; N_o, number of observed values.

```
  kable(format = "latex", booktabs = TRUE, digit = 2,
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2, "Diff. 95% CI" = 2,
                              " ", "g 95% CI" = 2,
                              "Welch 2-sample t-test" = 3)) %>%
  footnote(general = FN, footnote_as_chunk = TRUE, threeparttable = TRUE)
```

Table 15: T-Tests Comparing Potential Barriers Measures Between Unenrolled and Enrolled Groups of Eligible Applicants

		Diff. 95% CI			g 95% CI		Welch 2-sample t-test		
Variable	Difference	LL	UL	Hedges_g	g_LL	g_UL	t	df	p
Barrier_FO	0.29	0.01	0.57	0.28	0.01	0.55	2.03	105.28	0.045
Barrier_WR	0.21	-0.06	0.47	0.20	-0.06	0.47	1.54	113.97	0.126

Note: Difference, difference in means (Unenrolled - Enrolled); LL, lower limit; UL, upper limit.

5.2.4 Emotional Readiness

Table 16 shows the mean scores for emotional readiness measures separately for unenrolled and enrolled groups of eligible applicants, while Table 17 shows the t-tests for mean differences between the two groups, along with estimates and confidence intervals for the effect size.

```
```{r}
#| label: tbl-ProQOL-means
#| tbl-cap: Emotional Readness Means by Variable and Enrollment Status Among
#|          Eligible Applicants
#| message: false

FN <- paste("Applicants with missing data were omitted from comutation of",
            "Means, SDs, SEs, and CIs via listwise deletion.",
            "N, number of eligible applicants;",
            "N_m, number of missing values;",
            "N_o, number of observed values.")

Eligible_Applicants %>%
  mutate(Enrolled = as_factor(Enrolled),
         Enrolled = fct_expand(Enrolled, "Overall", after = 0)) %>%
```
```

```
  bind_rows(., Overall) %>%
  select(ID, Enrolled, ProQOL_BO, ProQOL_CS, ProQOL_STS) %>%
  pivot_longer(cols = c(ProQOL_BO, ProQOL_CS, ProQOL_STS), names_to = "Variable",
               values_to = "Value") %>%
  group_by(Variable, Enrolled) %>%
  summarize(N = n(),
            N_m = sum(is.na(Value)),
            N_o = sum(!is.na(Value)),
            Mean = mean(Value, na.rm = TRUE),
            SD = sd(Value, na.rm = TRUE),
            SE = SD/sqrt(N_o),
            LL = Mean - qt(0.975, df = N_o - 1)*SE,
            UL = Mean + qt(0.975, df = N_o - 1)*SE) %>%
  mutate(N = as.character(N),
         N_m = as.character(N_m),
         N_o = as.character(N_o)) %>%
  kable(format = "latex", booktabs = TRUE, digit = 2, align = "llrrrrrrrr",
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2, "Sample Size" = 3, " " = 3,
                              "95% CI" = 2)) %>%
  collapse_rows(columns = 1:2, valign = "top", latex_hline = "major",
                row_group_label_position = "first") %>%
  footnote(general = FN, footnote_as_chunk = TRUE, threeparttable = TRUE)
```

Table 16: Emotional Readness Means by Variable and Enrollment Status Among Eligible Applicants

		Sample Size						95% CI	
Variable	Enrolled	N	N_m	N_o	Mean	SD	SE	LL	UL
ProQOL_BO	Overall	327	3	324	17.15	3.94	0.22	16.72	17.58
	No	73	1	72	17.35	4.38	0.52	16.32	18.38
	Yes	254	2	252	17.09	3.81	0.24	16.62	17.56
ProQOL_CS	Overall	327	3	324	45.35	3.68	0.20	44.95	45.76
	No	73	1	72	45.24	3.75	0.44	44.35	46.12
	Yes	254	2	252	45.39	3.66	0.23	44.93	45.84
ProQOL_STS	Overall	327	3	324	17.47	3.91	0.22	17.04	17.90
	No	73	1	72	17.56	4.06	0.48	16.60	18.51
	Yes	254	2	252	17.44	3.87	0.24	16.96	17.93

