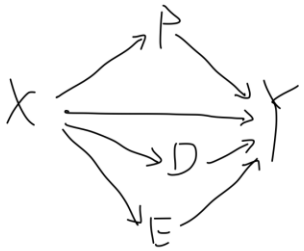
# JiayiShi_js6177_p8158hw6

Jiayi Shi

2023-03-26

## Problem 1

a. 0.56 is the correlations between improvement in Negative Symptoms (SANS Ratings) and positive symptoms for patients given Olanzapine, Haloperidol, or Placebo. 0.31 is the correlations between improvement in Negative Symptoms (SANS Ratings) and depressive symptoms for patients given Olanzapine, Haloperidol, or Placebo. 0.13 is the correlations between improvement in Negative Symptoms (SANS Ratings) and extrapyramidal symptoms for patients given Olanzapine, Haloperidol, or Placebo.




$$P = \alpha_P + \beta_{XP} X$$
$$D = \alpha_D + \beta_{XD} X$$
$$E = \alpha_E + \beta_{XE} X$$
$$Y = \alpha_Y + \beta_{XY} X + \beta_{PY} P + \beta_{DY} d + \beta_{EY} e$$

Figure 2: $Y = -1.91X + (-2.98 \times 0.51)P + (-0.49 \times 0.35)D + (0.25 \times 0.41)E + \varepsilon$

$= -1.91X - 1.52P - 0.17D + 0.1E + \varepsilon$

Figure 4: $Y = -1.92X + (-0.12 \times 0.44)P + (-0.07 \times 0.35)D + (-1.38 \times 0.21)E + \varepsilon$

$= -1.92X - 0.05P - 0.02D - 0.29E + \varepsilon$

b. linear regression. P, D, E represent change in positive, depressive, and extrapyramidal symptoms respectively. In figure 2, X represents high-dose olanzapine versus placebo and Y represents total differential treatment effect on negative symptoms. In figure 4, X represents high-dose olanzapine versus haloperidol and Y represents total differential treatment effect on negative symptoms.

c. The treatment effect denotes the additional change in scores of olanzapine-treated subjects relative to that of subjects who received either placebo or haloperidol. The on negative symptoms is the sum of both the direct effect and the indirect effects.

For high-dose versus placebo:
total effect: $1.52 + 1.91 + 0.17 - 0.10 = 3.5$;
direct effect: $1.91/3.5 * 100\% = 55\%$
indirect effect through positive symptoms: $1.52/3.5 * 100$

indirect effect through depressive symptoms: $0.17/3.5 * 100$
indirect effect through extrapyramidal symptoms: $-0.1/3.5 * 100$

For high-dose olanzapine versus haloperidol:
total effect: $0.05 + 1.92 + 0.02 + 0.29 = 2.28$;
direct effect: $1.92/2.28 * 100\% = 84\%$
indirect effect through positive symptoms: $0.05/2.28 * 100$
indirect effect through depressive symptoms: $0.02/2.28 * 100$
indirect effect through extrapyramidal symptoms: $0.29/2.28 * 100$

d. Compared to placebo, the high-dose olanzapine makes positive symptoms lower by 2.98 points in BPRS positive symptom subscale score, makes depressive symptoms lower by 0.49 points in BPRS item 9 (depressive mood) score, and makes extrapyramidal symptoms higher by 0.25 points in SimpsonAngus scale total score.

