Supplementary Information (SI) for Chemistry Education Research and Practice.

Electronic Supplementary Information (ESI)

Title: Beyond Content Knowledge: Affective Factors Continue to Influence Performance in Organic Chemistry II

Joshua S. Henrich[a], Lorraine Laguerre Van Sickle[b], and Regina F. Frey[a]*

[a] Department of Chemistry, University of Utah, Salt Lake City, UT 84112
[b]Florida Southern College, Lakeland, FL 33801
*Corresponding Author: Regina F. Frey, gina.frey@utah.edu

Table of Contents

A. Internal Structure
   I. Social Belonging
   II. Imposter Syndrome
   III. Academic Mindset
B. SEM Path Models
   I. Main text model fits with FDR corrections and bootstrapped results
   II. Additional Models tested (not included in the main text)
C. Mediation Analysis

A: Internal Structure

Descriptive statistics and determination of sampling adequacy

The investigation of data distributions enables early detection of non-normal properties that may affect correlations and subsequent analyses. In this study, data normality and variability were assessed using descriptive statistics, including histograms, means, standard deviations, skewness, and kurtosis. Data for social belonging, imposter syndrome, and academic mindset were collected at two time points: early semester (first two weeks) and end of semester (last two weeks, prior to finals). Because the social belonging instrument had been previously used at the same institution and course, only confirmatory factor analysis (CFA) was conducted to confirm its internal structure. For academic mindset, the factor structure has been determined in prior literature as well as in the General Chemistry 1 course at this institution. However, since this self-theory survey has not been widely used at this institution and in chemistry, to give more evidence in the literature about the survey's factor structure, we also performed an EFA on our Organic Chemistry II sample. While current literature recommends performing the EFA and CFA on different subsamples (e.g., half of the observations for EFA and half for CFA) when possible, we used the same sample for our test of EFA and CFA on the self-theory data due to limited sample size. For Imposter Syndrome, both exploratory and confirmatory factor analyses (EFA and CFA) were performed and on different subsamples.

For Imposter Syndrome, the total early-semester pre_IS dataset (N = 201) was randomly split in half: one subgroup (DS1) was used for exploratory factor analysis (EFA), while the other (DS2) was reserved for confirmatory factor analysis (CFA). For this construct, separate training and testing datasets were used. Since the internal structure of the Academic Mindset (pre_AME and pr_AMI) measure had been previously confirmed in the literature and in General Chemistry I at this institution, EFA followed by CFA was conducted for this construct using the matched dataset (N = 183) to evaluate its structure in the Organic Chemistry II context. The Social Belonging (pre_SB and pre_BU) measure's internal structure has been consistently supported at this institution; therefore, only CFAs for this construct were conducted and used the matched dataset (N = 183).

Further details for each construct are described below. Descriptive statistics (not-reverse scored/coded to show true distribution) were calculated for DS1 for Imposter Syndrome and for the matched set for Social Belonging and Academic Mindset, followed by assumption checks to determine suitability for factor analysis. For multidimensional models that include both positive and negative phrasing of survey questions such as academic mindset and social belonging, it is recommended that negatively worded items (i.e., academic mindset entity and social belonging: belonging uncertainty) be reverse scored/coded to avoid additional problems with skewness (Streiner and Kottner, 2014; Komperda*, et al.*, 2018). Multivariate normality was examined using Mardia's test ($p < .05$ indicating deviation from normality). Due to violations of multivariate normality detected by Mardia's test across all measures, Polychoric correlations were used for the EFAs. The Polychoric correlation matrix was examined to ensure that a substantial number of correlations exceeded .30, indicating sufficient intercorrelation among items to justify factor analysis (Watkins, 2018). Specifically, sampling adequacy was evaluated using the Kaiser-Meyer-Olkin (KMO) measure (KMO > 0.80), and factorability was assessed with Bartlett's test of sphericity (p < .001). Multicollinearity was assessed for measures with more than one factor via Variance Inflation Factors (VIFs < 10.0). Below, we present our descriptive statistics, assumption checks, EFAs, and CFAs for each measure used in this study.

EFA and Estimation Methods

EFA is a multivariate technique for uncovering the underlying structure of relationships among a set of observed variables (items) by identifying their common underlying dimensions (Beaujean, 2014; Hair*, et al.*, 2019). This method does not rely on prior assumptions about the factor structure; factors (latent variables) are extracted from a dataset without pre-specifying the number of factors or the pattern of factor loadings between the observed and latent factors. To identify the appropriate number of factors to retain for rotation, two post-estimations were conducted: minimum average partial (MAP) (Velicer, 1976) and scree plots (Cattell, 1978). For the extraction method, Principal Axis Factoring (PAF) was chosen. An oblique rotation was selected because it allows factors to be correlated, more accurately reflecting the interrelated nature of social science variables (Meehl, 1990). Specifically, Promax (an oblique rotation method) was used to permit factor correlations (Finch, 2020). Accordingly, both the pattern coefficients (to interpret the meaning of each factor) and the factor intercorrelations (to assess the degree of overlap among factors) were examined closely (Bandalos & Gerstner, 2016; Bandalos, 2018). To ensure both practical (10% variance explained) and statistical ($p < .05$) significance of eigen loadings, the threshold for salience was originally set at 0.32 (Morin*, et al.*, 2020). Factor models were checked to ensure that the proportion (no more than 20% of data) of residual coefficients did not exceed absolute values of .05 and .10 (Flora and Flake, 2017). Close model fit, meaning the ability of a model to reproduce the data, was indicated by an RMSR value of 0.05 or smaller.

CFA Guidelines

CFA is a theory-driven statistical method used to test hypotheses about the relationships between observed variables (items) and their underlying latent constructs (factors). This method involves specifying a predefined measurement model, estimating the covariance matrix implied by this model, and comparing it to the observed covariance matrix to evaluate how well the model fits the data. Model fit is assessed using standard indices as suggested by Hu and Bentler (1999). This includes: the Comparative Fit Index (CFI), which should be $\geq$ 0.95 (Yuan, et al., 2016); the Tucker-Lewis Index (TLI), which should be $\geq$ 0.95 (Watkins, 2021); the Root Mean Square Error of Approximation (RMSEA), which should be less than 0.06 (Browne and Cudeck, 1993); and the Standardized Root Mean Square Residual (SRMR), which should be less than 0.08 (Hu and Bentler, 1999). Model fit thresholds were evaluated using criteria established by Hu and Bentler and by Yuan and Marcoulides. Model fits were then described using the qualitative descriptors proposed by Yuan and Marcoulides, which assign a range of adjectives to specific values of RMSEA (.01 = "excellent," .05 = "close," .08 = "fair," and .10 = "poor") and CFI (.99 = "excellent," .95 = "close," .92 = "fair," and .90 = "poor") (Yuan, et al., 2016). It is important to note that these cutoffs function as flexible guidelines, not strict rules, and should be considered alongside additional evidence when evaluating model fit (Goretzko et al., 2024).

Goodness-of-fit Satorra Bentler

Due to the multivariate non-normality distribution of our data, a Satorra-Bentler scaled chi-squared test for model goodness of fit was selected, which is robust to non-normality (Satorra and Bentler, 1994). The chi-square goodness of fit statistic evaluates the size of the difference between the sample and fitted covariance matrices (Hu and Bentler, 1999). For good model fits, Schreiber et al. (2006) recommended a non-significant chi-square p-value.

RMSEA

The Root Mean Square Error of Approximation (RMSEA) is an index to evaluate a "badness of fit," where zero indicates a perfect fit and higher values indicate a lack of fit (Browne and Cudeck, 1993; Watkins, 2021). Hu and Bentler (1999) recommended RMSEA values <.06 to .08. Yuan and Marcoulides (2016) assigned the following adjectives to specific values of RMSEA (.01 = "excellent," .05 = "close," .08 = "fair," and .10 = "poor").

CFI

The Comparative Fit Index (CFI) represents the amount of variance that has been accounted for in a covariance matrix (Watkins, 2021). Higher CFI values indicate a better model fit. Hu and Bentler recommend a CFI index greater than or equal to 0.95 for a good model fit; whereas, a CFI index > 0.90 is an adequate model fit (Hu and Bentler, 1999; Sivo*, et al.*, 2006). Yuan and Marcoulides (2016) assigned the following adjectives to specific values of CFI (.99 = "excellent," .95 = "close," .92 = "fair," and .90 = "poor").

TLI

The Tucker-Lewis Index (TLI) (Tucker and Lewis, 1973), also known as the non-normed fit index (NNFI) (Bentler and Bonett, 1980), is an incremental fit index that measures the proportionate improvement of fit by comparing models to a more restricted, hypothetical baseline model (Woods and Edwards, 2007). Hu and Bentler recommend a TLI index greater than or equal to 0.95 for a good model fit; whereas, a TLI index > 0.90 is an adequate model fit (Hu and Bentler, 1999; Sivo*, et al.*, 2006)

SRMR

The Standardized Root Mean Square Residual is an absolute measure of fit and is the square root of the discrepancy between the sample covariance matrix and the model covariance matrix. Hu and Bentler (1999) recommend an SRMR value of < 0.08.

References

Bandalos D. L., (2018), *Measurement theory and applications for the social sciences*, New York: Guilford Publications.

Bandalos D. L. and Gerstner J. J., (2016), Using factor analysis in test construction.

Beaujean A. A., (2014), *Latent variable modeling using R: A step-by-step guide*, New York: Routledge.

Bentler P. M. and Bonett D. G., (1980), Significance tests and goodness of fit in the analysis of covariance structures, *Psychological bulletin*, **88**, 588.

Browne M. W. and Cudeck R., (1993), Alternative ways of assessing model fit, *Sociological methods & research*, **21**, 230-258.

Cattell R., (1978), *The scientific use of factor analysis in behavioral and life sciences*, New York: Springer Science & Business Media.

Finch W. H., (2020), *Exploratory factor analysis*, Thousand Oaks, California: Sage Publications.

Flora D. B. and Flake J. K., (2017), The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation, *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, **49**, 78.

Goretzko D., Siemund K. and Sterner P., (2024), Evaluating model fit of measurement models in confirmatory factor analysis, *Educational and Psychological Measurement*, **84**, 123-144.

Hair J. F., Black W. C., Babin B. J. and Anderson R. E., (2019), Multivariate data analysis, *Journal*.

Hu L. t. and Bentler P. M., (1999), Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural equation modeling: a multidisciplinary journal*, **6**, 1-55.

Komperda R., Pentecost T. C. and Barbera J., (2018), Moving beyond alpha: A primer on alternative sources of single-administration reliability evidence for quantitative chemistry education research, *Journal of Chemical Education*, **95**, 1477-1491.

Meehl P. E., (1990), Why summaries of research on psychological theories are often uninterpretable, *Psychological Reports*, **66**, 195–244.

Morin A. J., Myers N. D. and Lee S., (2020), Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM), and bifactor-ESEM, *Handbook of sport psychology*, 1044-1073.

Satorra A. and Bentler P. M., (1994), Latent Variables Analysis: Applications for Developmental Research, **171**, 399.

Sivo S. A., Fan X., Witta E. L. and Willse J. T., (2006), The search for" optimal" cutoff properties: Fit index criteria in structural equation modeling, *The journal of experimental education*, **74**, 267-288.

Streiner D. L. and Kottner J., (2014), Recommendations for reporting the results of studies of instrument and scale development and testing, *Journal of advanced nursing*, **70**, 1970-1979.

Tucker L. R. and Lewis C., (1973), A reliability coefficient for maximum likelihood factor analysis, *Psychometrika*, **38**, 1-10.

Velicer W. F., (1976), Determining the number of components from the matrix of partial correlations, *Psychometrika*, **41**, 321-327.

Watkins M. W., (2018), Exploratory factor analysis: A guide to best practice, *Journal of black psychology*, **44**, 219-246.

Watkins M. W., (2021), *A step-by-step guide to exploratory factor analysis with SPSS*, New York, New York: Routledge.

Woods C. M. and Edwards M. C., (2007), 12 Factor analysis and related methods, *Handbook of statistics*, **27**, 367-394.

Yuan K.-H., Chan W., Marcoulides G. A. and Bentler P. M., (2016), Assessing structural equation models by equivalence testing with adjusted fit indexes, *Structural Equation Modeling: A Multidisciplinary Journal*, **23**, 319-330.

I. Social Belonging

Analyses of the Social Belonging (pre_SB and pre_BU) measure were conducted using the matched dataset (N = 183), which includes students who completed both the beginning- and end-of-semester surveys and all course assessments. Because the internal structure of this instrument has been consistently supported across multiple prior studies conducted at this institution and within similar instructional contexts, only a confirmatory factor analysis (CFA) was performed to evaluate its structure in Organic Chemistry II. No exploratory factor analysis (EFA) was used for this construct.

