Variable Continuity Report

AUTHOR
Carly Adams, MS

PUBLISHED
March 11, 2025

Introduction

This analysis examines whether the survey question represented by V250 in 2019 and 2023 is consistent enough to be trended across survey years given the changes in wording of response options.

Question:
"During the past 30 days, how many times did you drive a car or other vehicle when you had been drinking alcohol?"

Changes in Response Option Wording

One potential source of inconsistency is the change in response option wording between 2019 and 2023:

Response Option	2019	2023
A	I did not drive a car or other vehicle during the past 30 days	I did not drive a car or other vehicle during the past 30 days
B	**0 times**	**I drove a car or other vehicle, but not when I had been drinking alcohol**
C	1 time	1 time
D	2 or 3 times	2 or 3 times
E	4 or 5 times	4 or 5 times
F	6 or more times	6 or more times

The key change is the change in words **"0 times"** meaning they have drove but not while drinking, changed to "**I drove a car or other vehicle, but not when I had been drinking alcohol"** which we believe means the same thing but written out. This may impact how respondents interpret and answer the question, potentially making a significant difference in interpretation.

Data Import and Recoding

```
# Load raw data
raw_data <- read.xlsx("raw_yrbs_allyears_statewide.xlsx")

# Recode Responses for Consistency
filtered_data_raw <- raw_data %>%
  mutate(Response = case_when(
    Survey_Year == 2019 & V250 == 1 ~ "Didn't Drive at all",
    Survey_Year == 2019 & V250 == 2 ~ "19-0 times/23-Drove Not Drunk",
    Survey_Year == 2019 & V250 == 3 ~ "1 time",
    Survey_Year == 2019 & V250 == 4 ~ "2 or 3 times",
    Survey_Year == 2019 & V250 == 5 ~ "4 or 5 times",
    Survey_Year == 2019 & V250 == 6 ~ "6 or more times",
    Survey_Year == 2023 & V250 == 1 ~ "Didn't Drive at all",
    Survey_Year == 2023 & V250 == 2 ~ "19-0 times/23-Drove Not Drunk",
    Survey_Year == 2023 & V250 == 3 ~ "1 time",
    Survey_Year == 2023 & V250 == 4 ~ "2 or 3 times",
    Survey_Year == 2023 & V250 == 5 ~ "4 or 5 times",
    Survey_Year == 2023 & V250 == 6 ~ "6 or more times",
```

```
        TRUE ~ NA_character_
    )) %>%
    filter(!is.na(Response))
```

Interpretation:

The response option wording for 2023 were **renamed for now**, from 2019. This step ensures comparability by aligning them before analysis and easier for visualization.

YRBS Complex Survey Design Setup

```
# Define YRBS complex survey design
survey_design <- filtered_data_raw %>%
  as_survey_design(
    ids = Primary_Samp_Unit,
    strata = Stratum,
    weights = Final_Weight,
    nest = TRUE
  )
```

Interpretation:

The dataset is now structured with survey weights, stratification, and PSU identifiers to maintain appropriate variance estimation.