## Problem 2

```
pacman::p_load(tidyverse, janitor, lavaan,semPlot)
lvsem_sem <- read_csv("data/SEM for uwcb.csv") %>%
  clean_names()

## Rows: 4746 Columns: 20
## ── Column specification
────────────────────────────────────────────────────────
## Delimiter: ","
## chr (20): id, GENDER, DIET_YR, ENJOY, FRNDWT, WT_IMPRT, TEASEWT, FAM_WT,
AGE...
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this
message.

model_fig1 <- '
  # measurement model
    outuwcb =~ diet_yr + uwcb_yr
    wtteas =~ teasewt + fam_wt

  # correlating the exogenous variables
    wtteas ~~ bmi_sf

  # structural model - direct effects
    bodydiss ~ a*bmi_sf + b*wtteas
    outuwcb ~ c*bmi_sf + d*wtteas + e*bodydiss

  #indirect
    indirect_bmi := a*e
    indirect_wtteas := b*e

  #total
    total_bmi := c+(a*e)
```

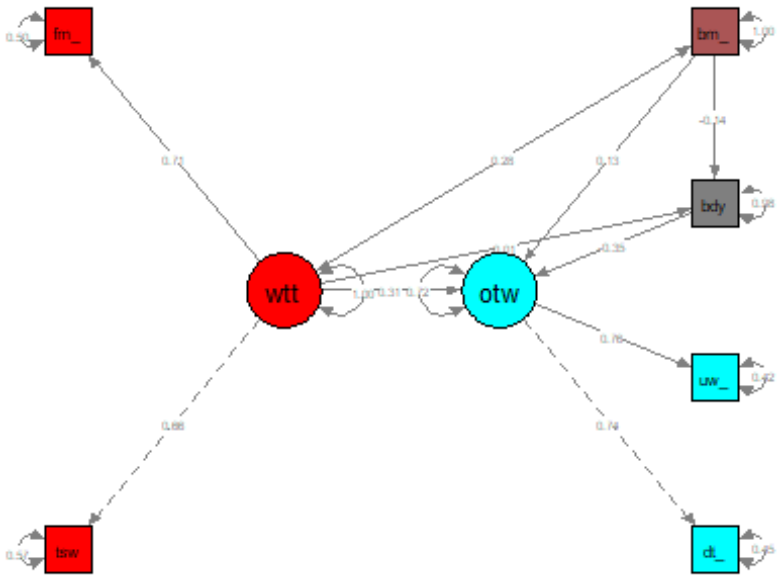
```
    total_wtteas := d+(b*e)
'

model.fit <- sem(model_fig1, data = lvsem_sem, sample.cov=TRUE, missing =
"ML")

summary(model.fit, standardized = TRUE, fit.measures=TRUE)
```

```
## lavaan 0.6.13 ended normally after 79 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        22
##
##   Number of observations                          4746
##   Number of missing patterns                         1
##
## Model Test User Model:
##
##   Test statistic                               118.927
##   Degrees of freedom                                 5
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              4339.632
##   Degrees of freedom                                15
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.974
##   Tucker-Lewis Index (TLI)                       0.921
##
##   Robust Comparative Fit Index (CFI)             0.974
##   Robust Tucker-Lewis Index (TLI)                0.921
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -67546.941
##   Loglikelihood unrestricted model (H1)             NA
##
##   Akaike (AIC)                               135137.881
##   Bayesian (BIC)                             135280.113
##   Sample-size adjusted Bayesian (SABIC)      135210.205
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.069
##   90 Percent confidence interval - lower         0.059
```

```
##   90 Percent confidence interval - upper         0.080
##   P-value H_0: RMSEA <= 0.050                    0.001
##   P-value H_0: RMSEA >= 0.080                    0.056
##
##   Robust RMSEA                                   0.069
##   90 Percent confidence interval - lower         0.059
##   90 Percent confidence interval - upper         0.080
##   P-value H_0: Robust RMSEA <= 0.050             0.001
##   P-value H_0: Robust RMSEA >= 0.080             0.056
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.022
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Observed
