Scale Psychometrics and Preliminary Analyses

KN

June 10, 2022

Table of Contents

[Sample 2](#_Toc105749530)

[Wave 1 Demographics 2](#_Toc105749531)

[Attrition 3](#_Toc105749532)

[Observed Scales 4](#_Toc105749533)

[Family Conflict 4](#_Toc105749534)

[Exposure to and Victim of Abuse 5](#_Toc105749535)

[Sibling Aggression 6](#_Toc105749536)

[Latent Variable Scales 7](#_Toc105749537)

[Finding Summary 8](#_Toc105749538)

[Sexual Harassment 9](#_Toc105749539)

[Depression 10](#_Toc105749540)

[Lack of Empathy 11](#_Toc105749541)

[Substance Use 11](#_Toc105749542)

[Delinquent Behavior 12](#_Toc105749543)

[School Belonging 13](#_Toc105749544)

[School Belonging 14](#_Toc105749545)

[Delinquency 16](#_Toc105749546)

[Model Variables 18](#_Toc105749547)

[Table 1. Correlations and Descriptive Statistics 18](#_Toc105749548)

[Intraclass Correlations 18](#_Toc105749549)

[Sexual Harassment Descriptives 19](#_Toc105749550)

[Race 19](#_Toc105749551)

[Gender 19](#_Toc105749552)

[School\_Belonging 19](#_Toc105749553)

[Overall 20](#_Toc105749554)

# Sample

## Wave 1 Demographics

|  | n | Percent |
| --- | --- | --- |
| Gender |  |  |
| Male | 797 | 51 |
| Female | 761 | 49 |
| Missing Gender | 6 | 0 |
| Race |  |  |
| Am Indian | 18 | 1 |
| Black | 460 | 29 |
| Asian | 36 | 2 |
| Hispanic | 605 | 39 |
| White | 298 | 19 |
| Pac Island | 1 | 0 |
| Multiple | 146 | 9 |
| Age at W1 |  |  |
| 11 years | 1,213 | 78 |
| 12 years | 320 | 20 |
| 13 years | 21 | 1 |
| Missing Age | 10 | 1 |
| Grade at W1 |  |  |
| 5th grade | 10 | 1 |
| 6th grade | 1,545 | 99 |
| 7th grade | 1 | 0 |
| Missing Grade | 8 | 1 |
| Academic Grades at W1 |  |  |
| Mostly D's & F's (< 64) | 25 | 2 |
| Mostly C's & D's (65-69) | 56 | 4 |
| Mostly C's (70-74) | 42 | 3 |
| Mostly B's & C's (75-79) | 214 | 14 |
| Mostly B's (80-84) | 94 | 6 |
| Mostly A's & B's (85-90) | 629 | 40 |
| Mostly A's (90-100) | 297 | 19 |
| Missing Academic Grades | 207 | 13 |
| Mom’s Education |  |  |
| Less than high school | 246 | 16 |
| High school diploma or GED | 362 | 23 |
| Some college | 269 | 17 |
| Graduated from college | 294 | 19 |
| Some graduate school | 71 | 5 |
| Graduate or professional | 64 | 4 |
| Missing Mom’s Education | 258 | 16 |
| Dad’s Education |  |  |
| Less than high school | 269 | 17 |
| High school diploma or GED | 352 | 23 |
| Some college | 218 | 14 |
| Graduated from college | 241 | 15 |
| Some graduate school | 65 | 4 |
| Graduate or professional | 64 | 4 |
| Missing Dad’s Education | 355 | 23 |

## Attrition

**Completed Surveys by Wave**

| Wave | n | Retention |
| --- | --- | --- |
| 1.00 | 1,697 | 1.00 |
| 2.00 | 1,461 | 0.86 |
| 3.00 | 1,186 | 0.70 |
| 4.00 | 1,010 | 0.60 |

**Youth Participation Pattern**

In the Pattern column, the four values represent, in order, participation in Waves 1 - 4 with 1 = took survey, 0 = attrition.

| Pattern | n | Percent |
| --- | --- | --- |
| 1, 1, 1, 1 | 859 | 51.00 |
| 1, 1, 0, 0 | 280 | 16.00 |
| 1, 1, 1, 0 | 249 | 15.00 |
| 1, 0, 0, 0 | 133 | 8.00 |
| 1, 1, 0, 1 | 73 | 4.00 |
| 1, 0, 1, 1 | 53 | 3.00 |
| 1, 0, 0, 1 | 25 | 1.00 |
| 1, 0, 1, 0 | 25 | 1.00 |

# Observed Scales

We are proposing three sources of hostility at home:

* Family Conflict - 3 items each on a 5-point scale
* Abuse - 3 dichotomous items (1/0)
* Sibling Aggression - 2 items each on a 6-point scale

All three sources have too few items to create a meaningful latent variable. A scale score was created for each hostility source. The mean score was calculated for Family Conflict and Sibling Aggression scales whereas the sum score was calculated for the Abuse scale. Participants must respond to all items in a scale in order to receive a score.

The home hostility scales are based on an underlying continuum of no experiences to many experiences, so the scale scores will be treated as continuous variables in the analysis. The distribution and reliability information for each source is presented below.