Prevalence Estimates and Visualization

```
# Calculate prevalence estimates
prevalence_results <- survey_design %>%
  group_by(Survey_Year, Response) %>%
  summarise(
    Prevalence = survey_mean(na.rm = TRUE,
                             proportion = TRUE,
                             prop_method = "logit",
                             vartype = "ci"),
    Numerator = sum(Response == Response, na.rm = TRUE),
    Denominator = sum(!is.na(Response), na.rm = TRUE)
  ) %>%
  rename(
    `CI Low` = Prevalence_low,
    `CI High` = Prevalence_upp
  ) %>%
  mutate(
    Prevalence = round(Prevalence * 100, 1),
    `CI Low` = round(`CI Low` * 100, 1),
    `CI High` = round(`CI High` * 100, 1)
  )
```

```
ggplot(prevalence_results, aes(x = factor(Response,
                                          levels = c(
    "Didn't Drive at all",
    "19-0 times/23-Drove Not Drunk",
    "1 time",
    "2 or 3 times",
    "4 or 5 times",
    "6 or more times"
  )),
  y = Prevalence,
```

```
    fill = factor(Survey_Year, levels = c(2019, 2023)))) +
    geom_bar(stat = "identity", position =
            position_dodge(width = 0.9)) +
    geom_errorbar(aes(ymin = `CI Low`, ymax = `CI High`),
                width = 0.2, position =
                    position_dodge(width = 0.9)) +
    scale_fill_manual(values = c("2019" = "#F8766D",
                                "2023" = "#00BFC4")) +
    labs(title = "Change from '0 times' to 'Drove but not drunk' Words -
    Response Distributions (2019 vs 2023)",
        x = "Response Options",
        y = "Prevalence Estimate (%)",
        fill = "Survey Year") +
    theme_minimal() +
    theme(plot.title = element_text(hjust = 0.5,
                                    size = 14,
                                    face = "bold")) +
    coord_flip()
```
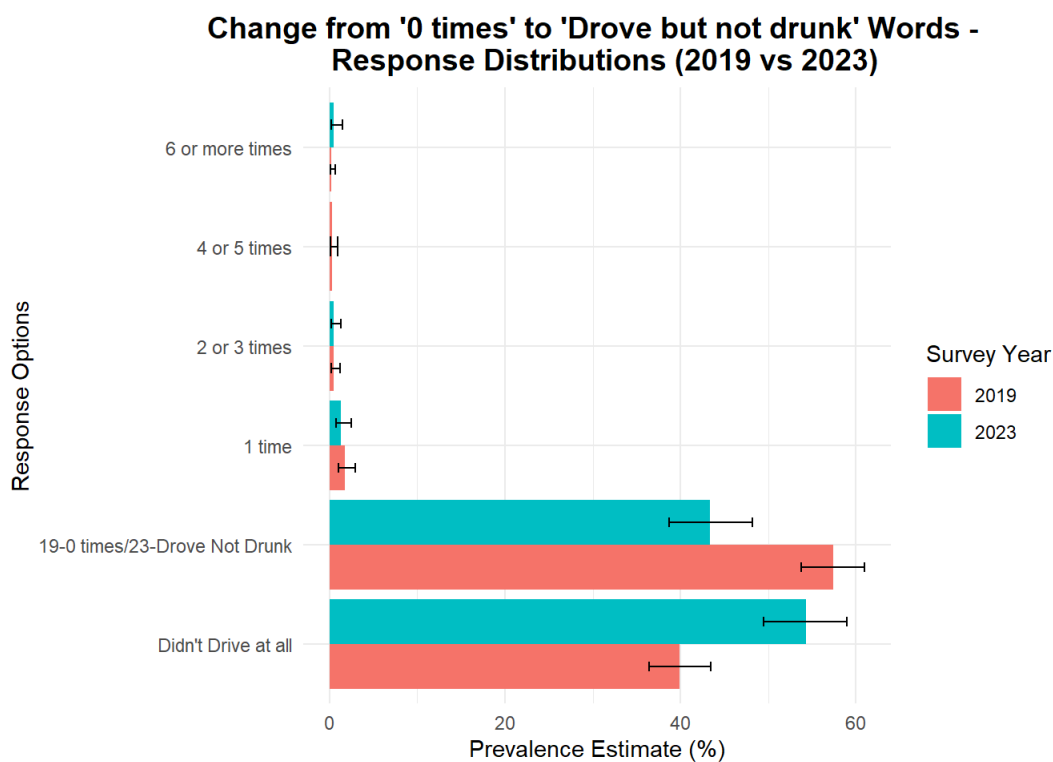

Interpretation:

The bar chart displays prevalence estimates for each response category in 2019 and 2023, accounting for survey weights. The change in wording from **"0 times"** to **"Drove but not drunk"** may have influenced how students categorized their behavior, which is reflected in the differences in response distributions.

Rao-Scott Chi-Square Test

The **Rao-Scott Chi-Square test** was selected for statistical significance testing in this analysis because it accounts for the **complex survey design** of the YRBS data. The **Centers for Disease Control and Prevention (CDC) also uses this method** when analyzing YRBS data, as it properly adjusts for **stratification, clustering, and survey weights**—all critical factors in obtaining valid results.

The Rao-Scott test **corrects for the design effects** present in the YRBS, ensuring accurate significance testing even when **sample sizes vary between survey years**. Given these considerations, this test provides the most appropriate approach for evaluating variable continuity across

survey cycles.

```
chi_test <- svychisq(~ Response + Survey_Year, survey_design)
chi_test
```

```
    Pearson's X^2: Rao & Scott adjustment

data:  NextMethod()
F = 7.2861, ndf = 3.785, ddf = 454.202, p-value = 1.687e-05
```

Interpretation:

Test Statistic (F) = 7.2861

- This is the adjusted chi-square statistic that tells us **how strong the difference is** between response distributions in 2019 and 2023.

- A **higher F-value** indicates a **stronger difference**.

- In general, an F-statistic above **4 or 5** suggests a meaningful difference—**7.2861 is significant**, meaning the response distributions **differ significantly**.

Numerator Degrees of Freedom (ndf) = 3.785

- This reflects the **number of independent comparisons** being made while adjusting for the survey design.

- Because of weighting and clustering, the effective degrees of freedom is slightly fractional rather than a whole number.

Denominator Degrees of Freedom (ddf) = 454.202

- This represents the **available information for statistical inference**, adjusted for the survey design.

- A **higher ddf (e.g., >100)** means that the test has **a strong basis for detecting real differences** rather than random noise.

p-value = 1.687e-05

- This is **far below 0.05**, meaning the probability of observing these differences **by random chance alone is extremely low**.

- The response distributions between 2019 and 2023 are **statistically different**.

This suggests that **the change in response wording significantly impacted how students reported their behavior**.