##   Observed information based on                Hessian
##
## Latent Variables:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   outuwcb =~
##     diet_yr          1.000                               0.728    0.739
##     uwcb_yr          1.543    0.057   27.205    0.000    1.124    0.761
##   wtteas =~
##     teasewt          1.000                               0.825    0.658
##     fam_wt           0.431    0.028   15.313    0.000    0.355    0.706
##
## Regressions:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   bodydiss ~
##     bmi_sf     (a)   -0.033    0.004   -8.950    0.000   -0.033   -0.140
##     wtteas     (b)    0.111    0.290    0.383    0.702    0.092    0.008
##   outuwcb ~
##     bmi_sf     (c)    0.002    0.000    7.609    0.000    0.003    0.132
##     wtteas     (d)    0.273    0.022   12.663    0.000    0.309    0.309
##     bodydiss   (e)   -0.023    0.001  -20.238    0.000   -0.032   -0.351
##
## Covariances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   wtteas ~~
##     bmi_sf          10.853    0.856   12.686    0.000   13.155    0.281
##
## Intercepts:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .diet_yr          2.981    0.042   71.714    0.000    2.981    3.024
##    .uwcb_yr          3.587    0.065   55.131    0.000    3.587    2.430
##    .teasewt          2.682    0.018  147.467    0.000    2.682    2.141
##    .fam_wt           2.144    0.007  293.338    0.000    2.144    4.258
```

```
##    .bodydiss          28.183    0.391   72.103    0.000   28.183    2.549
##     bmi_sf            97.004    0.680  142.708    0.000   97.004    2.071
##    .outuwcb            0.000                               0.000    0.000
##     wtteas             0.000                               0.000    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .diet_yr           0.441    0.020   21.888    0.000    0.441    0.454
##    .uwcb_yr           0.915    0.047   19.568    0.000    0.915    0.420
##    .teasewt           0.889    0.047   18.906    0.000    0.889    0.566
##    .fam_wt            0.127    0.008   15.046    0.000    0.127    0.502
##    .bodydiss        119.890    2.461   48.710    0.000  119.890    0.981
##     bmi_sf         2192.862   45.016   48.713    0.000 2192.862    1.000
##    .outuwcb           0.383    0.019   19.694    0.000    0.721    0.721
##     wtteas            0.681    0.051   13.391    0.000    1.000    1.000
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##     indirect_bmi      0.001    0.000    7.966    0.000    0.001    0.049
##     indirect_wttes   -0.003    0.007   -0.381    0.703   -0.003   -0.003
##     total_bmi         0.003    0.000    9.862    0.000    0.004    0.181
##     total_wtteas      0.270    0.021   12.727    0.000    0.306    0.306
```

```
semPaths(model.fit,intercept = FALSE, whatLabels="std", reorder = FALSE,
         rainbow = .5, groups = "latents",rotation = 2,layout = "tree")
```