Prompt: Using a 6-point Likert-Scale where (1) “Strongly Disagree” (2) “Disagree” (3) “Somewhat Disagree” (4) “Somewhat Agree” (5) “Agree” (6) “Strongly Agree.” rate your agreement to the following questions:

<u>Factor 1: Social Belonging (4 items)</u>

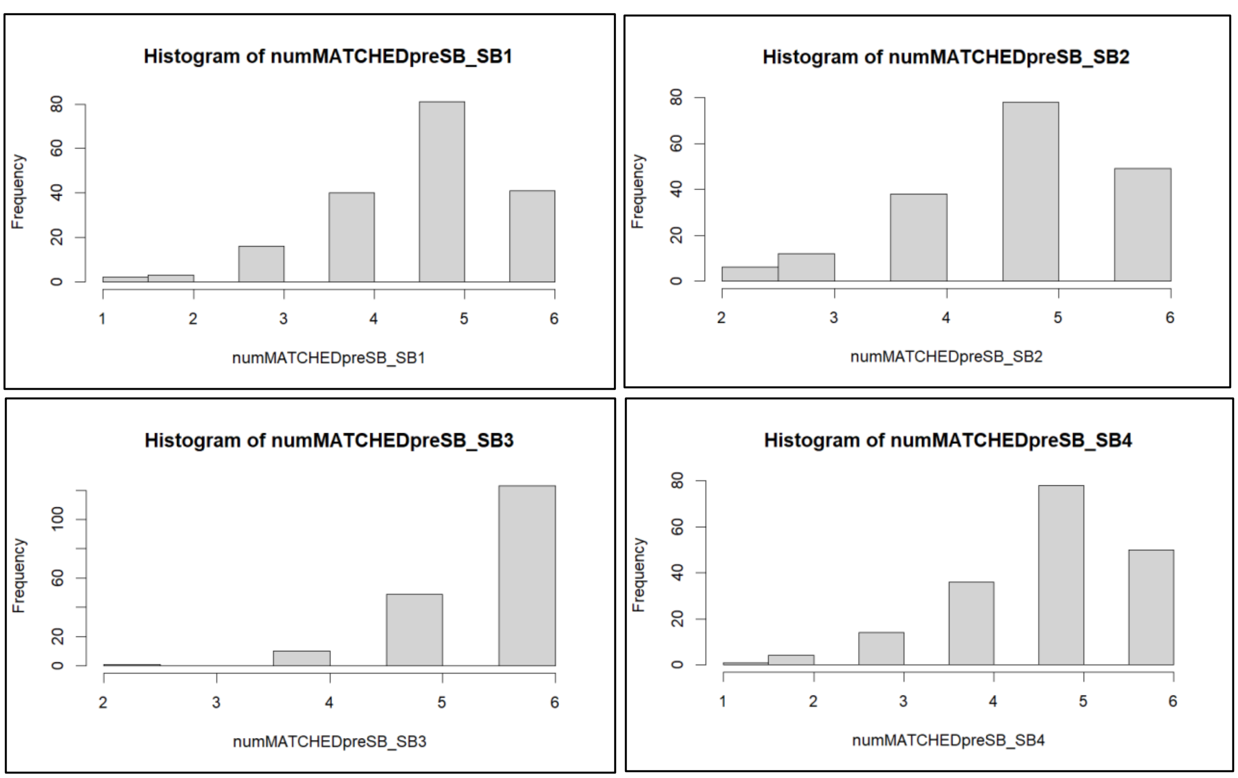
1. In (Course name), I feel like I fit in.
2. In (Course name), I feel comfortable with my peers and classmates.
3. In (Course name), I feel comfortable with my instructor.
4. In (Course name), setting aside my performance in class, I feel like belong.

<u>Factor 2: Belonging Uncertainty (2 items)</u>

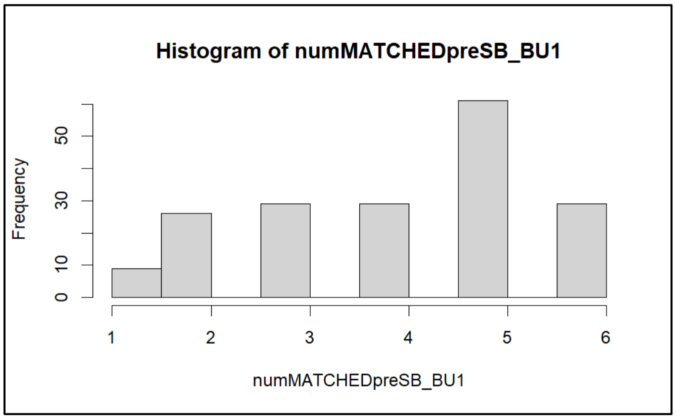
5. In (Course name), I feel uncertain about my belonging (i.e., sometimes I feel like I belong, and sometimes I don’t)
6. In (Course name), when I don’t perform well, I feel like maybe I don’t belong.

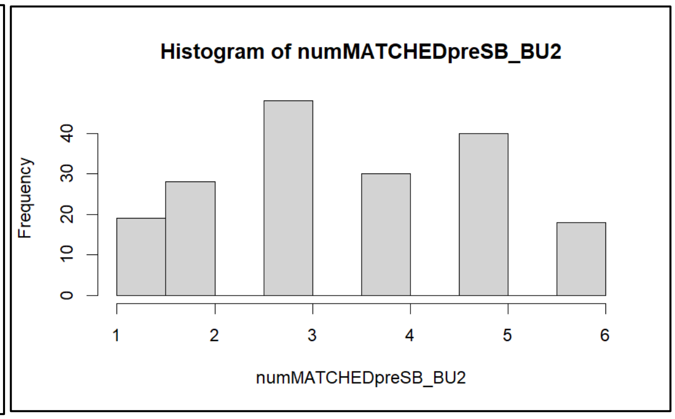
**The actual course name in the course catalog was provided to students taking the survey*

<u>Social Belonging: Beginning-of-Semester Sense of Belonging Histograms</u>



Social Belonging: Beginning-of-Semester Belonging Uncertainty Histograms



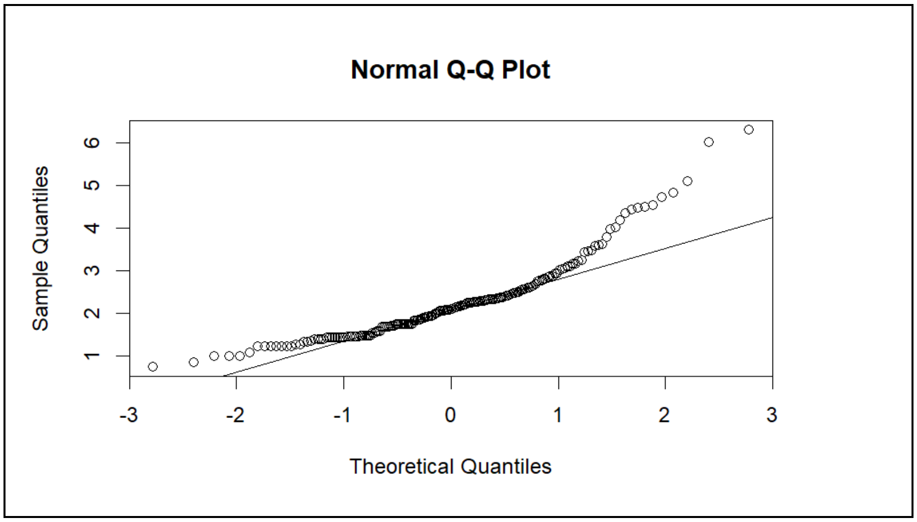


Skewness and Kurtosis for pre_SB (6 items. Belonging Uncertainty items are reverse coded)

	Skewness	Kurtosis	Mean	SD
Item SB1	-0.9301256	4.101205	4.737705	1.0255254
Item SB2	-0.8294947	3.413127	4.830601	1.0047925
Item SB3	-1.8537590	7.735807	5.601093	0.6456949
Item SB4	-0.9167750	3.807397	4.83066	1.0192703
Item BU1	-0.4515647	2.104789	4.060109	1.4644858
Item BU2	-0.03401401	2.016567	3.535519	1.4890057

Mardia's Test of Social Belonging (Belonging Uncertainty Items are Reverse Coded)

```
Mardia tests of multivariate skew and kurtosis
Use describe(x) the to get univariate tests
n.obs = 183   num.vars =  6
b1p =  11.37   skew =  346.64  with probability
 <=  1.6e-43
 small sample skew =  353.98  with probability <
=  7.3e-45
b2p =  68.43   kurtosis =  14.1  with probabilit
y <=  0
```



Mardia's test revealed significant multivariate skewness (b1p = 11.37, skew = 346.64, p < .001) and kurtosis (b2p = 68.43, kurtosis = 14.1, p < .001) in the sample of 183 cases with 6 variables, indicating that the data violate the assumption of multivariate normality. To account for the non-normality of the data, we used a polychoric correlation matrix in subsequent analyses.

Polychoric Correlation Matrix

	SB1	**SB2**	**SB3**	**SB4**	**BU1**	**BU2**
Item SB1	1.0000000					
Item SB2	0.7991296	1.0000000				
Item SB3	0.5630156	0.5134998	1.0000000			
Item SB4	0.7681316	0.8042976	0.4844906	1.0000000		
Item BU1	0.6361503	0.6603969	0.3392654	0.6471144	1.0000000	
Item BU2	0.4810976	0.4392290	0.303418	0.3984689	0.597442	1.0000000

Kaiser-Myer Olkin Test (KMO)

```
> kmo_MATCHEDpreSocBelong[["results"]]
$overall
[1] 0.861353

$individual
                                  MSA
numMATCHEDpreSB_SB1 0.8708961
numMATCHEDpreSB_SB2 0.8550396
numMATCHEDpreSB_SB3 0.9092360
numMATCHEDpreSB_SB4 0.8634604
numMATCHEDpreSB_BU1 0.8537580
numMATCHEDpreSB_BU2 0.8205908
```

Bartlett's Test of Sphericity

```
> BartRes <- cortest.bartlett(MATCHEDAllpreSocBelongEX
PolMatrix, n=183)
> BartRes
$chisq
[1] 672.2724

$p.value
[1] 1.501664e-133

$df
[1] 15

> format(BartRes$p.value, scientific = TRUE)
[1] "1.501664e-133"
```

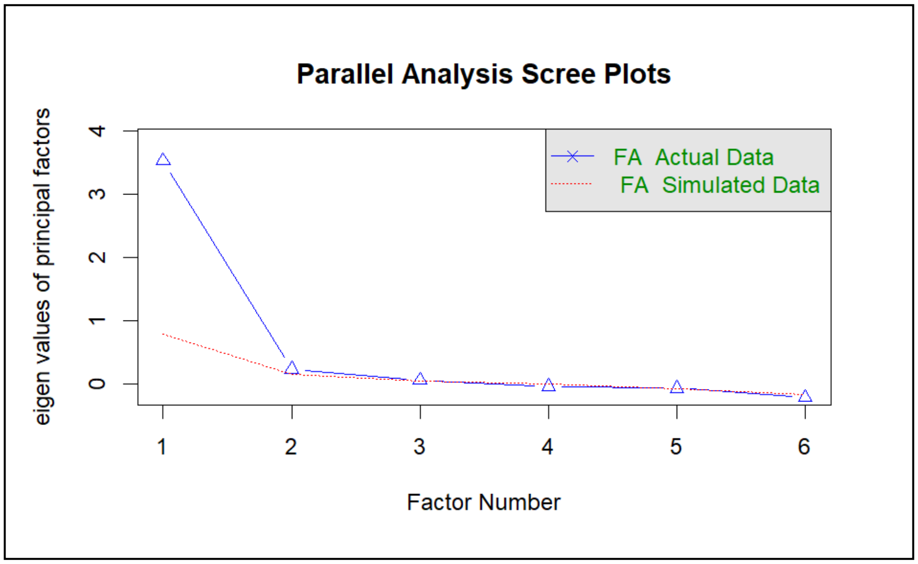
Check for Multicollinearity Using VIFs

```
> vif(sense_model)
numMATCHEDpreSB_SB1 numMATCHEDpreSB_SB2 numMATCHEDpreSB_SB3 numMATCHEDpreSB_SB4
       3.396500        3.695306        1.490932        3.235336

> vif(Uncertainty_model)
numMATCHEDpreSB_BU1 numMATCHEDpreSB_BU2
       1.555057        1.555057
```

Minimum Average Partial (MAP) Test

```
Very Simple Structure
Call: vss(x = MATCHEDAllpreSocBelongEXPolMatrix, rotate = "none", n.obs =
183,
    plot = FALSE)
VSS complexity 1 achieves a maximimum of 0.92  with  5  factors
VSS complexity 2 achieves a maximimum of 0.95  with  4  factors

The Velicer MAP achieves a minimum of 0.07  with  1  factors
BIC achieves a minimum of  -9.28  with  2  factors
Sample Size adjusted BIC achieves a minimum of  3.39  with  2  factors
```

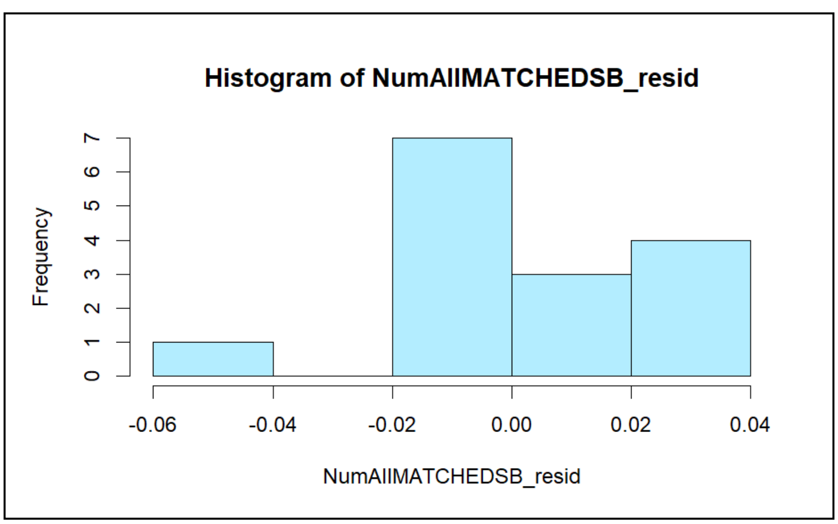
Scree Plot



Parallel analysis suggests that the number of factors = 2

Residuals for pre_SB:

	SB1	SB2	SB3	SB4	BU1	BU2
Item SB1	1.0000000					
Item SB2	-0.15815135	1.0000000				
Item SB3	0.025902068	-0.019806942	1.0000000			
Item SB4	-0.015373236	-0.008278707	0.015695952	1.0000000		
Item BU1	-0.010085539	0.005166158	0.026846362	0.0377314	1.0000000	
Item BU2	0.033838864	-0.016341073	-0.0184617	-0.0412921	0.007464214	1.0000000