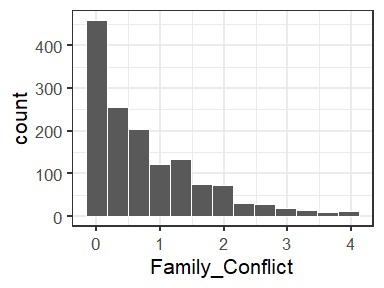
## Family Conflict

The Family Conflict scale consists of 3 items, each with response codes 0 = never, 1 = seldom, 2 = sometimes, 3 = often, 4 = always.

* How often is there yelling, quarreling, or arguing in your household?
* How often do family members lose their temper or blow up for no good reason?
* How often are there physical fights in the household, like people hitting, shoving, throwing?

**Score Distribution**

Due to missing data on all 3 items, 126 (8%) students did not receive a scale score.



**Reliability**

* = 0.87
* = 0.81

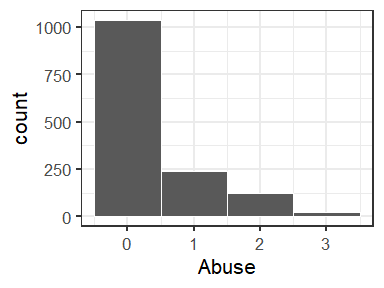
## Exposure to and Victim of Abuse

The Abuse scale consists of 3 items with response codes 0 = no and 1 = yes.

* Before you were 10 years old, did you ever have injuries such as bruises, cuts or broken bones, as a result of being spanked, struck, or shoved by your parents, guardians or parent’s partner?
* Before you were 10 years old, did you see or hear one of your parents or guardians being hit, slapped, punched, shoved, kicked, or otherwise physically hurt by their spouse or partner?
* Before you were 10 years old, did anyone ever force you to have sex or do something sexual that you did not want to do?

**Score Distribution**

Due to missing data on all 3 items, 144 (9%) students did not receive a scale score.



**Reliability**

* = 0.78
* = 0.57

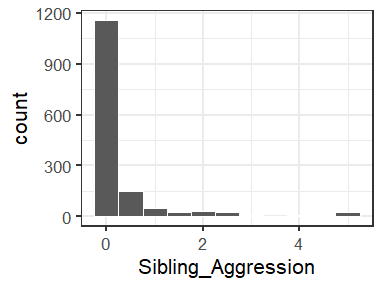
## Sibling Aggression

The Sibling Aggression scale consists of 2 items with response codes 0 = never, 1 = once in a while, 2 = about once a week, 3 = 1 or 2 days a week, 4 = Most days, 5 = Everyday

* I am scared that my sister or brother will hurt me bad someday.
* My sister or brother beats me up.

**Score Distribution**

Students were instructed to skip the sibling questions if they did not have a sibling. Nonetheless, only 80 (5%) students did not receive a total score.



**Reliability**

* = 0.85
* = 0.70

# Latent Variable Scales

In contrast to the Hostility at Home scales where we use the observed responses, the other scales are incorporated into the final model as latent variables.

* Sexual Harassment - 6 items, all dichotomoized (1/0); removed items 5 (poor conceptual fit) and 8-12 (low frequencies)
* Depression - 8 items (removed item 6 [happy])
* Lack of Empathy - 5 items (all reverse coded)
* Delinquent Behavior - 7 items (response categories collapsed to dichotomized to Never and 1+)
* Substance Use - 8 items (response categories collapsed to Never, 1-2, 3+)
* School Belonging - 4 items

To use these scales in our model, we need to demonstrate the survey items 1) measured the constructs adequately and 2) did so consistently over time. For each scale, a one-factor confirmatory factor model was fit at each wave for each scale to assess overall model fit and identify any problematic items. Additionally, alpha and omega reliabilities were calculated based on the CFA models.

Building on the preliminary CFAs, we establish longitudinal measurement invariance by fitting models that sequentially impose stricter constraints. The configural model estimates all parameters freely across time, except those described in the next paragraph to identify the model. The metric model is the same as the configural but the factor loadings are constrained to be equal over time. The scalar model constrains both the factor loadings and the item thresholds. Finally, the strict model constrains the loadings, thresholds, and unique variances (i.e. item residual variances). Only when strict invariance is achieved can we conclude that changes in the latent means over time are due to changes in the population rather than the product of measurement error.

There are a variety of common indicators to evaluate overall model fit for both the CFA and the measurement invariance models. One caveat is these are guidelines, not definitive cutoffs, and were derived with continuous indicators and maximum likelihood estimation (Kline, 2016). More guidance is needed for ordinal items and WLSMV estimation, which is what we use.

* Non-significant test ( > .05), with larger sample sizes, such as ours, the test is commonly significant
* CFI and TLI > .95
* RMSEA < .06
* SRMR < .08

For comparing between the configural, metric, scalar, and strict models to establish measurement invariance, a difference test is the traditional, albeit overly conservative, approach. Criteria for differences in the fit indices have also been suggested (Chen, 2007; Rutkowski & Svetina, 2014). As with the overall fit measures, more research is needed in this area with ordinal items, such as those in our scales.