| | direct effect | indirect effect | total effect |
|---|---|---|---|

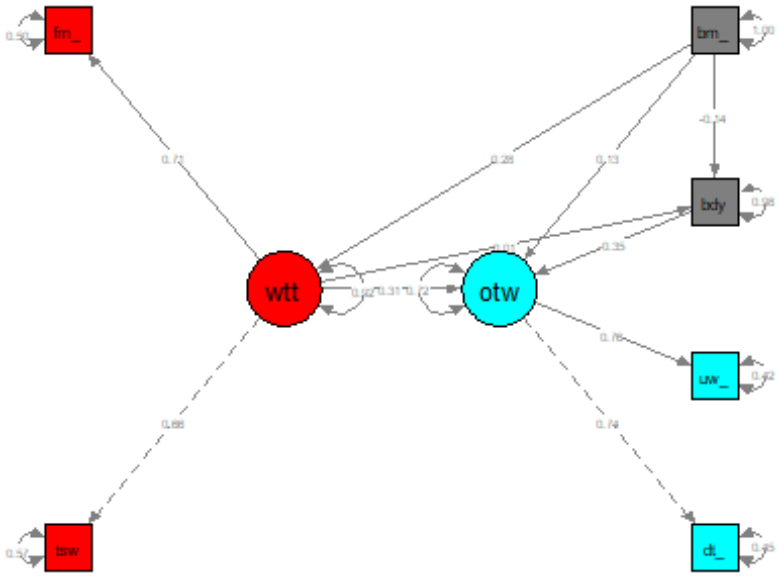
| | direct effect | indirect effect | total effect |
|---|---|---|---|
| wtteas to outuwcb | 0.309 | -0.003 | 0.306 |
| bmi_sf to outuwcb | 0.132 | 0.049 | 0.181 |