Note: Applicants with missing data were omitted from comutation of Means, SDs, SEs, and CIs via listwise deletion. N, number of eligible applicants; N_m, number of missing values; N_o, number of observed values.

```
```{r}
#| label: tbl-ProQOL-ttests
#| tbl-cap: T-Tests Comparing Emotional Readiness Measures Between Unenrolled
#|          and Enrolled Groups of Eligible Applicants
#| message: false

FN <- paste("Applicants with missing data were omitted via listwise deletion.",
            "Difference, difference in means (unenrolled - enrolled);",
            "LL, lower limit; UL, upper limit.")

Eligible_Applicants %>%
  select(Enrolled, ProQOL_BO, ProQOL_CS, ProQOL_STS) %>%
  pivot_longer(cols = c(ProQOL_BO, ProQOL_CS, ProQOL_STS), names_to = "Variable",
               values_to = "Value") %>%
  nest(.by = Variable) %>%
  mutate(t.test = map(data, ~tidy(t.test(Value ~ Enrolled, data = .x))),
         Hedges_G = map(data, ~tibble(hedges_g(Value ~ Enrolled, data = .x,
                                               pooled_sd = FALSE,
                                               na.action = na.omit)))) %>%
  unnest(t.test) %>%
  rename(Difference = estimate, t = statistic, p = p.value, df = parameter,
         LL = conf.low, UL = conf.high) %>%
  unnest(Hedges_G) %>%
  rename(g_LL = CI_low, g_UL = CI_high) %>%
  select(Variable, Difference, LL, UL, Hedges_g, g_LL, g_UL, t, df, p) %>%
```
```

```
  mutate(p = display_num(p)) %>%
  kable(format = "latex", booktabs = TRUE, digit = 2,
        format.args = list(nsmall = 2)) %>%
  kable_styling() %>%
  add_header_above(header = c(" " = 2, "Diff. 95% CI" = 2,
                              " ", "d 95% CI" = 2,
                              "Welch 2-sample t-test" = 3)) %>%
  footnote(general = FN, footnote_as_chunk = TRUE, threeparttable = TRUE)
```
```

**Table 17: T-Tests Comparing Emotional Readiness Measures Between Unenrolled and Enrolled Groups of Eligible Applicants**

| | | Diff. 95% CI | | | d 95% CI | | Welch 2-sample t-test | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Difference | LL | UL | Hedges_g | g_LL | g_UL | t | df | p |
| ProQOL_BO | 0.26 | -0.87 | 1.39 | 0.06 | -0.21 | 0.33 | 0.45 | 103.57 | 0.654 |
| ProQOL_CS | -0.15 | -1.14 | 0.84 | -0.04 | -0.30 | 0.22 | -0.31 | 112.54 | 0.760 |
| ProQOL_STS | 0.11 | -0.95 | 1.17 | 0.03 | -0.24 | 0.29 | 0.21 | 110.73 | 0.836 |

*Note:* Applicants with missing data were omitted via listwise deletion. Difference, difference in means (unenrolled - enrolled); LL, lower limit; UL, upper limit.

### 5.2.5 Conclusions

We did not detect differences between unenrolled versus enrolled applicants with respect to any of the background characteristics examined. Instead, those analyses primarily show that:

- Over half of each group had a bachelor's degree.
- The vast majority of applicants were licensed as registered nurses (RNs), currently employed as nurses, and employed full-time.
- The applicants in both groups had on average been employed as nurses less than 1.5 years.
- The applicants in both groups were roughly evenly split between practicing in the three types of settings examined.
- Very few applicants in either group had received any prior SANE training.

Enrolled applicants reported being motivated by a need for SANE services in their community more often than unenrolled applicants. Similarly, the enrolled applicants reported being motivated by a personal connection to sexual assault more often than unenrolled applicants. For both these measures there were subtle discrepancies between whether zero fell just inside versus just outside the 95% CI and the p-values fell just above versus below the $\alpha = 0.05$ cutoff. The American Statistical Association encourages basing conclusions on thoughtful evaluation of available evidence rather than solely on whether p-values pass a specific threshold [Wasserstein & Lazar (2016); Wasserstein-RN3844]. Here, the p-values for the two motivation comparisons lie just on either side of the =.05 cutoff but are not meaningfully different values (Gelman & Stern, 2006). We interpret the results as evidence for differences on both variables because the effect size ($\Delta p$) confidence intervals are very similar, span almost exclusively negative values (see Table 1), and are more scientifically informative than p-values (Cummins & Marks, 2020; Hubbard & Lindsay, 2008).

Regarding potential barriers to participation, we found that enrolled applicants reported slightly lower mean levels of family obligations than unenrolled applicants. Our best estimate is that it was a small effect size, but it could be anywhere from trivially small to medium. We did not
```