* Non-significant difference test ( > .05)
* CFI and TLI < - .01
* RMSEA < .01
* SRMR < .01

We use guidance from Liu et al. (2017) for identifying and estimating the models. To identify the location and scale of the common factor (e.g., Depression) at each time point, the first item’s intercept and factor loading were fixed to 0 and 1, respectively. The location and scale of each latent item-response underlying the ordinal responses were identified using the “theta” parameterization, and the identification constraints proposed by Millsap & Tein (2004). Lastly, we used a DWLS estimator with robust standard errors and a mean and variance adjusted test statistic (i.e., WLSMV). The models were fit in R using the lavaan package and pairwise deletion for missing data.

## Finding Summary

* Sexual Harassment - Of the original 12 items, item 5 was removed as a poor conceptual fit (homophobic name calling) and items 8-12 were removed due to low frequencies overall or by group. The remaining 6 items had good CFA model fit and reliabiliy > .70 at each wave. With dichotomous items the metric and scalar models are equivalent. SRMR is a bit high even in the configural model. This could be corrected by estimating additional parameters. Which parameter, however, would be empirically driven which limits generalizability. Nonetheless, the consistency over the models suggests measurement invariance.
* Depression - Item 6 (happy) had a loading near 0 and was removed. The configural model had trouble converging at first. Using different identification constraints I determined that 1) Item 3 (worried) had a much higher frequency of 0 responses at Wave 1 than the other waves and 2) Item 1 (sad) is a poor marker variable. The model converged after adjusting the model to use item 2 as the marker and constrain threshold 4 instead of threshold 1. There is evidence of good model fit and measurment invariance. These adjustments were also made to the CFAs, which produced adequate model fit and reliability > .85 at each wave.
* Lack of Empathy - RMSEA is a bit high at each wave, but the other indices suggest good fit and reliability > .70 The most constrained model (“strict”) showed good fit based on the fit indices, and there was no change in fit as the models became more constrained. This suggests measurement invariance. However, the tests were significant for overall fit and the difference tests were ambiguous.
* Substance Use - The models converged after collapsing response categories to Never, 1-2, and 3+. Like the sexual harassment scale the SRMR is a bit high, but model fit is consistent as constraints are added.
* Delinquent Behavior - Initially collapsed response categories to Never, 1-2 and 3+ due to low frequencies in certain groups of students (e.g., Black). The latent growth models, however, showed poor fit so the items were dichotomized.
* School Belonging - Same as Lack of Empathy

## Sexual Harassment

**CFA Model Fit**

| Wave | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1,534.00 | 12.00 | 14.58 | 9.00 | 0.10 | 1.00 | 0.02 | [0.00, 0.04] | 0.05 |
| 2 | 1,348.00 | 12.00 | 25.39 | 9.00 | 0.00 | 0.98 | 0.04 | [0.02, 0.05] | 0.09 |
| 3 | 1,089.00 | 12.00 | 8.80 | 9.00 | 0.46 | 1.00 | 0.00 | [0.00, 0.03] | 0.04 |
| 4 | 992.00 | 12.00 | 22.58 | 9.00 | 0.01 | 0.99 | 0.04 | [0.02, 0.06] | 0.06 |

**Reliability**

| wave | alpha | omega |
| --- | --- | --- |
| 1 | 0.94 | 0.77 |
| 2 | 0.91 | 0.73 |
| 3 | 0.92 | 0.75 |
| 4 | 0.94 | 0.78 |

**Measurement Invariance Model Fit**

| model | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Configural | 732.00 | 93.00 | 239.94 | 207.00 | 0.06 | 0.98 | 0.01 | [0.00, 0.02] | 0.11 |
| Metric/Scalar | 732.00 | 79.00 | 257.15 | 221.00 | 0.05 | 0.98 | 0.01 | [0.00, 0.02] | 0.11 |
| Strict | 732.00 | 63.00 | 278.21 | 237.00 | 0.03 | 0.98 | 0.02 | [0.00, 0.02] | 0.11 |

**Measurement Invariance Model Comparison**

| constraint\_added | delta.chisq | delta.df | delta.pvalue | delta.cfi | delta.rmsea | delta.srmr |
| --- | --- | --- | --- | --- | --- | --- |
| loadings | 21.930 | 14 | 0.080 | -0.002 | 0.005 | 0.004 |
| residual variance | 26.394 | 16 | 0.049 | -0.005 | 0.006 | 0.004 |

## Depression

**CFA Model Fit**

| Wave | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1,409.00 | 40.00 | 227.79 | 20.00 | 0.00 | 0.98 | 0.09 | [0.08, 0.10] | 0.05 |
| 2 | 1,304.00 | 40.00 | 179.31 | 20.00 | 0.00 | 0.98 | 0.08 | [0.07, 0.09] | 0.04 |
| 3 | 1,077.00 | 40.00 | 195.88 | 20.00 | 0.00 | 0.98 | 0.09 | [0.08, 0.10] | 0.04 |
| 4 | 957.00 | 40.00 | 259.52 | 20.00 | 0.00 | 0.98 | 0.11 | [0.10, 0.12] | 0.05 |