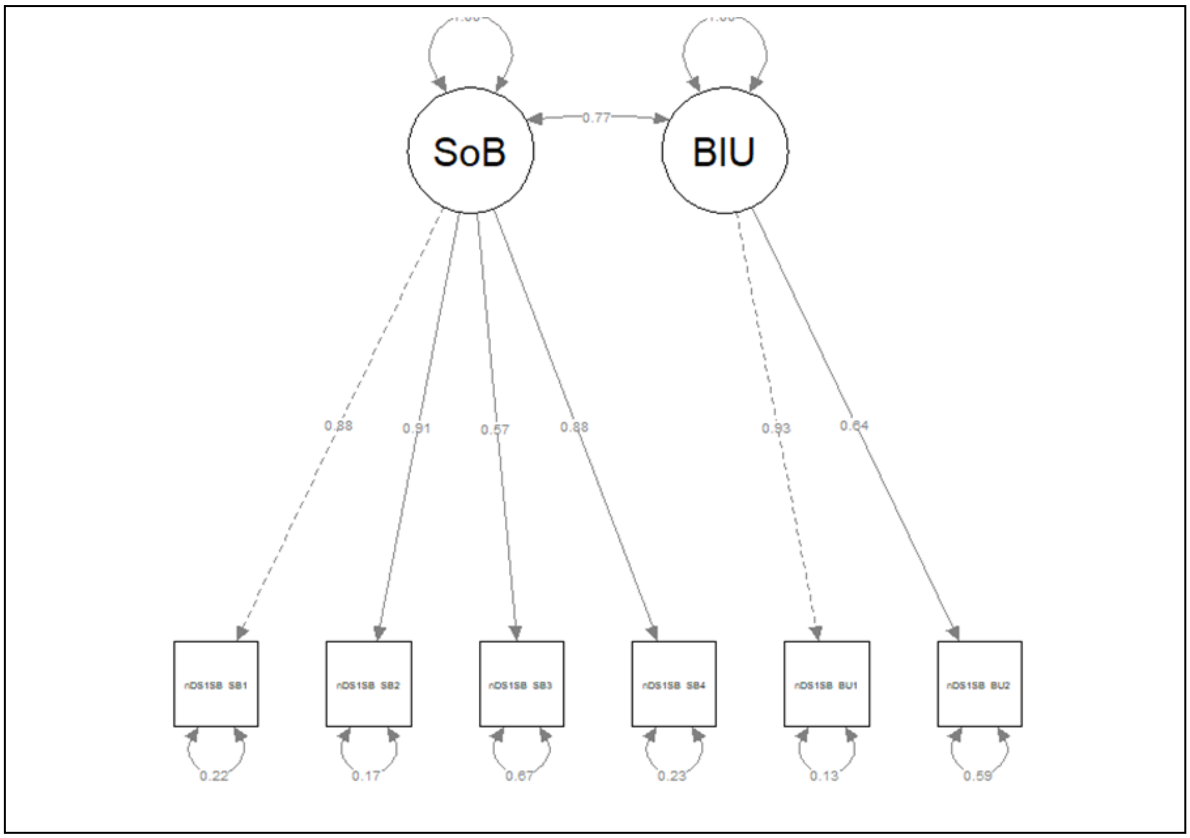


- No residual is over |0.05|
- Residuals are bimodal (not normally distributed)

Note: EFA was not conducted on SB as this instrument's internal structure has been confirmed in previous studies at the same institution (Edwards, *et al.*, 2021; Edwards, *et al.*, 2022; Edwards, *et al.*, 2023; Bustamante & Frey, 2025).

Multidimensional Confirmatory Factor Analysis of pre_ Social Belonging (pre_SB and pre_BU) Construct

Estimator: MLM



N	χ^2 (*df*, p)	RMSEA	CFI	TLI	SRMR
183	9.360 (8, 0.313)	0.038	0.997	0.994	0.025

This model demonstrates an *excellent* fit.

<u>Unidimensional Confirmatory Factor Analysis of pre_Belonging Uncertainty Data</u>

2 items on one factor

Estimator: MLM

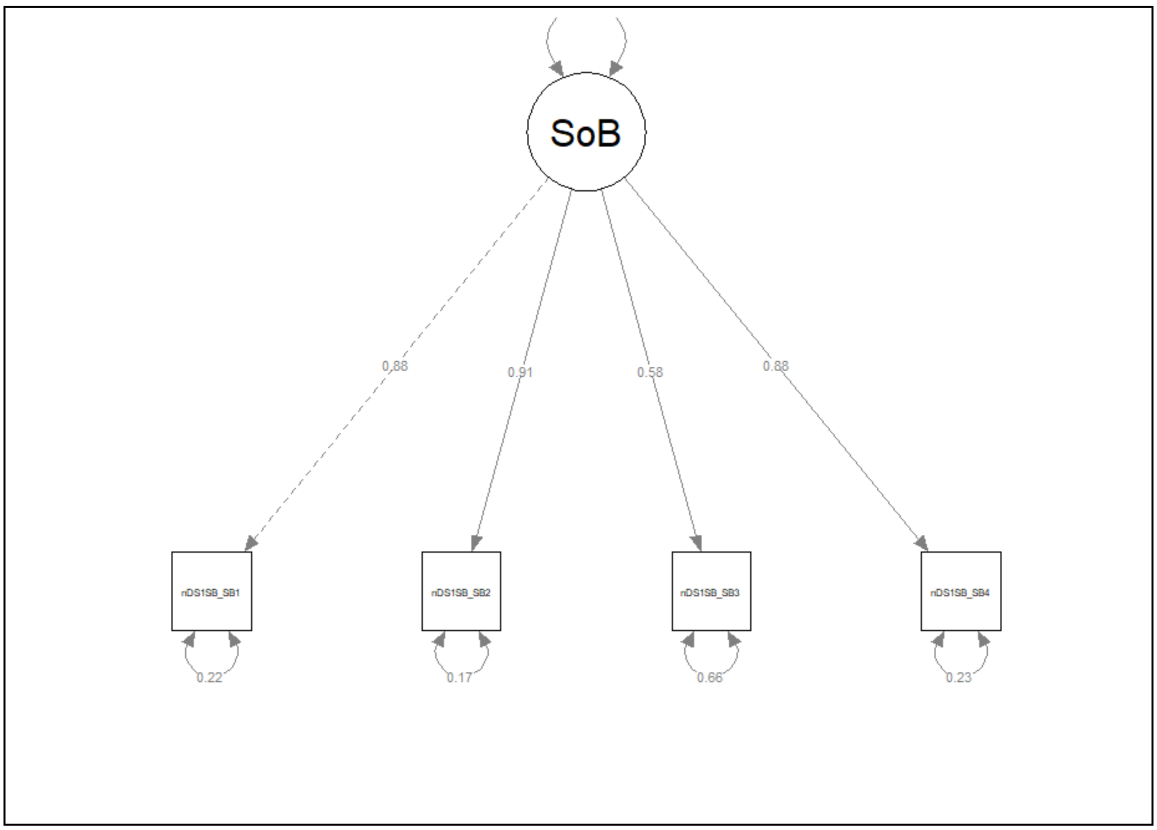
```
Warning messages: 1: lavaan->ldw_parse_model_string(): having identifiers
with spaces ('Belonging Uncertainty') is deprecated at line 1, pos 1
Belonging Uncertainty =~ numMATCHEDpreSB_BU1 + numMATCHEDpreSB_BU2 ^ 2:
lavaan->ldw_parse_model_string(): having identifiers with spaces ('Belonging
Uncertainty') is deprecated at line 1, pos 1 Belonging Uncertainty =~
numMATCHEDpreSB_BU1 + numMATCHEDpreSB_BU2 ^ 3: lavaan->lav_model_vcov():
Could not compute standard errors! The information matrix could not be
inverted. This may be a symptom that the model is not identified.
```

Note: Belonging uncertainty unidimensional model does not have an adequate internal structure to support 2 items on one factor. Hence, we did not use it in our SEM path analysis.

<u>Unidimensional Confirmatory Factor Analysis of pr_Sense of Belonging Data</u>

4 items on one factor

Estimator: MLM



N	χ^2 (*df*, p)	RMSEA	CFI	TLI	SRMR
183	1.898 (2, 0.387)	0.0*	1.0*	1.0*	0.018

**To further support these fit indices, we provide additional "checkpoints" to support the use of this model. Please see modification indices, as well as residuals.*

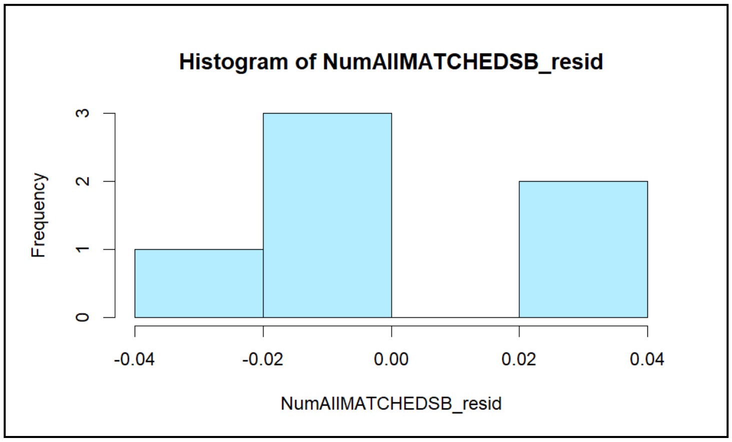
Modification Indices

```
> mi
                 lhs op                rhs    mi    epc sepc.lv sepc.all
11 numMATCHEDpreSB_SB1 ~~ numMATCHEDpreSB_SB3 4.590  0.051   0.051    0.202
14 numMATCHEDpreSB_SB2 ~~ numMATCHEDpreSB_SB4 4.590  0.121   0.121    0.589
15 numMATCHEDpreSB_SB3 ~~ numMATCHEDpreSB_SB4 1.306 -0.027  -0.027   -0.106
10 numMATCHEDpreSB_SB1 ~~ numMATCHEDpreSB_SB2 1.306 -0.067  -0.067   -0.334
13 numMATCHEDpreSB_SB2 ~~ numMATCHEDpreSB_SB3 0.947 -0.022  -0.022   -0.101
12 numMATCHEDpreSB_SB1 ~~ numMATCHEDpreSB_SB4 0.947 -0.052  -0.052   -0.221
```

Note: Modification indices are not large enough to warrant kicking out any items, and expected parameter change (EPC) are not large enough to impact assessment (Gunzler and Morris, 2015).

Residuals:

	SB1	**SB2**	**SB3**	**SB4**
Item SB1	1			
Item SB2	-0.01348160	1		
Item SB3	0.03577154	-0.01015437	1	
Item SB4	-0.1497690	0.02419212	-0.02166290	1



- No residual is over |0.05|
- Residuals are bimodal (not normally distributed)

Note: Residual correlations are all very small (none exceed |0.05|), so there is no evidence of substantial localized misfit among these four indicators despite the bimodal residual distribution. Given this, the model can be treated as a reasonable approximation to the data.

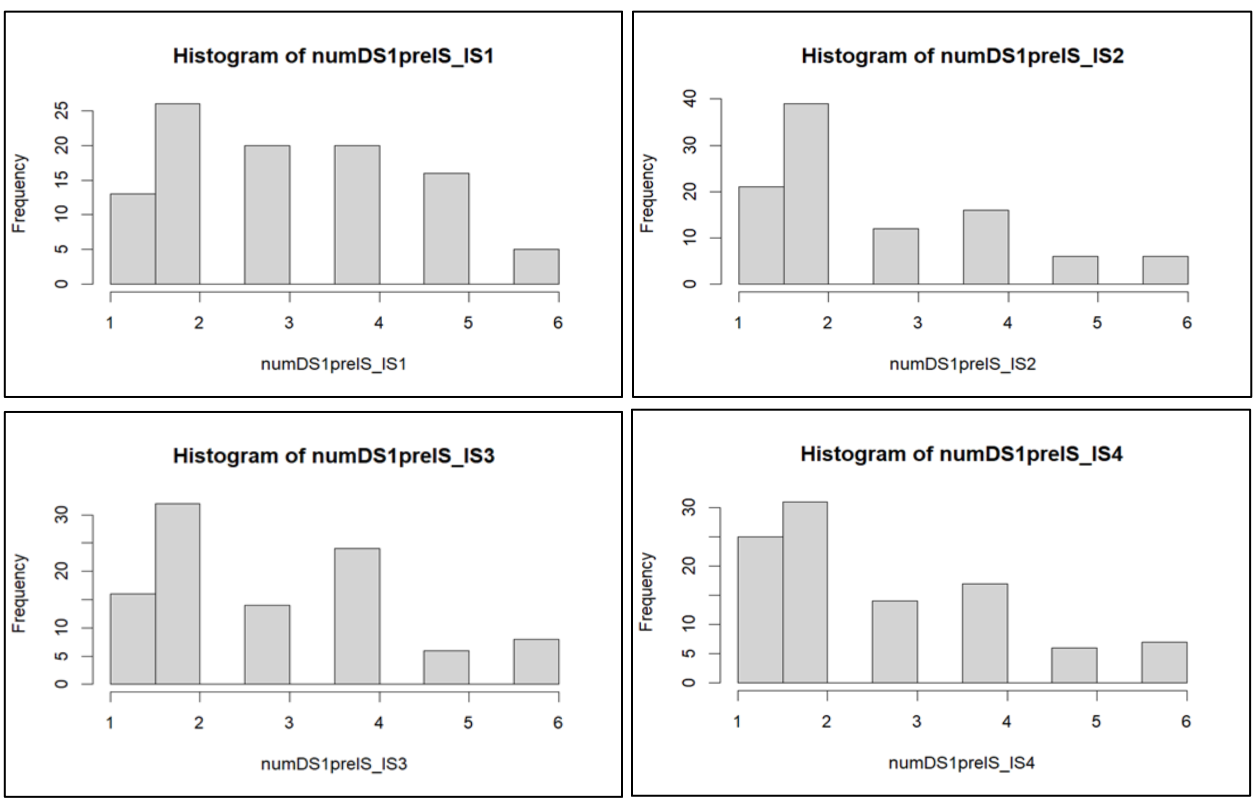
II. Imposter Syndrome

Analyses of the Impostor Syndrome (pre_IS) measure were conducted using the total early-semester dataset (N = 201), which includes all students who completed the beginning-of-semester survey and relevant course assessments, regardless of end-of-semester survey completion. While the internal structure of this measure has previously been evaluated in General Chemistry 1 at this institution, it has not been widely evaluated elsewhere especially in STEM; therefore, a split-sample approach was used. The dataset was randomly divided into a training subset (DS1; N = 100) for the EFA and a testing subset (DS2; N = 101) for the CFA, providing independent evidence for the stability of the factor structure.