Logistic Regression Results

The **logistic regression model** was used to check if responses in **2023 were significantly different from those in 2019**. The model predicts response patterns based on the survey year while accounting for the **survey design (weights, clustering, stratification)**.

```
# Convert Response to an ordered factor
survey_design <- survey_design %>%
  mutate(Response = factor(Response, levels = c(
    "Didn't Drive at all",
    "19-0 times/23-Drove Not Drunk",
    "1 time",
    "2 or 3 times",
    "4 or 5 times",
    "6 or more times"
    ), ordered = TRUE))


logit_model <- svyglm(as.numeric(Response) ~ factor(Survey_Year),
                  design = survey_design,
                  family = quasipoisson(link = "log"))
```

```
summary(logit_model)
```

```
Call:
svyglm(formula = as.numeric(Response) ~ factor(Survey_Year),
    design = survey_design, family = quasipoisson(link = "log"))

Survey design:
Called via srvyr

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)              0.49777    0.01272  39.126  < 2e-16 ***
factor(Survey_Year)2023 -0.09121    0.02465  -3.701 0.000327 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasipoisson family taken to be 0.2507421)

Number of Fisher Scoring iterations: 4
```

Interpretation:

Estimate (-0.09121) – This tells us the **direction and size of the difference** between 2019 and 2023 responses.

- A **positive value** would mean that responses were slightly **higher in 2023 than in 2019** (i.e., students were more likely to choose a response further down the scale).

- Since the estimate is **negative**, it suggests that **students were slightly less likely to report higher drinking and driving frequency in 2023 compared to 2019**.

Standard Error (0.02465) – This measures **how precise** the estimate is.

- The **standard error (SE)** tells us **how much variation** there is in our estimate due to sampling. It shows how precise our results are.

- Since this value is **relatively small**, it means that **our estimate is precise** and **unlikely to vary much if we repeated the survey**.

- If the SE were **greater than 0.1 or 0.2**, it would indicate a **much higher level of uncertainty**, making the estimate less reliable.

p-value (0.000327)

- This tells us **if the difference between 2019 and 2023 is real or just due to random chance**.

- A **very small p-value (less than 0.05)** means that the difference is **statistically significant**—it is **unlikely to have happened by chance**.

- Here, **0.000327 is extremely small**, meaning that the difference is **highly significant**.

What This Means for Our Data

- **Responses changed significantly** between 2019 and 2023.

- The shift is **not due to random variation**—it is a **real difference**.

- This confirms that **V250 (2019) and V250 (2023) should not be treated as the same variable for trending**.

- However, **a new V-code was not assigned** for 2023. This could impact the accuracy of trend interpretations.

By using this model, we highlight that **a new V-code should be created for 2023** to maintain trend integrity and avoid misrepresenting the data.

Subgroup Analysis

Chi-square tests were conducted separately within demographic subgroups (Sex, Grade, Race_Group_3) to determine whether differences persisted across populations.

```r
# Look at Subgroups Chi Square test

subgroups <- c("Sex", "Grade", "Race_Group_3")
subgroup_results <- list()

for (subgroup in subgroups) {
  filtered_data <- filtered_data_raw %>%
    filter(!is.na(!!sym(subgroup)))
  subgroup_survey_design <- filtered_data %>%
    as_survey_design(
      ids = Primary_Samp_Unit,
      strata = Stratum,
      weights = Final_Weight,
      nest = TRUE
    )
  chi_test <- svychisq(
    as.formula(paste("~ Response + Survey_Year")), subgroup_survey_design)
  subgroup_results[[subgroup]] <- chi_test
}

# Print results
subgroup_results
```

$Sex

```
	Pearson's X^2: Rao & Scott adjustment

data:  NextMethod()
F = 6.8116, ndf = 3.818, ddf = 458.162, p-value = 3.502e-05
```

$Grade

```
	Pearson's X^2: Rao & Scott adjustment

data:  NextMethod()
F = 6.9206, ndf = 3.7442, ddf = 449.3050, p-value = 3.381e-05
```

$Race_Group_3

```
	Pearson's X^2: Rao & Scott adjustment

data:  NextMethod()
F = 7.3502, ndf = 3.7718, ddf = 452.6116, p-value = 1.56e-05
```

Interpretation:

Subgroup	F-statistic	df	p-value
Sex	6.8116	3.818	3.502e-05
Grade	6.9206	3.7442	3.381e-05
Race_Group_3	7.3502	3.7718	1.56e-05

- In **every subgroup**, the differences remain **highly significant**.

- This suggests that **the issue is not isolated to a particular demographic group but reflects a broader inconsistency in the response structure**.

Conclusion

- The **Chi-Square and Logistic Regression** models show **significant differences** between 2019 and 2023 responses.

- These findings suggest that **V250 (2019) and V250 (2023) should not be trended together**.

- The change in responses indicates that students interpreted the question differently or that actual behavior changed drastically between years.

- It is **unlikely** that actual behavior changed that much between 2019 and 2023 for these two response options. Instead, it is **far more likely** that the response wording change led to differences in how students interpreted and categorized their behavior.

- The **reclassification of "0 times"** to a phrase implying **some level of driving but no intoxication** may have made students **hesitate or select different responses** than they would have previously.

- **A new V-code was not created for 2023**, which may impact long-term trend analyses.

Thus, **the 2023 version should receive a new V-code to preserve trend integrity**, ensuring that changes in responses are accurately represented over time.