firmly establish the presence or direction an effect of enrollment status on mean levels of work responsibilities, but the point estimate of the effect size was only a bit smaller than the effect on family obligations and the CI is wide enough to suggest that there could be up to a medium size effect such that enrolled applicants have fewer work responsibilities than unenrolled applicants.

We did not detect any differences between unenrolled and enrolled applicants in their emotional readiness for the training program, as measured by burnout, compassion satisfaction, and secondary traumatic stress.

6 References

Cummins, K. M., & Marks, C. (2020). Farewell to bright-line: A guide to reporting quantitative results without the s-word. *Frontiers in Psychology*, *11*(815). https://doi.org/10.3389/fpsyg.2020.00815

Delacre, M., Lakens, D., Ley, C., Liu, L., & Leys, C. (2021). *Why Hedges' g*s based on the non-pooled standard deviation should be reported with Welch's t-test* [Preprint manuscript]. PsyArXiv. https://doi.org/10.31234/osf.io/tu6mp

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology*, *30*(1), 92–101. https://doi.org/10.5334/irsp.82

Dontje, K., & Campbell, R. (07/01/2021–06/30/2025). *Increasing access, recruitment, and retention of sexual assault nurse examiners in rural michigan* (Grant No. T96HP42059). Health Resources and Services Administration.

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician*, *60*(4), 328–331. https://doi.org/10.1198/000313006X152649

Hubbard, R., & Lindsay, R. M. (2008). Why p values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology*, *18*(1), 69–88. https://doi.org/10.1177/0959354307086923

Newcombe, R. G. (1998). Interval estimation for the difference between independent proportions: Comparison of eleven methods. *Statistics in Medicine*, *17*(8), 873–890. https://doi.org/10.1002/(SICI)1097-0258(19980430)17:8%3C873::AID-SIM779%3E3.0.CO;2-I

Newcombe, R. G. (2012). *Confidence intervals for proportions and related measures of effect size*. Chapman & Hall/CRC. https://doi.org/10.1201/b12670-1

Pierce, S. J. (2026). *SANETPA: Research compendium for a study of sexual assault nurse examiner training program attrition* (Version 1.0.1) [Reproducible Research Materials and Computer Program, R Package, Public Repository]. GitHub. https://github.com/sjpierce/SANETPA. Zenodo. https://doi.org/10.5281/zenodo.18643254

Stamm, B. H. (2010). *The concise ProQOL manual*. The ProQOL.org. https://ProQOL.org

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: Context, process, and purpose. *The American Statistician*, *70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, *22*(158), 209–212. https://doi.org/10.1080/01621459.1927.10502953

7 Software Information

This section documents information that is important for reproducibility. Most users will not need to read it. It is primarily here for use by the statistician on the team if we need to troubleshoot reproducibility issues because someone else is unable to get the same results from the same code. Start by checking for differences in package versions.

We used R as our main computing environment and Quarto scripts to enhance reproducibility. We used RStudio as the editor to interface with R and Quarto.

- Software chain: **qmd file > RStudio > Quarto > R > knitr > md file > Pandoc > tex file > TinyTeX > PDF file**.
- Source file: `RQ1_Analyses.qmd`
- Output file: `RQ1_Analyses_2026-02-14.pdf`
- Quarto 1.8.27 runs `*.qmd` files through R and knitr to produce `*.md` markdown files.
- Pandoc 3.6.3 converts markdown files (`*.md`) to other formats, including LaTeX (`*.tex`) and HTML (`*.html`) among others.
- TinyTeX compiles LaTeX files (`*.tex`) into PDF files. It should be viable to use MiKTeX or another LaTeX distribution instead.