**Reliability**

| wave | alpha | omega |
| --- | --- | --- |
| 1 | 0.85 | 0.81 |
| 2 | 0.85 | 0.83 |
| 3 | 0.86 | 0.86 |
| 4 | 0.87 | 0.87 |

**Measurement Invariance Model Fit**

| model | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Configural | 1,564.00 | 214.00 | 1,102.15 | 410.00 | 0.00 | 0.98 | 0.03 | [0.03, 0.04] | 0.04 |
| Metric | 1,564.00 | 193.00 | 1,145.87 | 431.00 | 0.00 | 0.98 | 0.03 | [0.03, 0.03] | 0.04 |
| Scalar | 1,564.00 | 124.00 | 1,211.64 | 500.00 | 0.00 | 0.98 | 0.03 | [0.03, 0.03] | 0.04 |
| Strict | 1,564.00 | 100.00 | 1,265.78 | 524.00 | 0.00 | 0.98 | 0.03 | [0.03, 0.03] | 0.04 |

**Measurement Invariance Model Comparison**

| constraint\_added | delta.chisq | delta.df | delta.pvalue | delta.cfi | delta.rmsea | delta.srmr |
| --- | --- | --- | --- | --- | --- | --- |
| loadings | 40.52 | 21.00 | 0.01 | 0.00 | 0.00 | 0.00 |
| thresholds | 100.71 | 69.00 | 0.01 | 0.00 | 0.00 | 0.00 |
| residual variance | 98.88 | 24.00 | 0.00 | 0.00 | 0.00 | 0.00 |

## Lack of Empathy

**CFA Model Fit**

| Wave | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1,416.00 | 25.00 | 109.52 | 5.00 | 0.00 | 0.96 | 0.12 | [0.10, 0.14] | 0.05 |
| 2 | 1,352.00 | 25.00 | 73.91 | 5.00 | 0.00 | 0.97 | 0.10 | [0.08, 0.12] | 0.04 |
| 3 | 1,109.00 | 25.00 | 101.92 | 5.00 | 0.00 | 0.96 | 0.13 | [0.11, 0.16] | 0.05 |
| 4 | 972.00 | 25.00 | 87.71 | 5.00 | 0.00 | 0.98 | 0.13 | [0.11, 0.16] | 0.04 |

**Reliability**

| wave | alpha | omega |
| --- | --- | --- |
| 1 | 0.80 | 0.77 |
| 2 | 0.78 | 0.76 |
| 3 | 0.81 | 0.79 |
| 4 | 0.85 | 0.84 |

**Measurement Invariance Model Fit**

| model | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Configural | 1,563.00 | 136.00 | 599.91 | 134.00 | 0.00 | 0.97 | 0.05 | [0.04, 0.05] | 0.04 |
| Metric | 1,563.00 | 124.00 | 590.51 | 146.00 | 0.00 | 0.97 | 0.04 | [0.04, 0.05] | 0.04 |
| Scalar | 1,563.00 | 82.00 | 668.04 | 188.00 | 0.00 | 0.97 | 0.04 | [0.04, 0.04] | 0.04 |
| Strict | 1,563.00 | 67.00 | 752.57 | 203.00 | 0.00 | 0.96 | 0.04 | [0.04, 0.04] | 0.04 |

**Measurement Invariance Model Comparison**

| constraint\_added | delta.chisq | delta.df | delta.pvalue | delta.cfi | delta.rmsea | delta.srmr |
| --- | --- | --- | --- | --- | --- | --- |
| loadings | 15.54 | 12.00 | 0.21 | 0.00 | 0.00 | 0.00 |
| thresholds | 101.22 | 42.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| residual variance | 101.14 | 15.00 | 0.00 | 0.00 | 0.00 | 0.00 |

## Substance Use

**CFA Model Fit**

| Wave | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1,438.00 | 24.00 | 54.80 | 20.00 | 0.00 | 0.99 | 0.03 | [0.02, 0.05] | 0.05 |
| 2 | 1,065.00 | 24.00 | 106.01 | 20.00 | 0.00 | 0.98 | 0.06 | [0.05, 0.08] | 0.10 |
| 3 | 1,006.00 | 24.00 | 97.22 | 20.00 | 0.00 | 0.99 | 0.06 | [0.05, 0.07] | 0.05 |
| 4 | 967.00 | 24.00 | 103.49 | 20.00 | 0.00 | 0.99 | 0.07 | [0.05, 0.08] | 0.05 |

**Reliability**

| wave | alpha | omega |
| --- | --- | --- |
| 1 | 0.96 | 0.89 |
| 2 | 0.94 | 0.92 |
| 3 | 0.96 | 0.91 |
| 4 | 0.96 | 0.92 |

**Measurement Invariance Model Fit**

| model | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Configural | 1,562.00 | 150.00 | 597.32 | 410.00 | 0.00 | 0.99 | 0.02 | [0.01, 0.02] | 0.13 |
| Metric | 1,562.00 | 129.00 | 653.95 | 431.00 | 0.00 | 0.99 | 0.02 | [0.02, 0.02] | 0.13 |
| Scalar | 1,562.00 | 108.00 | 673.65 | 452.00 | 0.00 | 0.99 | 0.02 | [0.01, 0.02] | 0.13 |
| Strict | 1,562.00 | 84.00 | 744.86 | 476.00 | 0.00 | 0.99 | 0.02 | [0.02, 0.02] | 0.13 |