<u>Prompt:</u> Using a 6-point Likert-Scale where (1) “Strongly Disagree” (2) “Disagree” (3) “Somewhat Disagree” (4) “Somewhat Agree” (5) “Agree” (6) “Strongly Agree.” rate your agreement to the following questions.

1. In class, I felt like people might find out that I am not as capable as they think I am.
2. Today, I felt like my successes in class were due to some kind of luck.
3. In class, I felt afraid others would discover how much knowledge or ability I really lack.
4. In class, I felt like an "imposter."

Beginning-of-Semester Imposter Syndrome Histograms

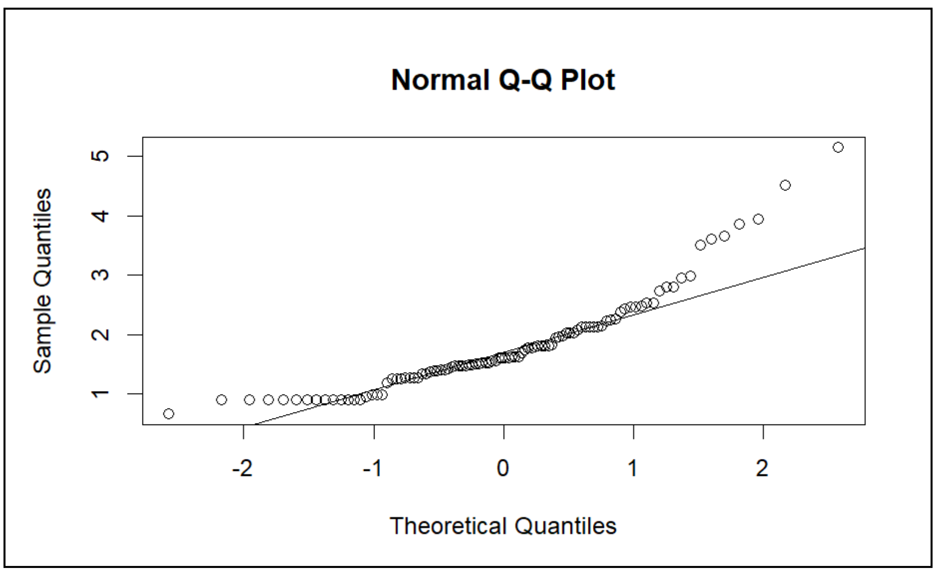


DS1-Skewness and Kurtosis for pre_IS (4 items)

	Skewness	Kurtosis	Mean	SD
Item IS1	0.2065851	2.029917	3.23	1.441485
Item IS2	0.8132847	2.754109	2.66	1.464875
Item IS3	0.4853567	2.315142	3.20	1.556998
Item IS4	0.6962388	2.504398	2.90	1.623688

Mardia's Test

```
Call: mardia(x = DS1_AllpreIS)

Mardia tests of multivariate skew and kurtosis
Use describe(x) the to get univariate tests
n.obs = 100   num.vars =  4
b1p =  2.77   skew =  46.16  with probability  <=  0.00077
 small sample skew =  48.11  with probability <=  0.00041
b2p =  32.87   kurtosis =  6.4  with probability <=  1.6e-10
```



Note: Mardia's test indicated significant multivariate skewness (b1p = 2.77, skew = 46.16, $p < .001$) and kurtosis (b2p = 32.87, kurtosis = 6.4, $p < .001$) for the set of 4 variables ($n = 100$). These results suggest the data deviate significantly from multivariate normality. To account for the non-normality of the data, we used a polychoric correlation matrix in subsequent analyses.

Polychoric Correlation Matrix

	IS1	IS2	IS3	IS4
Item IS1	1			
Item IS2	0.4870654	1		
Item IS3	0.7353917	0.6147035	1	
Item IS4	0.6745451	0.6353210	0.7871189	1

Kaiser-Myer Olkin Test (KMO)

```
> kmo_DS1preIS[["results"]]
$overall
[1] 0.8126946

$individual
                       MSA
numDS1preIS_IS1 0.8361959
numDS1preIS_IS2 0.8783846
numDS1preIS_IS3 0.7676486
numDS1preIS_IS4 0.8015372
```

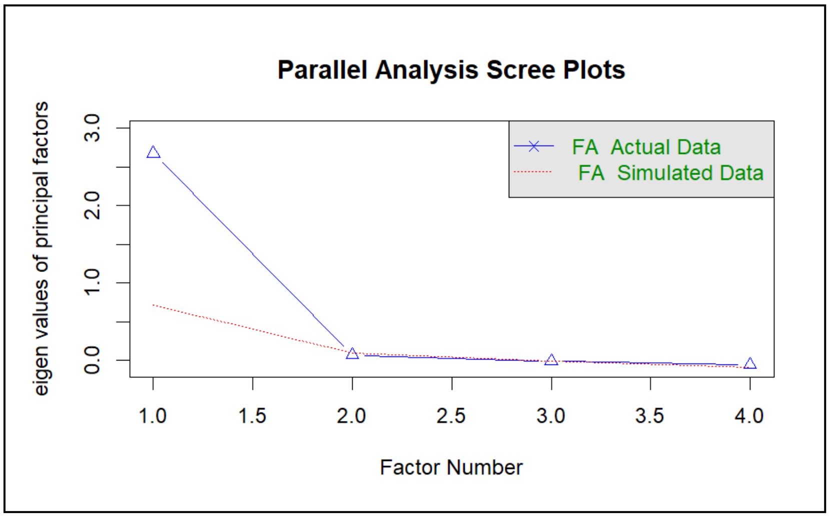
Barlett’s Test of Sphericity

```
> BartRes <- cortest.bartlett(DS1AllpreISEXPolMatrix, n=100)
> BartRes
$chisq
[1] 229.9927

$p.value
[1] 7.6844e-47

$df
[1] 6

> format(BartRes$p.value, scientific = TRUE)
[1] "7.6844e-47"
```

Minimum Average Partial (MAP) Test

```
Very Simple Structure
Call: vss(x = DS1AllpreISEXPolMatrix, rotate = "none", n.obs = 100,
    plot = FALSE)
VSS complexity 1 achieves a maximimum of 0.95  with  2  factors
VSS complexity 2 achieves a maximimum of 0.96  with  2  factors

The Velicer MAP achieves a minimum of 0.12  with  1  factors
BIC achieves a minimum of  -5.59  with  1  factors
Sample Size adjusted BIC achieves a minimum of  0.72  with  1  factors

Statistics by number of factors
  vss1 vss2  map dof   chisq prob sqresid  fit RMSEA  BIC SABIC complex
1 0.94 0.00 0.12   2 3.6e+00 0.16    0.54 0.94 0.089 -5.6  0.72     1.0
2 0.95 0.96 0.42  -1 9.6e-07   NA    0.37 0.96    NA   NA    NA     1.1
3 0.95 0.96 1.00  -3 0.0e+00   NA    0.37 0.96    NA   NA    NA     1.1
4 0.95 0.96   NA  -4 0.0e+00   NA    0.37 0.96    NA   NA    NA     1.1
```

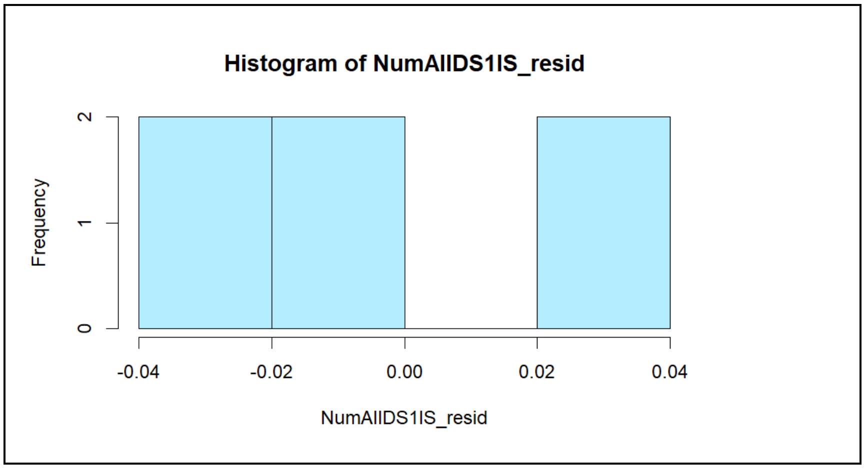
Scree Plot



Parallel analysis suggests that the number of factors = 1

Residuals for pre_IS:

	IS1	IS2	IS3	IS4
Item IS1	1			
Item IS2	-0.035127637	1		
Item IS3	0.031095940	-0.005923893	1	
Item IS4	-0.005579803	0.036022195	-0.021171500	1



- No residual is over |0.05|
- Residuals are bimodal (not normally distributed)

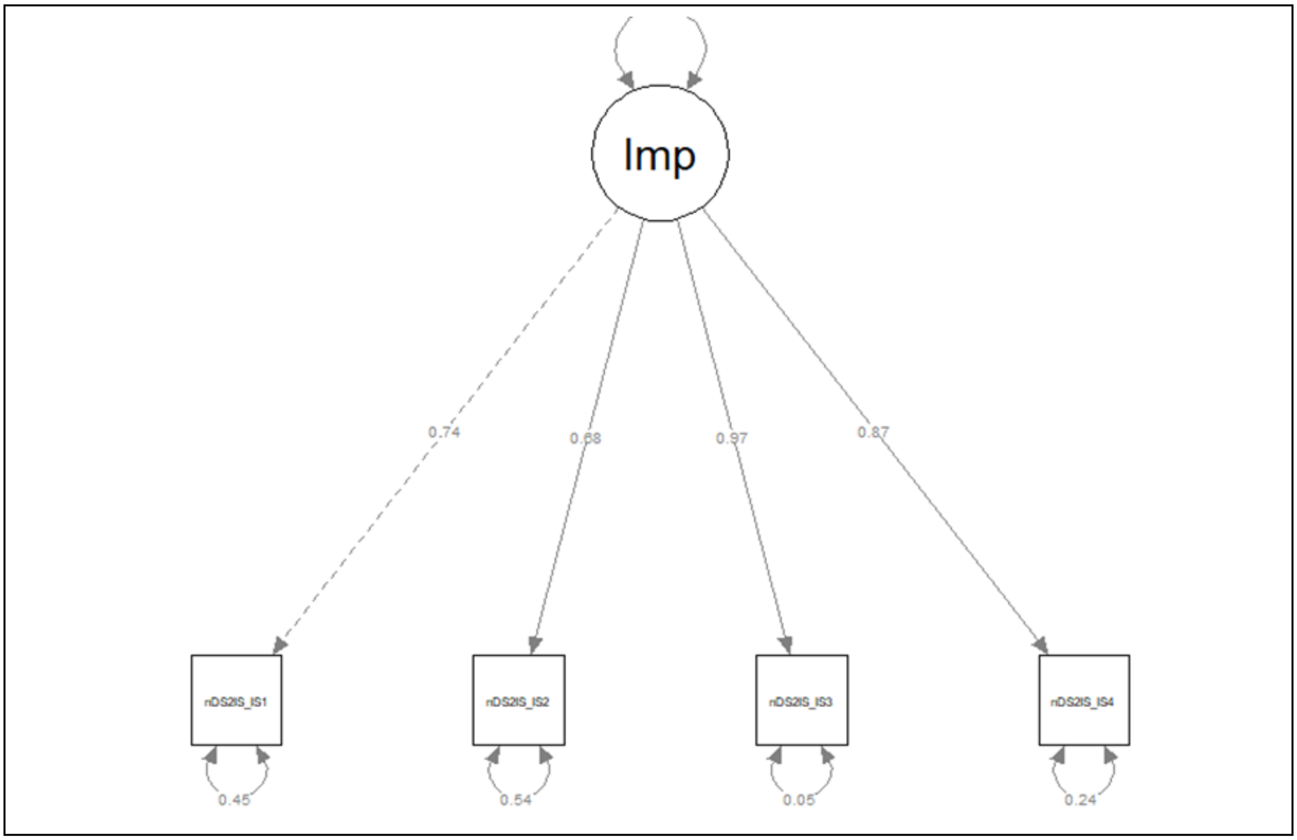
Note: Residual correlations are all very small (none exceed |0.05|), so there is no evidence of substantial localized misfit among these four indicators despite the bimodal residual distribution. Given this, the model can be treated as a reasonable approximation to the data.

Exploratory Factor Analysis

Factor Loadings of Exploratory Factor Analysis (EFA) of Early-Semester Imposter Syndrome			
Survey Item	**Factor 1**	**Uniqueness** u^2	**Communality** h^2
DS1-item IS1	**0.77**	0.41	0.59
DS1-item IS2	**0.68**	0.54	0.46
DS1-item IS3	**0.91**	0.16	0.84
DS1-item IS4	**0.88**	0.22	0.78

Confirmatory Factor Analysis on DS2 Data

Estimator: MLM



N	χ^2 (*df*, p)	RMSEA	CFI	TLI	SRMR
101	2.77 (2, 0.249)	0.086	0.994	0.983	0.023

This model demonstrates a *close* fit.