7.1 Versions

This document was generated using the following computational environment and dependencies:

```
# Check and report whether we used TinyTex or other LaTeX software.
which_latex()
```

```
is_tinytex = TRUE. We used TeX Live 2025 (TinyTeX) with tlmgr 2025-11-06.
```

```
tlmgr revision 76773 (2025-11-06 20:43:29 +0100)
tlmgr using installation: C:/Users/pierces1/AppData/Roaming/TinyTeX
TeX Live (https://tug.org/texlive) version 2025
```

```
# Get R and R package version numbers in use.
devtools::session_info()
```

```
Warning in system2("quarto", "-V", stdout = TRUE, env = paste0("TMPDIR=", :
running command '"quarto"
TMPDIR=C:/Users/pierces1/AppData/Local/Temp/RtmpEnHqG7/file5728727a230d -V' had
status 1
```

```
- Session info ---------------------------------------------------------------
 setting  value
 version  R version 4.5.2 (2025-10-31 ucrt)
 os       Windows 11 x64 (build 26100)
 system   x86_64, mingw32
 ui       RTerm
 language (EN)
 collate  English_United States.utf8
 ctype    English_United States.utf8
 tz       America/New_York
 date     2026-02-14
 pandoc   3.6.3 @ C:/Program Files/RStudio/resources/app/bin/quarto/bin/tools/ (via rmarkdown)
 quarto   NA @ C:\\PROGRA~1\\Quarto\\bin\\quarto.exe

- Packages -------------------------------------------------------------------
 package      * version  date (UTC) lib source
 assertthat     0.2.1    2019-03-21 [1] CRAN (R 4.5.0)
 backports      1.5.0    2024-05-23 [1] CRAN (R 4.5.0)
 bayestestR     0.17.0   2025-08-29 [1] CRAN (R 4.5.1)
 broom        * 1.0.12   2026-01-27 [1] CRAN (R 4.5.2)
 cachem         1.1.0    2024-05-16 [1] CRAN (R 4.5.0)
```

```
 cli           3.6.5    2025-04-23 [1] CRAN (R 4.5.0)
 coda          0.19-4.1 2024-01-31 [1] CRAN (R 4.5.0)
 codetools     0.2-20   2024-03-31 [1] CRAN (R 4.5.0)
 datawizard    1.3.0    2025-10-11 [1] CRAN (R 4.5.1)
 devtools    * 2.4.6    2025-10-03 [1] CRAN (R 4.5.1)
 digest        0.6.39   2025-11-19 [1] CRAN (R 4.5.2)
 dplyr       * 1.2.0    2026-02-03 [1] CRAN (R 4.5.2)
 effectsize  * 1.0.1    2025-05-27 [1] CRAN (R 4.5.0)
 ellipsis      0.3.2    2021-04-29 [1] CRAN (R 4.5.0)
 emmeans       2.0.1    2025-12-16 [1] CRAN (R 4.5.2)
 estimability  1.5.1    2024-05-12 [1] CRAN (R 4.5.0)
 evaluate      1.0.5    2025-08-27 [1] CRAN (R 4.5.1)
 farver        2.1.2    2024-05-13 [1] CRAN (R 4.5.0)
 fastmap       1.2.0    2024-05-15 [1] CRAN (R 4.5.0)
 forcats     * 1.0.1    2025-09-25 [1] CRAN (R 4.5.1)
 fs            1.6.6    2025-04-12 [1] CRAN (R 4.5.0)
 generics      0.1.4    2025-05-09 [1] CRAN (R 4.5.0)
 ggplot2     * 4.0.2    2026-02-03 [1] CRAN (R 4.5.2)
 git2r         0.36.2   2025-03-29 [1] CRAN (R 4.5.0)
 glue          1.8.0    2024-09-30 [1] CRAN (R 4.5.0)