Goodness of fit statistics:
Chi-square=118.927, d.f.=5, RMSEA=0.069, CFI=0.974

```
model_fig1_mod <- '
  # measurement model
    outuwcb =~ diet_yr + uwcb_yr
    wtteas =~ teasewt + fam_wt

  # structural model - direct effects
    bodydiss ~ a*bmi_sf + b*wtteas
    outuwcb ~ c*bmi_sf + d*wtteas + e*bodydiss
    wtteas ~ f*bmi_sf

  #indirect
    indirect_bmi := a*e + f*b*e + f*d
    indirect_wtteas := b*e

  #total
    total_bmi := c + (a*e + f*b*e + f*d)
    total_wtteas := d + (b*e)
'

model.fit2 <- sem(model_fig1_mod, data = lvsem_sem, sample.cov=TRUE, missing
= "ML", fixed.x=FALSE)

summary(model.fit2, standardized = TRUE, fit.measures=TRUE)

## lavaan 0.6.13 ended normally after 71 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        22
##
##   Number of observations                          4746
##   Number of missing patterns                         1
##
## Model Test User Model:
##
##   Test statistic                               118.927
##   Degrees of freedom                                 5
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              4339.632
```

```
##   Degrees of freedom                                15
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.974
##   Tucker-Lewis Index (TLI)                       0.921
##
##   Robust Comparative Fit Index (CFI)             0.974
##   Robust Tucker-Lewis Index (TLI)                0.921
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -67546.941
##   Loglikelihood unrestricted model (H1)             NA
##
##   Akaike (AIC)                              135137.881
##   Bayesian (BIC)                            135280.113
##   Sample-size adjusted Bayesian (SABIC)     135210.205
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.069
##   90 Percent confidence interval - lower         0.059
##   90 Percent confidence interval - upper         0.080
##   P-value H_0: RMSEA <= 0.050                    0.001
##   P-value H_0: RMSEA >= 0.080                    0.056
##
##   Robust RMSEA                                   0.069
##   90 Percent confidence interval - lower         0.059
##   90 Percent confidence interval - upper         0.080
##   P-value H_0: Robust RMSEA <= 0.050             0.001
##   P-value H_0: Robust RMSEA >= 0.080             0.056
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.022
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Observed
##   Observed information based on                Hessian
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   outuwcb =~
##     diet_yr           1.000                               0.728    0.739
##     uwcb_yr           1.543    0.057   27.205    0.000    1.124    0.761
##   wtteas =~
```

```
##     teasewt           1.000                                0.825    0.658
##     fam_wt            0.431    0.028   15.313    0.000    0.355    0.706
##
## Regressions:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   bodydiss ~
##     bmi_sf     (a)   -0.033    0.004   -8.950    0.000   -0.033   -0.140
##     wtteas     (b)    0.111    0.290    0.383    0.702    0.092    0.008
##   outuwcb ~
##     bmi_sf     (c)    0.002    0.000    7.609    0.000    0.003    0.132
##     wtteas     (d)    0.273    0.022   12.663    0.000    0.309    0.309
##     bodydiss   (e)   -0.023    0.001  -20.238    0.000   -0.032   -0.351
##   wtteas ~
##     bmi_sf     (f)    0.005    0.000   13.139    0.000    0.006    0.281
##
## Intercepts:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .diet_yr           2.850    0.041   70.116    0.000    2.850    2.891
##    .uwcb_yr           3.385    0.063   53.569    0.000    3.385    2.293
##    .teasewt           2.202    0.041   54.127    0.000    2.202    1.757
##    .fam_wt            1.937    0.016  124.100    0.000    1.937    3.847
##    .bodydiss         28.130    0.366   76.939    0.000   28.130    2.545
##     bmi_sf           97.004    0.680  142.708    0.000   97.004    2.071
##    .outuwcb           0.000                                0.000    0.000
##    .wtteas            0.000                                0.000    0.000
##
## Variances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .diet_yr           0.441    0.020   21.888    0.000    0.441    0.454
##    .uwcb_yr           0.915    0.047   19.568    0.000    0.915    0.420
##    .teasewt           0.889    0.047   18.906    0.000    0.889    0.566
##    .fam_wt            0.127    0.008   15.046    0.000    0.127    0.502
##    .bodydiss        119.890    2.461   48.710    0.000  119.890    0.981
##    .outuwcb           0.383    0.019   19.694    0.000    0.721    0.721
##    .wtteas            0.627    0.046   13.535    0.000    0.921    0.921
##     bmi_sf         2192.862   45.016   48.713    0.000 2192.862    1.000
##
## Defined Parameters:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##     indirect_bmi      0.002    0.000   13.391    0.000    0.003    0.135
##     indirect_wttes   -0.003    0.007   -0.381    0.703   -0.003   -0.003
##     total_bmi         0.004    0.000   14.992    0.000    0.006    0.267
##     total_wtteas      0.270    0.021   12.727    0.000    0.306    0.306

semPaths(model.fit2,intercept = FALSE, whatLabels="std", reorder = FALSE,
         rainbow = .5, groups = "latents",rotation = 2,layout = "tree")
```