III. Academic Mindset

Analyses of the Academic Mindset (pre_AME and pre_AMI) measure were conducted using the matched dataset (N = 183). While, the factor structure has been determined in prior literature as well as in the General Chemistry 1 course at this institution,, the EFA and CFA were conducted to evaluate its internal structure specifically within the Organic Chemistry II context.

Prompt: Using a 6-point Likert-Scale where (1) "Strongly Disagree" (2) "Disagree" (3) "Somewhat Disagree" (4) "Somewhat Agree" (5) "Agree" (6) "Strongly Agree." rate your agreement to the following questions.

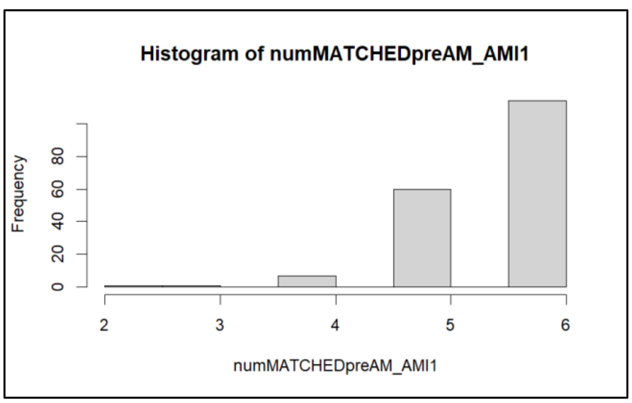
Self-Theory Incremental: (**4 questions total**)

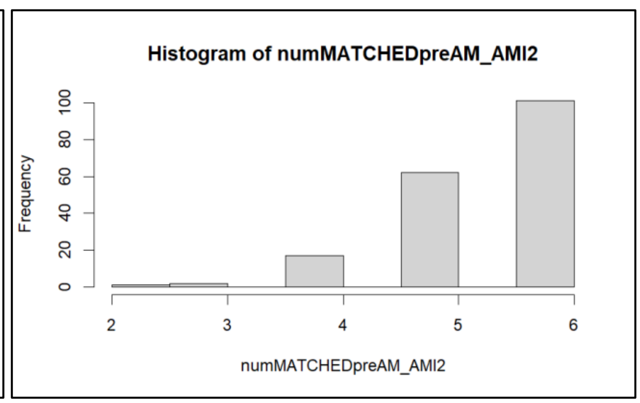
1. With enough time and effort, I think I could significantly improve my intelligence level in chemistry.
2. I believe I can always substantially improve my intelligence in chemistry.
3. Regardless of my current chemistry intelligence level, I think I have the capacity to change it quite a bit.
4. I believe I have the ability to change my basic intelligence level in chemistry considerably over time.

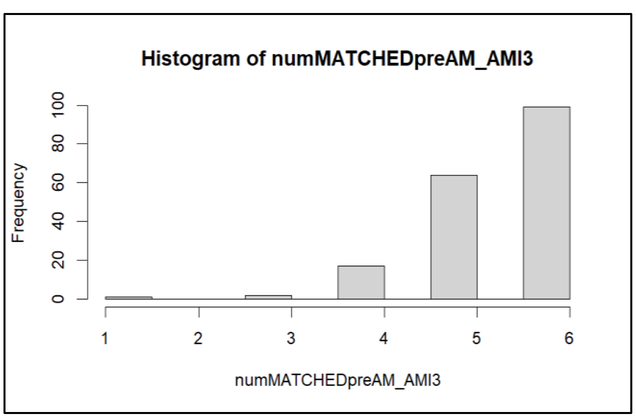
Self-Theory Entity: (**4 questions total**)

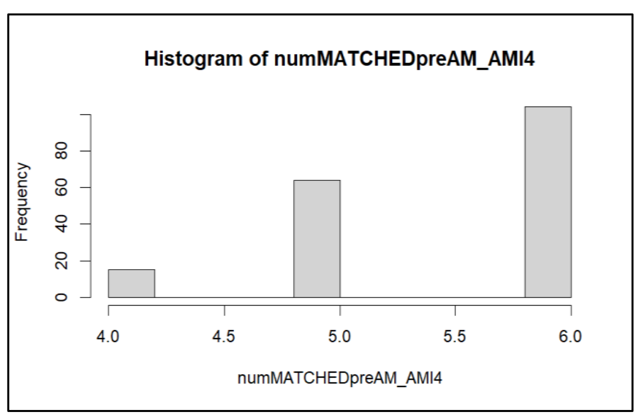
1. I don't think I personally can do much to increase my chemistry intelligence.
2. My chemistry intelligence is something about me that I personally can't change very much.
3. To be honest, I don't think I can really change how intelligent I am in chemistry.
4. I can learn new things, but I don't have the ability to change my basic intelligence in chemistry.

Academic Mindset: Beginning-of-Semester Incremental Histograms

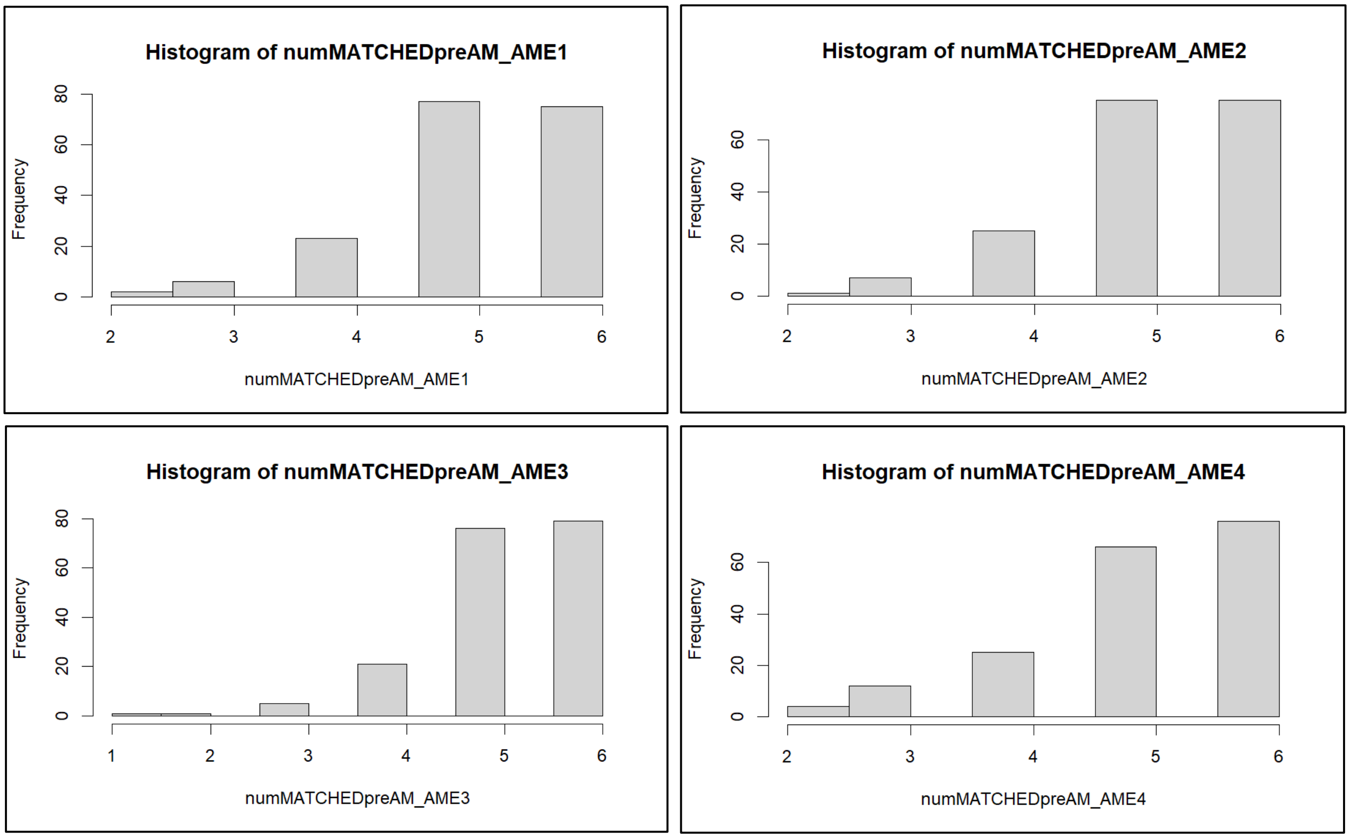








Academic Mindset: Beginning-of-Semester Entity Histograms

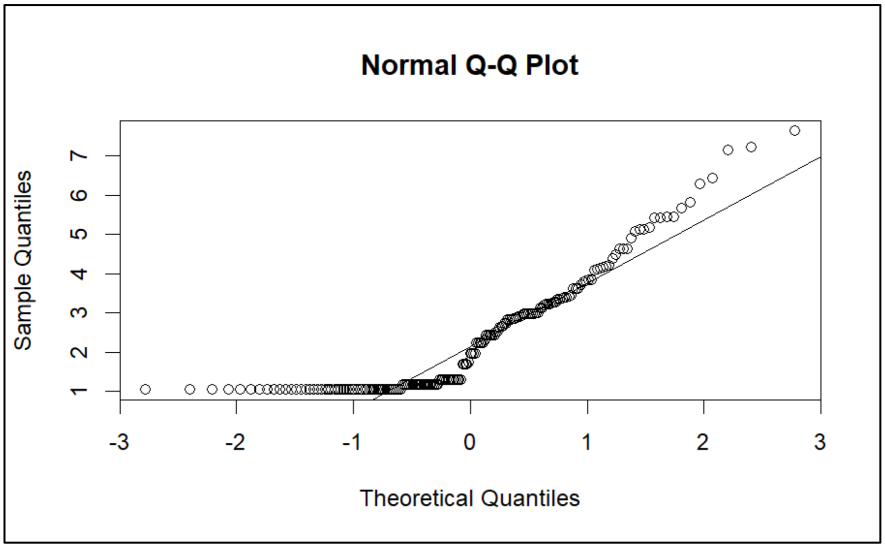


Skewness, Kurtosis, SD, and Means for pre_Academic Mindset (8 items). Entity items are reverse coded.

	Skewness	Kurtosis	Mean	SD
Item I1	-1.770567	7.818007	5.557377	0.6510665
Item I2	-1.329583	5.009891	5.420765	0.7508277
Item I3	-1.682464	7.905487	5.404372	0.7776991
Item I4	-0.8730783	2.682300	5.486339	0.6449970
Item E1	-1.100009	4.316145	5.185792	0.8570428
Item E2	-0.9465277	3.654405	5.180328	0.8485579
Item E3	-1.421714	6.197799	5.224044	0.8638472
Item E4	-1.076746	3.655475	5.081967	1.0048522

Mardia’s Test of Academic Mindset (Incremental and Entity; Entity Items are Reverse Coded)

```
Call: mardia(x = MATCHED_AllpreAM)

Mardia tests of multivariate skew and kurtosis
Use describe(x) the to get univariate tests
n.obs = 183   num.vars =  8
b1p =  41.73   skew =  1272.92  with probability  <=  8.2e-192
small sample skew =  1298.49  with probability <=  7.4e-197
b2p =  167.29   kurtosis =  46.68  with probability <=  0
```



Mardia’s test of multivariate skewness (b1p = 41.73, $p < .001$) and kurtosis (b2p = 167.29, $p < .001$; kurtosis = 46.68) indicated significant departures from multivariate normality for the 8 variables in the sample (n = 183). These results suggest the data do not meet the assumption of multivariate normality. Therefore, we used a polychoric correlation matrix in subsequent analyses.

Polychoric Correlation Matrix

	AME1	AME2	AME3	AME4	AMI1	AMI2	AMI3	AMI4
Item AME1	1.0000000							
Item AME2	0.8905201	1.0000000						
Item AME3	0.8859923	0.8965308	1.0000000					
Item AME4	0.7669654	0.7751626	0.8142604	1.0000000				
Item AMI1	0.5519122	0.5530289	0.5289871	0.5260757	1.0000000			
Item AMI2	0.5609336	0.5615472	0.4976788	0.5220784	0.8439148	1.0000000		
Item AMI3	0.59560810	0.6299084	0.5677661	0.5268612	0.7243852	0.7797213	1.0000000	
Item AMI4	0.5314150	0.5416110	0.4837963	0.5315821	0.6593554	0.7323855	0.8106906	1.0000000

Kaiser-Myer Olkin Test (KMO)

```
> kmo_MATCHEDpreAM[["results"]]
$overall
[1] 0.8835621

$individual
                   MSA
numMATCHEDpreAM_AME1 0.9109350
numMATCHEDpreAM_AME2 0.8998332
numMATCHEDpreAM_AME3 0.8589596
numMATCHEDpreAM_AME4 0.9261286
numMATCHEDpreAM_AMI1 0.8695784
numMATCHEDpreAM_AMI2 0.8496850
numMATCHEDpreAM_AMI3 0.8786280
numMATCHEDpreAM_AMI4 0.8759294
```

Barlett's Test of Sphericity

```
> BartRes
$chisq
[1] 1535.599

$p.value
[1] 1.863155e-306

$df
[1] 28

> format(BartRes$p.value, scientific = TRUE)
[1] "1.863155e-306"
```