**Measurement Invariance Model Comparison**

| constraint\_added | delta.chisq | delta.df | delta.pvalue | delta.cfi | delta.rmsea | delta.srmr |
| --- | --- | --- | --- | --- | --- | --- |
| loadings | 76.64 | 21.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| thresholds | 26.85 | 21.00 | 0.18 | 0.00 | 0.00 | 0.00 |
| residual variance | 84.47 | 24.00 | 0.00 | 0.00 | 0.00 | 0.00 |

## Delinquent Behavior

**CFA Model Fit**

| Wave | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1,447.00 | 14.00 | 31.22 | 14.00 | 0.01 | 0.99 | 0.03 | [0.02, 0.04] | 0.04 |
| 2 | 1,213.00 | 14.00 | 53.71 | 14.00 | 0.00 | 0.98 | 0.05 | [0.04, 0.06] | 0.06 |
| 3 | 1,088.00 | 14.00 | 33.89 | 14.00 | 0.00 | 0.99 | 0.04 | [0.02, 0.05] | 0.04 |
| 4 | 969.00 | 14.00 | 18.97 | 14.00 | 0.17 | 1.00 | 0.02 | [0.00, 0.04] | 0.04 |

**Reliability**

| wave | alpha | omega |
| --- | --- | --- |
| 1 | 0.89 | 0.73 |
| 2 | 0.87 | 0.71 |
| 3 | 0.87 | 0.73 |
| 4 | 0.88 | 0.74 |

**Measurement Invariance Model Fit**

| model | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Configural | 1,564.00 | 107.00 | 478.04 | 299.00 | 0.00 | 0.98 | 0.02 | [0.02, 0.02] | 0.06 |
| Metric/Scalar | 1,564.00 | 89.00 | 534.03 | 317.00 | 0.00 | 0.98 | 0.02 | [0.02, 0.02] | 0.07 |
| Strict | 1,564.00 | 68.00 | 549.67 | 338.00 | 0.00 | 0.98 | 0.02 | [0.02, 0.02] | 0.07 |

**Measurement Invariance Model Comparison**

| constraint\_added | delta.chisq | delta.df | delta.pvalue | delta.cfi | delta.rmsea | delta.srmr |
| --- | --- | --- | --- | --- | --- | --- |
| loadings | 59.51 | 18.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| residual variance | 32.48 | 21.00 | 0.05 | 0.00 | 0.00 | 0.00 |

## School Belonging

**CFA Model Fit**

| Wave | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1,461.00 | 16.00 | 4.45 | 2.00 | 0.11 | 1.00 | 0.03 | [0.00, 0.07] | 0.01 |
| 2 | 1,402.00 | 16.00 | 9.16 | 2.00 | 0.01 | 0.99 | 0.05 | [0.02, 0.09] | 0.02 |
| 3 | 1,147.00 | 16.00 | 13.57 | 2.00 | 0.00 | 0.99 | 0.07 | [0.04, 0.11] | 0.03 |
| 4 | 987.00 | 16.00 | 19.44 | 2.00 | 0.00 | 0.99 | 0.09 | [0.06, 0.13] | 0.03 |

**Reliability**

| wave | alpha | omega |
| --- | --- | --- |
| 1 | 0.73 | 0.64 |
| 2 | 0.72 | 0.65 |
| 3 | 0.74 | 0.67 |
| 4 | 0.80 | 0.74 |

**Measurement Invariance Model Fit**

| model | ntotal | npar | chisq | df | pvalue | cfi | rmsea | rmsea.ci | srmr |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Configural | 763.00 | 94.00 | 110.43 | 74.00 | 0.00 | 0.99 | 0.03 | [0.01, 0.03] | 0.03 |
| Metric | 763.00 | 85.00 | 122.76 | 83.00 | 0.00 | 0.99 | 0.03 | [0.01, 0.03] | 0.03 |
| Scalar | 763.00 | 70.00 | 145.32 | 98.00 | 0.00 | 0.99 | 0.03 | [0.02, 0.03] | 0.03 |
| Strict | 763.00 | 59.00 | 199.26 | 109.00 | 0.00 | 0.99 | 0.03 | [0.03, 0.04] | 0.03 |