 gtable        0.3.6    2024-10-25 [1] CRAN (R 4.5.0)
 haven       * 2.5.5    2025-05-30 [1] CRAN (R 4.5.0)
 here        * 1.0.2    2025-09-15 [1] CRAN (R 4.5.1)
 hms           1.1.4    2025-10-17 [1] CRAN (R 4.5.1)
 htmltools     0.5.9    2025-12-04 [1] CRAN (R 4.5.2)
 httr          1.4.8    2026-02-13 [1] CRAN (R 4.5.2)
 insight       1.4.6    2026-02-04 [1] CRAN (R 4.5.2)
 janitor     * 2.2.1    2024-12-22 [1] CRAN (R 4.5.0)
 jsonlite      2.0.0    2025-03-27 [1] CRAN (R 4.5.0)
 kableExtra  * 1.4.0    2024-01-24 [1] CRAN (R 4.5.0)
 knitr       * 1.51     2025-12-20 [1] CRAN (R 4.5.2)
 later         1.4.5    2026-01-08 [1] CRAN (R 4.5.2)
 lattice       0.22-9   2026-02-09 [1] CRAN (R 4.5.2)
 lifecycle     1.0.5    2026-01-08 [1] CRAN (R 4.5.2)
 lubridate   * 1.9.5    2026-02-04 [1] CRAN (R 4.5.2)
 magrittr      2.0.4    2025-09-12 [1] CRAN (R 4.5.1)
 MASS          7.3-65   2025-02-28 [1] CRAN (R 4.5.0)
 Matrix        1.7-4    2025-08-28 [1] CRAN (R 4.5.1)
 MBESS         4.9.42   2026-01-08 [1] CRAN (R 4.5.2)
 memoise       2.0.1    2021-11-26 [1] CRAN (R 4.5.0)
 mnormt        2.1.2    2026-01-27 [1] CRAN (R 4.5.2)
 multcomp      1.4-29   2025-10-20 [1] CRAN (R 4.5.1)
 mvtnorm       1.3-3    2025-01-10 [1] CRAN (R 4.5.0)
 nlme          3.1-168  2025-03-31 [1] CRAN (R 4.5.0)
 otel          0.2.0    2025-08-29 [1] CRAN (R 4.5.1)
 parameters    0.28.3   2025-11-25 [1] CRAN (R 4.5.2)
 pbivnorm      0.6.0    2015-01-23 [1] CRAN (R 4.5.0)
 piercer     * 0.23.0   2025-09-07 [1] Github (sjpierce/piercer@7e53e10)
 pillar        1.11.1   2025-09-17 [1] CRAN (R 4.5.1)
 pkgbuild      1.4.8    2025-05-26 [1] CRAN (R 4.5.0)
 pkgconfig     2.0.3    2019-09-22 [1] CRAN (R 4.5.0)
 pkgload       1.5.0    2026-02-03 [1] CRAN (R 4.5.2)
 pROC          1.19.0.1 2025-07-31 [1] CRAN (R 4.5.1)
 processx      3.8.6    2025-02-21 [1] CRAN (R 4.5.0)
 PropCIs       0.3-0    2018-02-23 [1] CRAN (R 4.5.0)
 ps            1.9.1    2025-04-12 [1] CRAN (R 4.5.0)
 psych       * 2.6.1    2026-02-03 [1] CRAN (R 4.5.2)
 purrr       * 1.2.1    2026-01-09 [1] CRAN (R 4.5.2)
 quarto      * 1.5.1    2025-09-04 [1] CRAN (R 4.5.1)
 R6            2.6.1    2025-02-15 [1] CRAN (R 4.5.0)
 RColorBrewer  1.1-3    2022-04-03 [1] CRAN (R 4.5.0)
 Rcpp          1.1.1    2026-01-10 [1] CRAN (R 4.5.2)
 readr       * 2.1.6    2025-11-14 [1] CRAN (R 4.5.2)
 remotes       2.5.0    2024-03-17 [1] CRAN (R 4.5.0)
 rlang         1.1.7    2026-01-09 [1] CRAN (R 4.5.2)
 rmarkdown   * 2.30     2025-09-28 [1] CRAN (R 4.5.1)
 rprojroot     2.1.1    2025-08-26 [1] CRAN (R 4.5.1)
 rstudioapi    0.18.0   2026-01-16 [1] CRAN (R 4.5.2)
 S7            0.2.1    2025-11-14 [1] CRAN (R 4.5.2)
 sandwich      3.1-1    2024-09-15 [1] CRAN (R 4.5.0)
 SANETPA     * 1.0.1    2026-02-14 [1] Github (sjpierce/SANETPA@482345e)