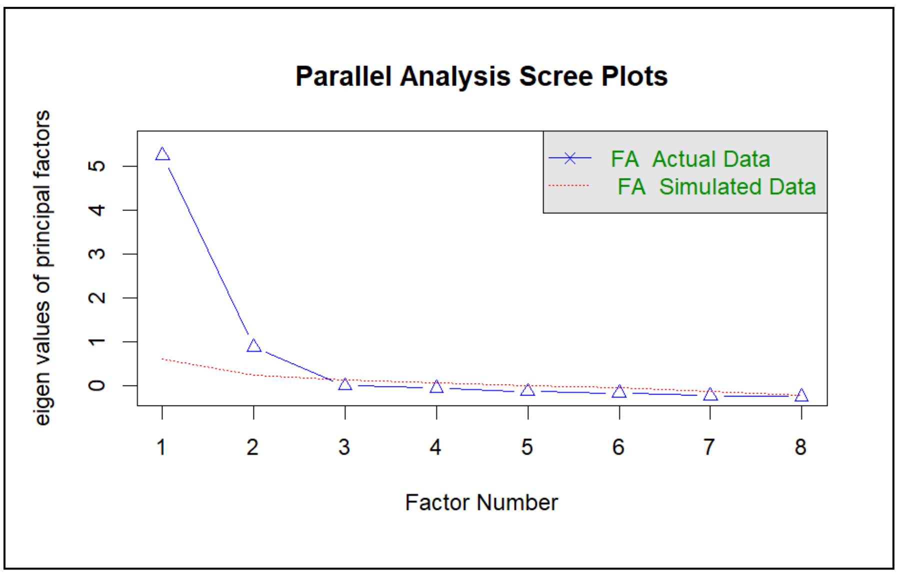
Check for Multicollinearity Using VIFs

```
MATCHEDpreAM_AMI1 MATCHEDpreAM_AMI2 MATCHEDpreAM_AMI3 MATCHEDpreAM_AMI
4
      3.615708        4.586640        3.830895        3.153265

MATCHEDpreAM_AME1 MATCHEDpreAM_AME2 MATCHEDpreAM_AME3 MATCHEDpreAM_A
ME4
      6.017583        6.608701        7.204670        3.092836
```

Minimum Average Partial (MAP) Test

```
Very Simple Structure
Call: vss(x = MATCHED_AllpreAM, rotate = "varimax", fm = "minres", n.obs = nrow(DS1_
AllpreAM))
VSS complexity 1 achieves a maximimum of 0.94  with  1  factors
VSS complexity 2 achieves a maximimum of 0.99  with  2  factors

The Velicer MAP achieves a minimum of 0.07  with  2  factors
BIC achieves a minimum of  -15.99  with  3  factors
Sample Size adjusted BIC achieves a minimum of  2.7  with  4  factors

Statistics by number of factors
  vss1 vss2   map dof   chisq    prob sqresid  fit RMSEA   BIC SABIC
1 0.94 0.00 0.217  20 4.8e+02 3.1e-89    1.93 0.94  0.35 376.5 439.9
2 0.68 0.99 0.072  13 6.6e+01 4.9e-09    0.41 0.99  0.15  -2.0  39.1
3 0.57 0.88 0.108   7 2.0e+01 4.6e-03    0.24 0.99  0.10 -16.0   6.2
4 0.56 0.84 0.175   2 6.8e+00 3.4e-02    0.14 1.00  0.11  -3.6   2.7
5 0.51 0.86 0.321  -2 5.3e-02      NA    0.16 1.00    NA    NA    NA
6 0.54 0.82 0.524  -5 4.4e-10      NA    0.15 1.00    NA    NA    NA
7 0.54 0.82 1.000  -7 9.1e-11      NA    0.14 1.00    NA    NA    NA
8 0.54 0.82    NA  -8 9.1e-11      NA    0.14 1.00    NA    NA    NA
```

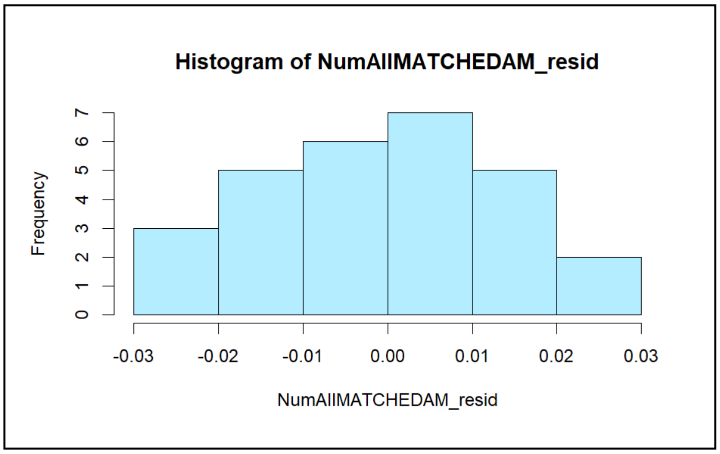
Scree Plot



Parallel analysis suggests that the number of factors = 2.

Residuals for pre_AM:

	AME1	AME2	AME3	AME4	AMI1	AMI2	AMI3	AMI4
Item AME1	1.0000000							
Item AME2	0.015737006	1.0000000						
Item AME3	-0.001964518	0.003167541	1.0000000					
Item AME4	-0.010996479	-0.003401225	-0.001195644	1.0000000				
Item AMI1	-0.020564683	-0.025528969	0.004710819	0.001741323	1.0000000			
Item AMI2	-0.001009944	0.007303489	0.000880345	0.015677099	-0.007118931	1.0000000		
Item AMI3	0.017881700	-0.012714612	0.015315456	-0.012673259	0.010564590	-0.014648236	1.0000000	
Item AMI4	0.020181734	0.006388838	-0.017237614	-0.022407203	0.025364562	0.008657831	-0.001879275	1.0000000



- No residual is over |0.05|.
- Residuals are normally distributed.

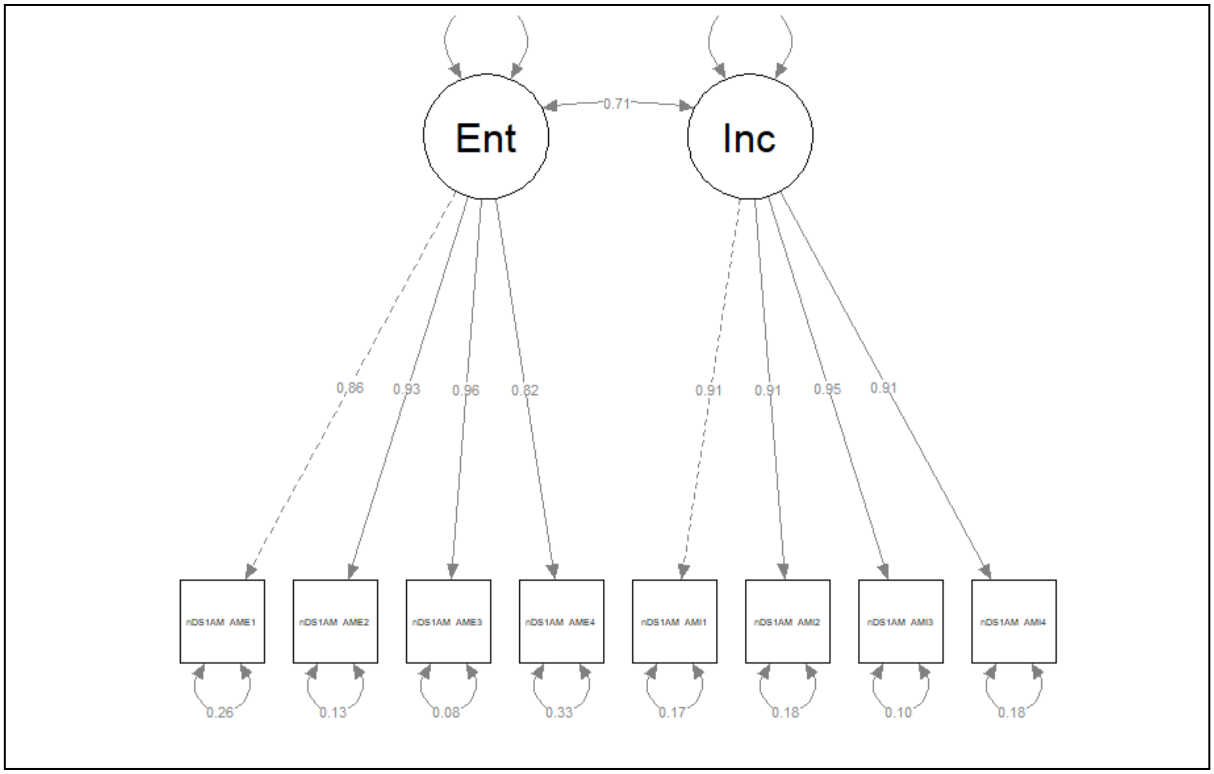
Note: Standardized residuals were all below 0.05 and approximately normally distributed, suggesting that the specified measurement model provides an adequate representation of the observed covariance structure.

<u>Exploratory Factor Analysis, *promax rotation*</u>

Factor Loadings of Exploratory Factor Analysis (EFA) of Early-Semester Academic Mindset				
Survey Item	**Factor 1**	**Factor 2**	**Uniqueness u^2**	**Communality h^2**
Item E1	**0.87**	0.01	0.136	0.86
Item E2	**0.88**	0.01	0.114	0.89
Item E3	**0.98**	-0.02	0.068	0.93
Item E4	**0.75**	0.03	0.304	0.70
Item I1	-0.05	**0.80**	0.280	0.72
Item I2	-0.02	**0.93**	0.152	0.85
Item I3	0.08	**0.82**	0.209	0.79
Item I4	0.04	**0.79**	0.312	0.69
N = 183; Factor loadings reported following a promax (oblique) factor rotation.				

Multidimensional Confirmatory Factor Analysis

Estimator: MLM

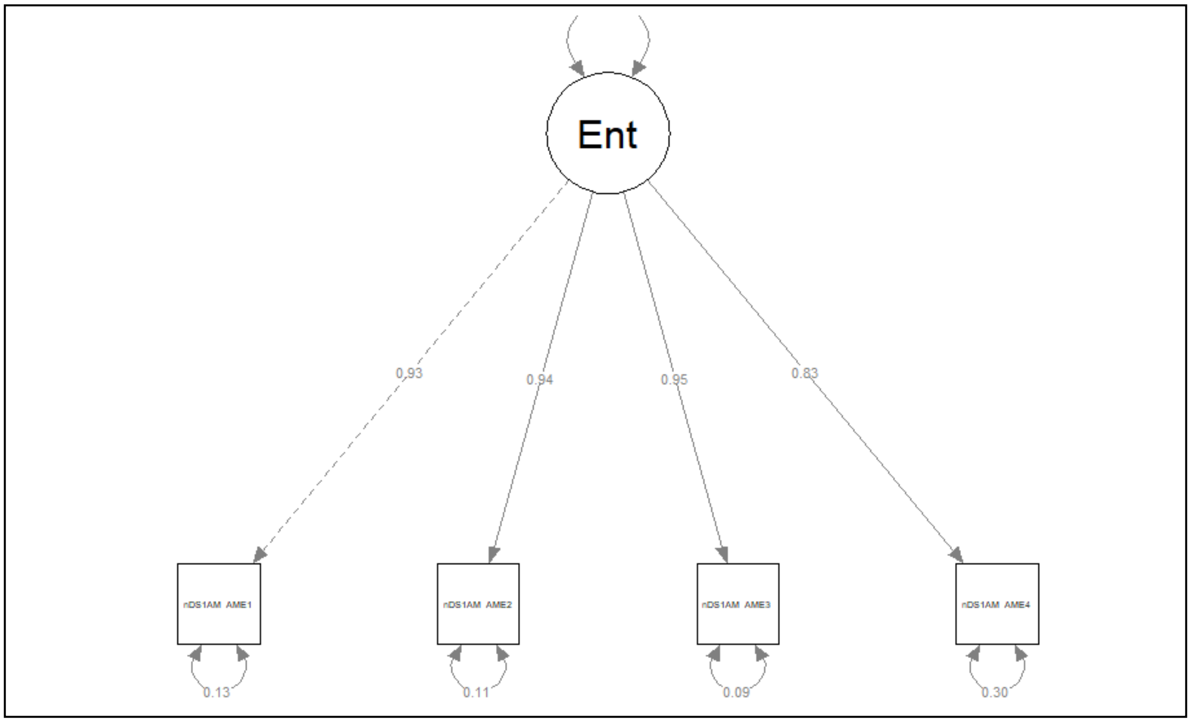


N	χ^2 (*df*, p)	RMSEA	CFI	TLI	SRMR
183	27.978 (19, 0.084)	0.070	0.989	0.984	0.041

This model demonstrates a *close* fit.

<u>Unidimensional Confirmatory Factor Analysis</u>
Estimator: MLM

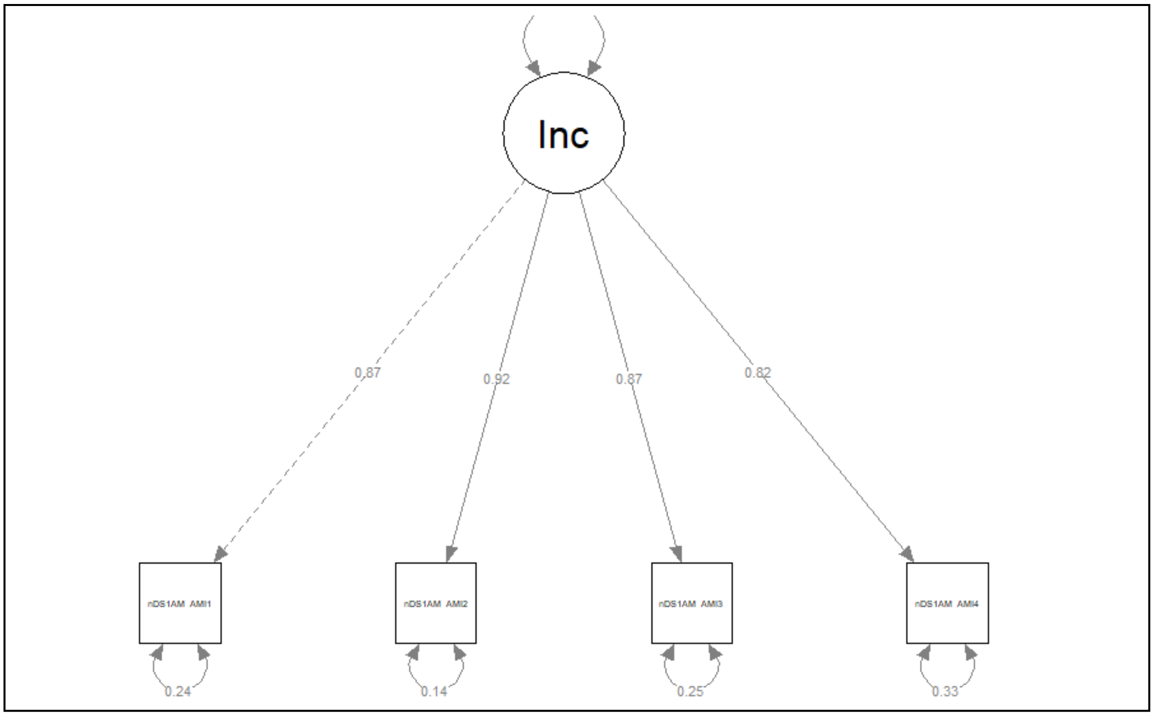
Academic Mindset Entity (4 items loading on one factor)



N	χ^2 (*df*, p)	RMSEA	CFI	TLI	SRMR
183	3.363 (2, 0.186)	0.076	0.997	0.992	0.009

This model demonstrates an *excellent* fit.