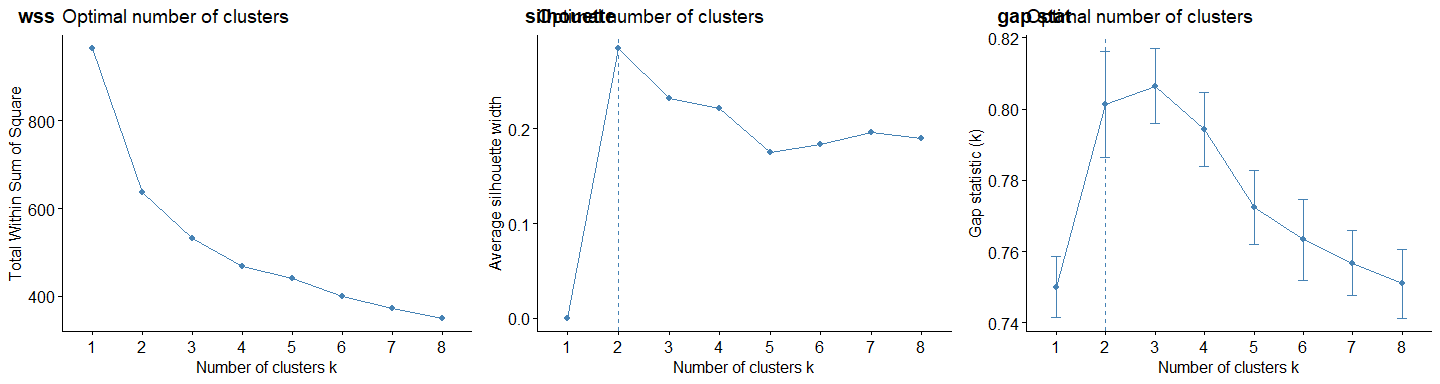
**Measurement Invariance Model Comparison**

| constraint\_added | delta.chisq | delta.df | delta.pvalue | delta.cfi | delta.rmsea | delta.srmr |
| --- | --- | --- | --- | --- | --- | --- |
| loadings | 13.746 | 9.000 | 0.132 | 0.000 | 0.008 | 0.000 |
| thresholds | 25.487 | 15.000 | 0.044 | 0.000 | 0.002 | 0.001 |
| residual variance | 53.841 | 11.000 | 0.000 | -0.004 | 0.013 | 0.005 |

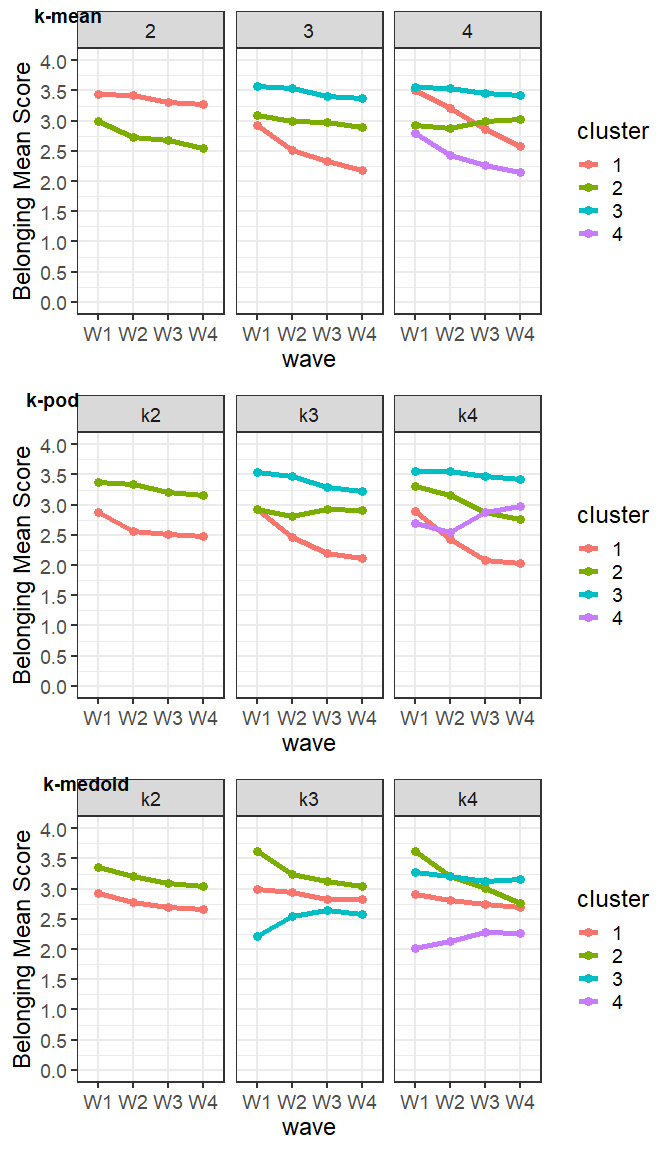
# School Belonging

Three approaches were examined to categorize students’ pattern of school belongingness over the 4 waves of data. *k*-means clustering is a common approach, but requires complete (or imputed) data. *k*-POD (Chi et al., 2016) is a recent development for conducing k-means clustering with partially observed data (i.e., missing data). *k*-medoid is an alternative clustering method that selects an exemplar data point to center each cluster rather than a mean value.

There is no definitive method for selecting the optimal *k*. Utilizing the factoextra package in R, we employed 3 methods 1) within sum of squares (wss) - similar to a scree plot, visual inspection of an “elbow” indicates the optimal *k* given the variance explained, 2) silhouette - incorporates average distance of observations to cluster center and distance between cluster centers with the optimal *k* maximizing the silhouette score, and 3) gap stat - compares the within cluster variation from the proposed clustering and a reference distribution with no clustering with the optimal *k* maximizing the gap.