 scales        1.4.0    2025-04-24 [1] CRAN (R 4.5.0)
 sessioninfo   1.2.3    2025-02-05 [1] CRAN (R 4.5.0)
 snakecase     0.11.1   2023-08-27 [1] CRAN (R 4.5.0)
 stringi       1.8.7    2025-03-27 [1] CRAN (R 4.5.0)
 stringr     * 1.6.0    2025-11-04 [1] CRAN (R 4.5.2)
 survival      3.8-6    2026-01-16 [1] CRAN (R 4.5.2)
 svglite       2.2.2    2025-10-21 [1] CRAN (R 4.5.1)
 systemfonts   1.3.1    2025-10-01 [1] CRAN (R 4.5.1)
 texreg        1.39.5   2025-12-22 [1] CRAN (R 4.5.2)
 textshaping   1.0.4    2025-10-10 [1] CRAN (R 4.5.1)
 TH.data       1.1-5    2025-11-17 [1] CRAN (R 4.5.2)
 tibble      * 3.3.1    2026-01-11 [1] CRAN (R 4.5.2)
 tidyr       * 1.3.2    2025-12-19 [1] CRAN (R 4.5.2)
 tidyselect    1.2.1    2024-03-11 [1] CRAN (R 4.5.0)
 tidyverse   * 2.0.0    2023-02-22 [1] CRAN (R 4.5.0)
 timechange    0.4.0    2026-01-29 [1] CRAN (R 4.5.2)
 tinytex       0.58     2025-11-19 [1] CRAN (R 4.5.2)
 tzdb          0.5.0    2025-03-15 [1] CRAN (R 4.5.0)
 usethis     * 3.2.1    2025-09-06 [1] CRAN (R 4.5.1)
 vctrs         0.7.1    2026-01-23 [1] CRAN (R 4.5.2)
 viridisLite   0.4.3    2026-02-04 [1] CRAN (R 4.5.2)
```

```
 withr        3.0.2    2024-10-28 [1] CRAN (R 4.5.0)
 xfun         0.56     2026-01-18 [1] CRAN (R 4.5.2)
 xml2         1.5.2    2026-01-17 [1] CRAN (R 4.5.2)
 xtable       1.8-4    2019-04-21 [1] CRAN (R 4.5.0)
 yaml         2.3.12   2025-12-10 [1] CRAN (R 4.5.2)
 zoo          1.8-15   2025-12-15 [1] CRAN (R 4.5.2)

 [1] C:/Users/pierces1/AppData/Local/R/win-library/4.5
 [2] C:/Program Files/R/R-4.5.2/library
 * -- Packages attached to the search path.

------------------------------------------------------------------------------
```

7.2 Git Details

The current Git commit details and status are:

```
git_report()
```

```
Local:    main P:/Consulting/Cases_1600-1799/C1788/SANETPA
Remote:   main @ origin (https://github.com/sjpierce/SANETPA.git)
Head:     [482345e] 2026-02-14: Update version number.



Untracked files:
    Untracked:  scripts/Production_Run.rmarkdown
    Untracked:  scripts/RQ1_Analyses.rmarkdown
```

This is useful because it tells us exactly which commit in the Git history we would need to be using to make sure we are running the exact same code. Sometimes another person is not using the most current code, or has changed the code in some way since it was last committed.

Tip

- Untracked files are files located in the repository that Git has not been told to entirely ignore, but have also not been committed into the version history.
- Unstaged changes to files indicate that some of the contents have been modified since the last time the file was committed to Git. In production runs, we want the Git output to not show any unstaged changes to key files!