Academic Mindset Incremental (4 items loading on one factor)



N	χ^2 (*df*, p)	RMSEA	CFI	TLI	SRMR
183	9.326 (2, 0.009)	0.299	0.946	0.837	0.041

This model demonstrates a *poor* fit. Therefore, we did not use Academic Mindset-Incremental in our SEM path analysis.

References

Bustamante V. P. and Frey R. F., (2025), Before the lecture begins: unpacking how affective measures impact performance in general chemistry 1, *Chemistry Education Research and Practice* (ASAP online)

Edwards J. D., Barthelemy R. S. and Frey R. F., (2021), Relationship between course-level social belonging (sense of belonging and belonging uncertainty) and academic performance in General Chemistry 1, *Journal of Chemical Education*, **99**, 71-82.

Edwards J. D., Laguerre L., Barthelemy R. S., De Grandi C. and Frey R. F., (2022), Relating students' social belonging and course performance across multiple assessment types in a calculus-based introductory physics 1 course, *Phys Rev Phys Edu Res*, **18**, 020150.

Edwards J. D., Torres H. L. and Frey R. F., (2023), The effect of social belonging on persistence to General Chemistry 2, *Journal of Chemical Education*. **100** (11), 4190-4199.

Gunzler D. D. and Morris N., (2015), A tutorial on structural equation modeling for analysis of overlapping symptoms in co-occurring conditions using MPlus, *Statistics in medicine*, **34**, 3246-3280.

B: SEM Path Models

Structural Equation Modeling (SEM) is a statistical technique used to explore relationships among observed (measured) and latent (unobserved) variables, whereby direct and indirect effects can be evaluated (Hancock and Mueller, 2013). Path Analysis**,** on the other hand, is a specialized form of multiple regression analysis that uses diagrams (path models) to visually represent hypothesized causal relationships among observed variables (Beaujean, 2014). In the SEM path models: ovals represent latent variables, squares represent observed variables, directional relations are indicated by straight, single-headed, and double-headed arrows represent covariance. It is important to note that path analysis is not to be used to establish causality or to determine whether a specific model is correct (Streiner, 2005). Path analysis can only determine whether the data are consistent with the path model specified. When paired together, SEM path analysis provides a comprehensive approach for analyzing both simple and intricate relationships displayed in a given path model (Chavance*, et al.*, 2010).

Just as with factor analysis, assumption checks were conducted for these path models to determine suitability. These include ensuring correct model tracings, ample sample size, and a properly identified model. According to Wright's tracing rules, there can be a maximum of one double-headed arrow included in one path (covariance), there is no going forward and then backward (i.e., once you have traveled along a route forward, you cannot travel backward to get to the variable at the end point), and there can be no loops (only non-recursive models are allowed) (Wright, 1918; Wright, 1934). Please see models presented in the main text or the example model below (**Example for Model 1B)** for models that follow Wright's Rules. To determine ample sample size, samples of 100 to 200 are generally considered the minimum necessary to obtain stable, unbiased estimates, while samples of fewer than 100 cases are often viewed as indefensible for most models (Kline, 2011). Others have indicated that at least 10 to 20 cases per free factor loadings (q; $\mathbf{Equation\ 1}$) are needed to provide a minimal basis for unbiased estimation and inference (Bentler and Chou, 1987; Wolf*, et al.*, 2013). All SEM path models use the entire sample size (*N* = 183), which satisfies the requirement provided by Kline. Additionally, we calculated the suggested sample size required by Bentler & Chou and Wolf et al. below.

Determination of an Ample Sample Size

Sample Size (***N***) = 183
Free Factor Loadings (q):

Constructs	**Indicators**	**Factors**	**(Indicators – Factors)**
Imposter Syndrome	4	1	3
Academic Mindset Entity	4	1	3
Sense of Belonging	4	1	3

$$Free\ Loadings\ (q) = \Sigma(indicators/latent\ variable\ - fixed\ loadings/latent\ variable) \quad (1)$$

$$(q) = 9$$

Minimum Recommended Sample Size = (# of cases) x (q: Free Factor Loadings)

= 10 x 9 = 90

Upper/Conservative Recommended Range = (# of cases) x (q: Free Factor Loadings)

= 20 x 9 = 180

Hence, our sample size of 183 is greater than the upper recommended range.

Model Identification

Model identification should be considered to ensure that parameters are uniquely estimated. That is, the number of free factor loadings (q; $\mathbf{Equation\ 1}$) should be less than or equal to the number of non-redundant elements in the sample covariance matrix (p^*). Note: p^* is calculated using **Equation 2,** and p is the sum of manifest variables (items/observed variables) in the model.

$$p^* = \frac{p(p+1)}{2} \quad (2)$$

Example calculation of p and p^*:

p = 4 items (sense of belonging)
+ 4 items (academic mindset entity)
+ 4 items (imposter syndrome)
+ 1 observed var. (prior knowledge)
+ 1 observed var. (exam variable: OC2 Final Exam)

p = 14

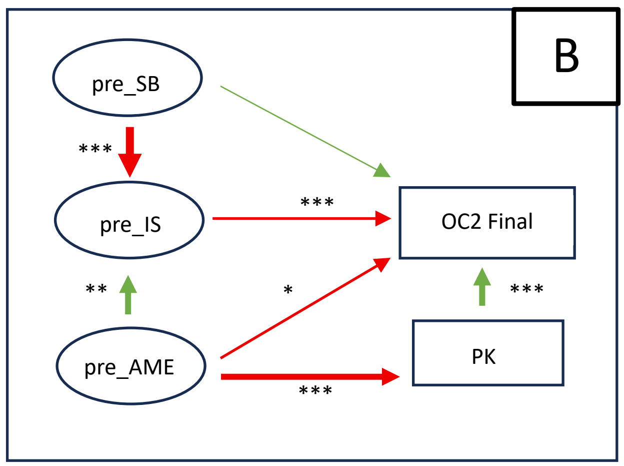
$$p^* = \frac{14(14+1)}{2} = 105$$

$9 < 105$, $q < p^*$ Therefore, **this path model is overidentified**

If $q > p^*$ the model is underidentified (not enough information to estimate all parameters). If $q = p^*$ the model is just-identified (a unique solution exists for each parameter, but the model cannot be tested for fit). If $q < p^*$ the model is overidentified (the model can be tested for fit, and parameter estimates are possible in multiple ways). The total parameters (k) can be found by summing the free factor loadings (q) + observed variable residual variances (Count each observed variable's residual variance as one estimated parameter) + latent variable variances + latent variable covariance + endogenous latent variable residual variances (Count each observed variable's residual variance as one estimated parameter) + number of structural paths (Bentler and Chou, 1987). It is also important to note that the number of parameters affects the degrees of freedom. A positive df is desirable because it indicates that your model is over-identified, meaning there is more information than unknowns, increasing the stability and reliability of the estimates. If df is zero, the model is just-identified, and if df is negative, the model is under-identified, meaning it's unlikely to produce stable or reliable results. The formula for degrees of freedom (*df*) is provided in **Equation 3**.

$$df = \frac{p(p+1)}{2} - k \tag{3}$$

Example Calculation for df:



Example for Model 1B

Degrees of Freedom

$$\begin{aligned} Total\ parameters\ (k) = \\ free\ factor\ loadings\ (q) \\ +\ observed\ variable\ residual\ variances \\ +\ latent\ variable\ variances\ +\ latent\ variable\ covariances \\ +\ endogenous\ \mathbf{latent}\ variable\ residual\ variances \\ +\ number\ of\ structural\ paths \end{aligned}$$

$$\begin{aligned} q &= 9 \\ p &= 14 \\ p^* &= 105 \\ k &= 42 \\ df &= 63 \end{aligned}$$

Therefore,

Total parameters (k) = 9 + 14 + 3 + 2 + 1 + 7 = 42

and

$$\boldsymbol{df} = \frac{p(p+1)}{2} - k = 105 - 42 = \mathbf{63}$$

Our df > 0; hence, our model is over-identified, which increases the stability and reliability of the estimates.

References

Beaujean A. A., (2014), *Latent variable modeling using R: A step-by-step guide*, New York: Routledge.

Bentler P. M. and Chou C.-P., (1987), Practical issues in structural modeling, *Sociological methods & research*, **16**, 78-117.

Chavance M., Escolano S., Romon M., Basdevant A., de Lauzon-Guillain B. and Charles M. A., (2010), Latent variables and structural equation models for longitudinal relationships: an illustration in nutritional epidemiology, *BMC medical research methodology*, **10**, 1-10.

Hancock G. R. and Mueller R. O., (2013), *Structural equation modeling: A second course*, Charlotte, North Carolina: Iap.

Kline R. B., (2011), in *The SAGE handbook of innovation in social research methods*SAGE Publications Ltd, pp. 562-589.

Streiner D. L., (2005), Finding our way: an introduction to path analysis, *The Canadian Journal of Psychiatry*, **50**, 115-122.

Wolf E. J., Harrington K. M., Clark S. L. and Miller M. W., (2013), Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety, *Educational and psychological measurement*, **73**, 913-934.

Wright S., (1918), On the nature of size factors, *Genetics*, **3**, 367.

Wright S., (1934), The method of path coefficients, *The annals of mathematical statistics*, **5**, 161-215.

I. Main Text Models with Bootstrapped Results

II. Path Model A (in Main Text)										
		CFI	TLI	RMSEA	SRMR	AIC	BIC	Estimate	*p-value*	FDR Corr. *p-value*
	Model Fit	0.994	0.992	0.030	0.062	4548.240	4660.188			
Path 1	pre_SB → pre_IS							**-0.541**	**< 0.001**	**4.195 * 10^{-9}**
Path 2	pre_AME → pre_IS							**0.232**	**0.008**	**8.309 * 10^{-3}**
Path 3	pre_AME → PK							**-0.106**	**< 0.001**	**1.274 * 10^{-4}**
Path 4	pre_AME → OC2 Avg. Mididterms							-0.027	0.096	9.605 * 10^{-2}
Path 5	pre_IS → OC2 Avg. Midterms							**-0.031**	**0.022**	**2.183 * 10^{-2}**
Path 6	pre_SB → OC2 Avg. Midterms							0.017	0.248	2.482 * 10^{-1}
Path 7	PK → OC2 Avg. Midterms							**0.202**	**< 0.001**	**6.165 * 10^{-6}**

Observations: 183

df: 70

Parameters: 35

Estimator: MLM, robust

Latent Variables: Sense of Belonging (SB), Imposter Syndrome (IS), Academic Mindset: Entity (AME)

Measured Variables: Prior Knowledge (PK) = Organic 1 Final, Organic 2 Average of Midterm Exams

Model A Bootstrapped Results

Estimator: ML
Test: Bollen.Stine

Iterations: 7500
Parameter Estimates: BCA.simple
*Significant paths bolded

	Paths	β [95% C.I.]	Boot S.E.
Path 1	pre_SB → pre_IS	**-0.541 [-0.370, -0.737]**	**0.092**
Path 2	pre_AME → pre_IS	**0.232 [0.034, 0.411]**	**0.095**
Path 3	pre_AME → PK	**-0.106 [-0.165 , -0.040]**	**0.032**
Path 4	pre_AME → AvgMid	-0.027 [-0.062, 0.004]	0.017
Path 5	pre_IS → AvgMid	**-0.031 [-0.060 , -0.003]**	**0.014**
Path 6	per_SB → AvgMid	0.017 [-0.014, 0.049]	0.016
Path 7	PK → AvgMid	**0.202 [0.123, 0.298]**	**0.045**

Path Model B (in Main Text)

		CFI	TLI	RMSEA	SRMR	AIC	BIC	Estimate	*p-value*	FDR Corr. *p-value*
	Model Fit	0.995	0.993	0.028	0.060	4716.556	4828.888			
Path 1	pre_SB → pre_IS							**-0.548**	**< 0.001**	**2.232 * 10^{-9}**
Path 2	pre_AME → pre_IS							**0.238**	**0.005**	**4.836 * 10^{-3}**
Path 3	pre_AME → PK							**-0.114**	**< 0.001**	**2.142 * 10^{-5}**
Path 4	pre_AME → OC2F							**-0.048**	**0.040**	**4.035 * 10^{-2}**
Path 5	pre_IS → OC2F							**-0.072**	**< 0.001**	**3.750 * 10^{-4}**
Path 6	pre_SB → OC2F							-0.001	0.978	9.784 * 10^{-1}
Path 7	PK → OC2F							**0.241**	**< 0.001**	**2.524 * 10^{-5}**