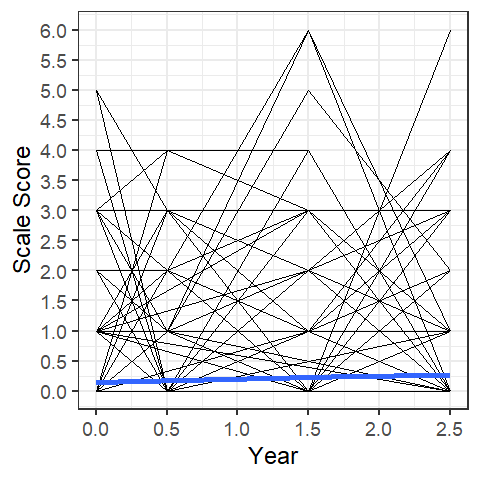
All 3 *k* selection methods suggest 2 as the optimal *k*. Running each of the 3 clustering approaches with *k* set to 2, 3, and 4 clusters produce the following cluster means at each time point.



**Conclusion** In addition to optimal cluster metrics suggesting 2 clusters, the 3 selection methods only produced similar trends with 2 clusters. Furthermore, 3 and 4 cluster patterns are highly unstable as evidenced by different patterns produced when varying the starting seed for the *k*-pod or *k*-medoid methods. Clusters identified from the *k*-pod method were added to the dataset as this method is designed explicitly to work with missing data.

Later models suggest no variation in sexual harassment perpetration slopes for students with high school belonging. All High Belonging students are included in the plot below. While there multiple lines show a variety of trends, the vast majority are along the zero line. Additionally, the plot includes a 90% confidence interval in grey around the blue trendline, but cannot be seen because it is so small, indicating minimal variation.

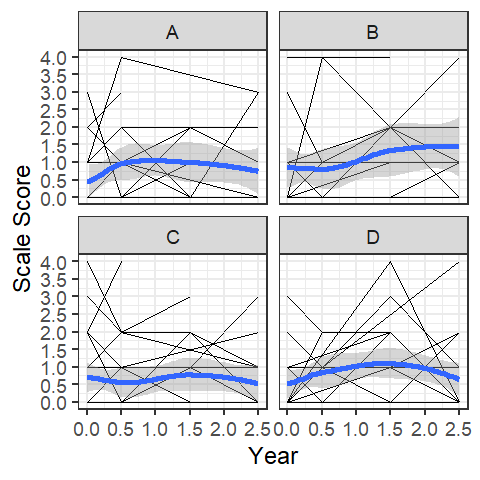
## `geom\_smooth()` using formula 'y ~ x'



# Delinquency

To explore the variability and trend in delinquency, 4 random samples of 25 students were drawn from the data to produce the spaghetti plot below. A loess trendline is highlighted in Blue.

## `geom\_smooth()` using formula 'y ~ x'



# Model Variables

## Table 1. Correlations and Descriptive Statistics

| var | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. Sexual Harassment W1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2. Sexual Harassment W2 | 0.36\* |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3. Sexual Harassment W3 | 0.29\* | 0.41\* |  |  |  |  |  |  |  |  |  |  |  |  |
| 4. Sexual Harassment W4 | 0.25\* | 0.30\* | 0.40\* |  |  |  |  |  |  |  |  |  |  |  |
| 5. Family Conflict | 0.16\* | 0.10\* | 0.13\* | 0.09\* |  |  |  |  |  |  |  |  |  |  |
| 6. Abuse | 0.23\* | 0.18\* | 0.09\* | 0.06 | 0.32\* |  |  |  |  |  |  |  |  |  |
| 7. Sibling Aggression | 0.11\* | 0.05 | 0.05 | 0.04 | 0.29\* | 0.17\* |  |  |  |  |  |  |  |  |
| 8. SH Victim | 0.34\* | 0.21\* | 0.17\* | 0.21\* | 0.27\* | 0.20\* | 0.13\* |  |  |  |  |  |  |  |
| 9. Female | -0.05\* | -0.11\* | -0.08\* | -0.07\* | 0.03 | -0.06\* | 0.00 | 0.04 |  |  |  |  |  |  |
| 10. Black | 0.05 | -0.01 | 0.04 | 0.07\* | 0.06\* | 0.08\* | -0.04 | 0.07\* | 0.02 |  |  |  |  |  |
| 11. White | -0.05 | -0.05 | -0.07\* | -0.03 | 0.05 | -0.05 | -0.01 | 0.02 | 0.00 | -0.31\* |  |  |  |  |
| 12. OtherR | 0.02 | 0.04 | 0.04 | 0.07\* | 0.04 | 0.01 | 0.02 | 0.05 | -0.01 | -0.25\* | -0.19\* |  |  |  |
| 13. High Belong | -0.13\* | -0.21\* | -0.18\* | -0.16\* | -0.13\* | -0.08\* | -0.05\* | -0.17\* | 0.08\* | -0.07\* | 0.02 | -0.04 |  |  |
| 14. GRADES W1 | -0.01 | -0.05 | -0.07\* | -0.04 | -0.02 | -0.01 | -0.02 | 0.02 | 0.06\* | -0.03 | 0.09\* | 0.07\* | 0.17\* |  |
| Descriptives |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| missing\_rate | 0.01 | 0.08 | 0.25 | 0.36 | 0.08 | 0.09 | 0.05 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.13 |
| mean | 0.23 | 0.27 | 0.35 | 0.37 | 0.78 | 0.39 | 0.34 | 0.28 | 0.49 | 0.29 | 0.19 | 0.13 | 0.60 | 4.48 |
| sd | 0.72 | 0.73 | 0.86 | 0.92 | 0.88 | 0.71 | 0.91 | 0.45 | 0.50 | 0.46 | 0.39 | 0.33 | 0.49 | 1.43 |
| skew | 4.26 | 3.52 | 3.19 | 3.40 | 1.36 | 1.80 | 3.54 | 0.96 | 0.05 | 0.90 | 1.58 | 2.22 | -0.40 | -1.22 |
| kurtosis | 21.42 | 14.45 | 11.73 | 13.38 | 1.58 | 2.41 | 13.02 | -1.09 | -2.00 | -1.18 | 0.49 | 2.94 | -1.85 | 0.97 |

## Intraclass Correlations

Our data is clustered with time nested within youth and youth nested within school. The ICC measures the % of total variation in construct scores explained by differences within youth over time compared to between youth within school or between schools. The latent growth models address the clustering of observations (i.e., time) within youth with a random intercept for youth. We account for the clustering of youth within schools by estimating cluster robust standard errors.

| Construct | School | Student | Time | Total | Design\_Effect |
| --- | --- | --- | --- | --- | --- |
| Delinquency | 0.02 | 0.40 | 0.58 | 1.00 | 2.34 |
| Depression | 0.03 | 0.43 | 0.54 | 1.00 | 3.85 |
| Lack of Empathy | 0.10 | 0.42 | 0.48 | 1.00 | 9.28 |
| School Belonging | 0.02 | 0.32 | 0.66 | 1.00 | 2.77 |
| SH Perpetration | 0.00 | 0.33 | 0.67 | 1.00 | 1.33 |

## Sexual Harassment Descriptives

### Race

| Characteristic | Black, N = 4601 | Hispanic, N = 6051 | White, N = 2981 | Other, N = 2011 | p-value2 | q-value3 |
| --- | --- | --- | --- | --- | --- | --- |
| sh\_scale\_W1\_di | 69 (15%) | 76 (13%) | 22 (7.4%) | 29 (15%) | 0.014 | 0.045 |
| sh\_scale\_W2\_di | 73 (17%) | 91 (16%) | 33 (12%) | 36 (19%) | 0.2 | 0.2 |
| sh\_scale\_W3\_di | 76 (24%) | 88 (19%) | 31 (13%) | 32 (22%) | 0.023 | 0.045 |
| sh\_scale\_W4\_di | 74 (26%) | 77 (19%) | 36 (18%) | 26 (25%) | 0.047 | 0.062 |
| 1n (%) | | | | | | |
| 2Pearson's Chi-squared test | | | | | | |
| 3False discovery rate correction for multiple testing | | | | | | |

### Gender

| Characteristic | Boy, N = 7971 | Girl, N = 7611 | p-value2 | q-value3 |
| --- | --- | --- | --- | --- |
| sh\_scale\_W1\_di | 121 (15%) | 74 (9.8%) | 0.001 | 0.002 |
| sh\_scale\_W2\_di | 150 (21%) | 82 (12%) | <0.001 | <0.001 |
| sh\_scale\_W3\_di | 134 (23%) | 92 (16%) | 0.004 | 0.005 |
| sh\_scale\_W4\_di | 121 (24%) | 92 (18%) | 0.025 | 0.025 |
| 1n (%) | | | | |
| 2Pearson's Chi-squared test | | | | |
| 3False discovery rate correction for multiple testing | | | | |

### School\_Belonging

| Characteristic | High, N = 9321 | Moderate, N = 6291 | p-value2 | q-value3 |
| --- | --- | --- | --- | --- |
| sh\_scale\_W1\_di | 82 (8.9%) | 114 (18%) | <0.001 | <0.001 |
| sh\_scale\_W2\_di | 87 (10%) | 146 (25%) | <0.001 | <0.001 |
| sh\_scale\_W3\_di | 103 (14%) | 124 (28%) | <0.001 | <0.001 |
| sh\_scale\_W4\_di | 107 (17%) | 106 (28%) | <0.001 | <0.001 |
| 1n (%) | | | | |
| 2Pearson's Chi-squared test | | | | |
| 3False discovery rate correction for multiple testing | | | | |

### Overall

| Characteristic | N = 1,5641 |
| --- | --- |
| sh\_scale\_W1\_di | 196 (13%) |
| Unknown | 15 |
| sh\_scale\_W2\_di | 233 (16%) |
| Unknown | 127 |
| sh\_scale\_W3\_di | 227 (19%) |
| Unknown | 397 |
| sh\_scale\_W4\_di | 213 (21%) |
| Unknown | 563 |
| 1n (%) | |