Observations: 183

df: 70

Parameters: 35

Estimator: MLM, robust

Latent Variables: Sense of Belonging (SB), Imposter Syndrome (IS), Academic Mindset: Entity (AME)

Measured Variables: Prior Knowledge (PK) = Organic 1 Final, Organic 2 Final

Model B Bootstrapped Results

Estimator: ML
Test: Bollen.Stine

Iterations: 7500
Parameter Estimates: BCA.simple
*Significant paths bolded

	Paths	β [95% C.I.]	Boot S.E.
Path 1	pre_SB → pre_IS	**-0.548 [-0.170, -0.050]**	**0.091**
Path 2	pre_AME → pre_IS	**0.238 [0.064, 0.417]**	**0.091**
Path 3	pre_AME → PK	**-0.114 [-0.165 , -0.040]**	**0.031**
Path 4	pre_AME → OC2F	**-0.048 [-0.094, -0.001]**	**0.024**
Path 5	pre_IS → OC2F	**-0.072 [-0.119 , -0.033]**	**0.021**
Path 6	pre_SB → OC2F	-0.001 [-0.042, 0.041]	0.021
Path 7	PK → OC2F	**0.241 [0.138, 0.373]**	**0.059**

II. Additional Models Tested (not included in the main text)

Switch Path: pre_IS → pre_SB (Outcome is OC2 Midterm Avg.)									
		CFI	TLI	RMSEA	SRMR	AIC	BIC	Estimate	*p-value*
	Model Fit	0.991	0.988	0.033	0.083	4554.026	4662.775		
Path 1	**pre_IS → pre_SB**							**-0.555**	**< 0.001**
Path 2	pre_AME → pre_IS							**0.500**	**< 0.001**

Path 3	pre_AME → PK							**-0.106**	**< 0.001**
Path 4	pre_AME → OC2 Avg. Mididterms							-0.027	0.095
Path 5	pre_IS → OC2 Avg. Midterms							**-0.031**	**0.038**
Path 6	pre_SB → OC2 Avg. Midterms							0.017	0.259
Path 7	PK → OC2 Avg. Midterms							**0.202**	**< 0.001**
Switch Path: pre_IS → pre_AME (Outcome is OC2 Midterm Avg.)									
		CFI	TLI	RMSEA	SRMR	AIC	BIC	Estimate	*p-value*
	Model Fit	0.991	0.988	0.037	0.083	4554.026	4662.775		
Path 1	pre_SB → pre_IS							**-0.632**	**< 0.001**
Path 2	**pre_IS → pre_AME**							**0.329**	**< 0.001**
Path 3	pre_AME → PK							**-0.106**	**< 0.001**
Path 4	pre_AME → OC2 Avg. Mididterms							-0.027	0.095
Path 5	pre_IS → OC2 Avg. Midterms							**-0.031**	**0.038**
Path 6	pre_SB → OC2 Avg. Midterms							0.017	0.259
Path 7	PK → OC2 Avg. Midterms							**0.202**	**< 0.001**

Switch Path: pre_IS → pre_SB (Outcome is OC2 Final)									
		CFI	TLI	RMSEA	SRMR	AIC	BIC	Estimate	*p-value*
	Model Fit	0.991	0.989	0.036	0.082	4722.870	4831.992		
Path 1	**pre_IS → pre_SB**							**-0.573**	**< 0.001**
Path 2	pre_AME → pre_IS							**0.518**	**< 0.001**
Path 3	pre_AME → PK							**-0.114**	**< 0.001**
Path 4	pre_AME → OC2F							**-0.047**	**0.044**
Path 5	pre_IS → OC2F							**-0.072**	**0.001**
Path 6	pre_SB → OC2F							-0.001	0.958
Path 7	PK → OC2F							**0.241**	**< 0.001**
Switch Path: pre_IS → pre_AME (Outcome is OC2 Final)									
		CFI	TLI	RMSEA	SRMR	AIC	BIC	Estimate	*p-value*
	Model Fit	0.991	0.989	0.039	0.085	4722.870	4831.992		
Path 1	pre_SB → pre_IS							**-0.646**	**< 0.001**
Path 2	**pre_IS → pre_AME**							**0.356**	**< 0.001**
Path 3	pre_AME → PK							**-0.114**	**< 0.001**
Path 4	pre_AME → OC2F							**-0.047**	**0.044**
Path 5	pre_IS → OC2F							**-0.072**	**0.001**

Path 6	pre_SB → OC2F							-0.001	0.958
Path 7	PK → OC2F							**0.241**	**< 0.001**

Path Model C (used to build more complex models)									
		CFI	TLI	RMSEA	SRMR	AIC	BIC	Estimate	*p-value*
	Model Fit	0.996	0.994	0.027	0.043	4787.414	4890.117		
Path 1	pre_SB → pre_IS							**-0.547**	**< 0.001**
Path 2	pre_AME → pre_IS							**0.237**	**0.005**
Path 3	pre_AME → PK							**-0.064**	**0.049**
Path 4	pre_SB → PK							0.058	0.069
Path 5	pre_IS → PK							-0.043	0.195

Switch Path: pre_AME → pre_SB									
		CFI	TLI	RMSEA	SRMR	AIC	BIC	Estimate	*p-value*
	Model Fit	0.992	0.990	0.036	0.066	4792.696	4892.190		
Path 1	pre_SB → pre_IS							**-0.643**	**< 0.001**
Path 2	**pre_AME → pre_SB**							**-0.489**	**< 0.001**
Path 3	pre_AME → PK							**-0.064**	**0.040**
Path 4	pre_SB → PK							0.058	0.100
Path 5	pre_IS → PK							-0.043	0.182

**Though AIC does not have a meaningful scale, the model with the smallest AIC value is the most likely to replicate and helpful for model selection (Klin, 2023)*

C. Mediation Analysis

Mediation analysis, often conceptualized as models incorporating direct and indirect effects, are widely used in the behavioral sciences (Mayer*, et al.*, 2014). Mediation analysis is useful when researchers assume that the effect of an explanatory variable on an outcome of interest is transmitted through one or more intermediate variables. However, many SEM parameters, particularly indirect effects, do not follow a normal distribution and are often skewed or kurtotic, violating the assumptions of standard statistical inference. Traditional methods for calculating standard errors and confidence intervals (like the delta method or Sobel test) rely on normality assumptions, which can lead to inaccurate or misleading results for non-normally distributed estimates. Additionally, due to the complex nature of SEM models, often involving multiple mediators, latent variables, and non-linear relationships, analytical formulas for standard errors and confidence intervals can become intractable or unavailable for such models. Therefore, bootstrapping is commonly used to obtain more accurate estimates of standard errors and confidence intervals in these situations.

In this study, we used the *manymome* package, an RStudio package developed by Cheung and Cheung (2024), which can be used to estimate and calculate confidence intervals (C.I) for indirect, direct, and total effects. This is done using a two-step approach where model parameters are estimated by SEM using *lavaan*, and then the coefficients are used to compute the requested effects and form confidence intervals (Cheung and Cheung, 2024).

Traditionally, there are two approaches to mediation analysis: the difference method and the product method. In our paper, we used the product method. Below is a brief description of the product method used in the *manymome* package, presented with its mathematical formulation. Consider the following equations (**equations 4-6;** MacKinnon et *al.*, 2004). In these equations, Y_0 is the dependent variable, X is the independent variable, and X_M is the mediator. The total effect of X on Y_0 is denoted by τ, the direct effect of X on Y_0 controlling for the mediator is denoted by τ′, α denotes the effect of X on the mediator, and β denotes the effect of the mediator on Y_0 controlling for X. Residuals (ε_1, ε_2, ε_3) have expected values of zero, and intercepts are β_{01}, β_{02}, and β_{03}. In the first regression equation (equation 4), the dependent variable (Y_0) is estimated based only on the independent variable (X). In the second regression equation (equation 5), the dependent variable (Y_0) is estimated based on both the independent variable (X) and the mediating variable (X_M). In the third regression equation (equation 6), the mediating variable (X_M) is estimated based on the independent variable (X). In this case, the indirect effect equals the product of α and β (Judd & Kenny, 1981).

$$Y_0 = \beta_{01} + \tau X + \varepsilon_1 \quad (4)$$

$$Y_0 = \beta_{02} + \tau' X + \beta X_M + \varepsilon_2 \quad (5)$$

$$X_M = \beta_{03} + \alpha X + \varepsilon_3 \quad (6)$$

Total effects consist of all paths from one variable to another mediated by at least one additional variable (Bollen, 1987). In other words, total effects are equal to the sum of the direct and indirect path effects (MacKinnon*, et al.*, 2004) as shown in **Figure 1** and **Equations 7-8.**

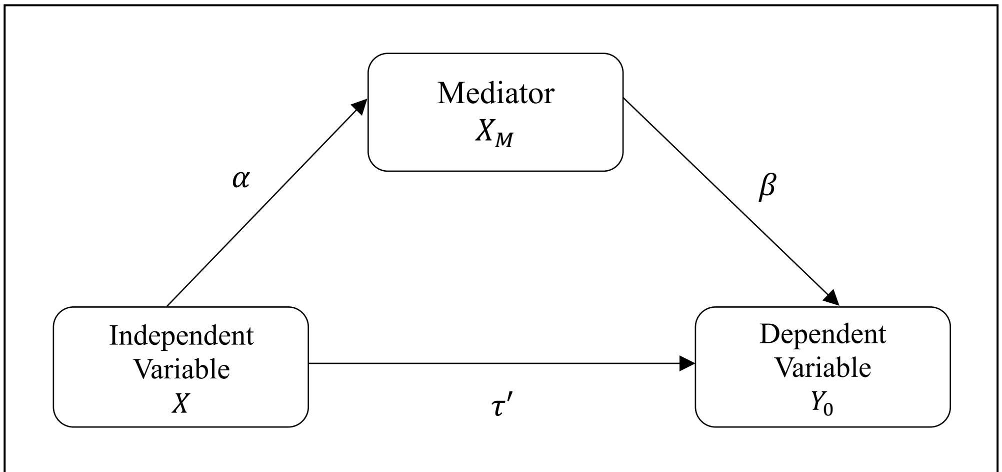


Figure 1: Visual depiction of path effects of a mediated model.

$$Direct\ Effect = \tau' \tag{7}$$

$$Indirect\ Effect = \alpha\beta \tag{8}$$

$$Total\ Effects = \alpha\beta + \tau' \tag{9}$$

Calculation of Indirect and Total Effects for the serial pathways of the two final models shown in the main text

The mediated models presented here are the serial pathways, whereas those in the main text **(Figure 1)** contain the parallel mediation pathways.

Path of Interest: **pre_AME → pre_IS → OC2 Avg. Midterms**

Pathway	Estimate	Boot S.E.
Direct pre_AME → OC2 Avg. Midterms	-0.027 [-0.059, 0.004]	0.016
Indirect pre_AME → pre_IS → OC2 Avg. Midterms	**-0.007 [-0.016, -0.000]**	**0.004**
Total Effects	**-0.035 [-0.065, -0.004]**	**0.016**

Path of Interest: **pre_SB → pre_IS → OC2 Avg. Midterms**

Pathway	Estimate	Boot S.E.
Direct pre_SB → OC2 Avg. Midterms	0.017 [-0.013, 0.048]	0.016
Indirect pre_SB → pre_IS → OC2 Avg. Midterms	**0.017 [0.002, 0.034]**	**0.008**
Total Effects	**0.034 [0.005, 0.063]**	**0.015**

Path of Interest: **pre_AME → pre_IS → OC2 Final**

Pathway	Estimate	Boot S.E.
Direct pre_AME → OC2 Final	**-0.048 [-0.091, -0.004]**	**0.022**
Indirect pre_AME → pre_IS → OC2 Final	**-0.017 [-0.034, -0.003]**	**0.008**
Total Effects	**-0.065 [-0.107, -0.020]**	**0.022**

Path of Interest: **pre_SB → pre_IS → OC2 Final**

Pathway	Estimate	Boot S.E.
Direct pre_SB → OC2 Final	-0.001 [-0.042, 0.042]	0.021
Indirect pre_SB → pre_IS → OC2 Final	**0.039 [0.017, 0.064]**	**0.012**
Total Effects	**-0.039 [0.003, 0.075]**	**0.018**

Path of Interest: **pre_AME → PK → OC2 Avg. Midterms**

Pathway	Estimate	Boot S.E.
Direct pre_AME → OC2 Avg. Midterms	-0.027 [-0.059, 0.004]	0.016
Indirect **pre_AME → PK → OC2 Avg. Midterms**	**-0.021 [-0.039, -0.008]**	**0.008**
Total Effects	**-0.058 [-0.081, -0.015]**	**0.017**

Path of Interest: **pre_AME → PK → OC2 Final**

Pathway	Estimate	Boot S.E.
Direct **pre_AME → OC2 Final**	**-0.048 [-0.092, -0.003]**	**0.023**
Indirect **pre_AME → PK → OC2 Final**	**-0.028 [-0.049, -0.011]**	**0.009**
Total Effects	**-0.076 [-0.121, -0.028]**	**0.024**

References

Bollen K. A., (1987), Total, direct, and indirect effects in structural equation models, *Sociological methodology*, 37-69.

Cheung S. F. and Cheung S.-H., (2024), manymome: An R package for computing the indirect effects, conditional effects, and conditional indirect effects, standardized or unstandardized, and their bootstrap confidence intervals, in many (though not all) models, *Behavior Research Methods*, **56**, 4862-4882.

MacKinnon D. P., Lockwood C. M. and Williams J., (2004), Confidence limits for the indirect effect: Distribution of the product and resampling methods, *Multivariate behavioral research*, **39**, 99-128.

Mayer A., Thoemmes F., Rose N., Steyer R. and West S. G., (2014), Theory and analysis of total, direct, and indirect causal effects, *Multivariate Behavioral Research*, **49**, 425-442.