| | direct effect | indirect effect | total effect |
|---|---|---|---|
| wtteas to outuwcb | 0.309 | -0.003 | 0.306 |
| bmi_sf to outuwcb | 0.132 | 0.135 | 0.267 |

Goodness of fit statistics remain the same as in (a):
Chi-square=118.927, d.f.=5, RMSEA=0.069, CFI=0.974

```
model_fig2 <- '
  # structural model - direct effects
    teasewt ~ a*bmi_sf
    bodydiss ~ b*bmi_sf + c*teasewt
    uwcb_yr ~ d*bmi_sf + e*bodydiss + f*teasewt
    diet_yr ~ g*bmi_sf + h*bodydiss + j*teasewt

  # indirect effects
    ind_bm_uwcb := b*e+a*f+a*c*e
    ind_bm_diet := b*h+a*j+a*c*h
    ind_tease_uwcb := c*e
    ind_tease_diet := c*h

  #total effects
    tot_bm_uwcb := d+b*e+a*f+a*c*e
    tot_bm_diet := g+b*h+a*j+a*c*h
    tot_tease_uwcb := f+c*e
    tot_tease_diet := j+c*h
```

```
'


### The default is listwise deletion for missing data
###adding the optoin missing = "ML" uses full information maximum likelihood
model.fit3 <- sem(model_fig2, data = lvsem_sem, sample.cov=TRUE, missing =
"ML", fixed.x=FALSE)

summary(model.fit3, standardized = TRUE, fit.measures=TRUE)

## lavaan 0.6.13 ended normally after 62 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        20
##
##   Number of observations                          4746
##   Number of missing patterns                         1
##
## Model Test User Model:
##
##   Test statistic                                 0.000
##   Degrees of freedom                                 0
##
## Model Test Baseline Model:
##
##   Test statistic                              2958.777
##   Degrees of freedom                                10
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    1.000
##   Tucker-Lewis Index (TLI)                       1.000
##
##   Robust Comparative Fit Index (CFI)             1.000
##   Robust Tucker-Lewis Index (TLI)                1.000
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -64699.696
##   Loglikelihood unrestricted model (H1)             NA
##
##   Akaike (AIC)                               129439.392
##   Bayesian (BIC)                             129568.693
##   Sample-size adjusted Bayesian (SABIC)      129505.141
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.000
##   90 Percent confidence interval - lower         0.000
```

```
##   90 Percent confidence interval - upper         0.000
##   P-value H_0: RMSEA <= 0.050                       NA
##   P-value H_0: RMSEA >= 0.080                       NA
##
##   Robust RMSEA                                   0.000
##   90 Percent confidence interval - lower         0.000
##   90 Percent confidence interval - upper         0.000
##   P-value H_0: Robust RMSEA <= 0.050                NA
##   P-value H_0: Robust RMSEA >= 0.080                NA
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.000
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Observed
##   Observed information based on                Hessian
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   teasewt ~
##     bmi_sf     (a)    0.005    0.000   14.026    0.000    0.005    0.200
##   bodydiss ~
##     bmi_sf     (b)   -0.029    0.003   -8.304    0.000   -0.029   -0.121
##     teasewt    (c)   -0.697    0.129   -5.400    0.000   -0.697   -0.079
##   uwcb_yr ~
##     bmi_sf     (d)    0.004    0.000    8.817    0.000    0.004    0.123
##     bodydiss   (e)   -0.035    0.002  -19.006    0.000   -0.035   -0.260
##     teasewt    (f)    0.180    0.016   11.023    0.000    0.180    0.153
##   diet_yr ~
##     bmi_sf     (g)    0.003    0.000   11.250    0.000    0.003    0.157
##     bodydiss   (h)   -0.021    0.001  -17.477    0.000   -0.021   -0.241
##     teasewt    (j)    0.093    0.011    8.446    0.000    0.093    0.118
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##  .uwcb_yr ~~
##    .diet_yr           0.638    0.021   30.873    0.000    0.638    0.501
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .teasewt           2.164    0.041   52.794    0.000    2.164    1.727
##    .bodydiss         29.638    0.459   64.541    0.000   29.638    2.681
##    .uwcb_yr           3.015    0.079   38.007    0.000    3.015    2.043
##    .diet_yr           2.569    0.053   48.219    0.000    2.569    2.607
##     bmi_sf           97.004    0.680  142.708    0.000   97.004    2.071
##
## Variances:
```

```
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .teasewt           1.507    0.031   48.713    0.000    1.507    0.960
##    .bodydiss        119.166    2.446   48.713    0.000  119.166    0.975
##    .uwcb_yr           1.893    0.039   48.713    0.000    1.893    0.869
##    .diet_yr           0.855    0.018   48.713    0.000    0.855    0.880
##     bmi_sf         2192.862   45.016   48.713    0.000 2192.862    1.000
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##     ind_bm_uwcb       0.002    0.000   12.013    0.000    0.002    0.066
##     ind_bm_diet       0.001    0.000   11.048    0.000    0.001    0.057
##     ind_tease_uwcb    0.024    0.005    5.194    0.000    0.024    0.021
##     ind_tease_diet    0.015    0.003    5.159    0.000    0.015    0.019
##     tot_bm_uwcb       0.006    0.000   13.250    0.000    0.006    0.189
##     tot_bm_diet       0.005    0.000   15.090    0.000    0.005    0.214
##     tot_tease_uwcb    0.204    0.017   12.095    0.000    0.204    0.173
##     tot_tease_diet    0.108    0.011    9.539    0.000    0.108    0.137
```

| | direct effect | indirect effect | total effect |
|---|---|---|---|
| bmi_sf to uwcb_yr | 0.123 | 0.066 | 0.189 |
| bmi_sf to diet_yr | 0.157 | 0.057 | 0.214 |
| teaswt to uwcb_yr | 0.153 | 0.021 | 0.173 |
| teaswt to diet_yr | 0.118 | 0.019 | 0.137 |

Goodness of fit statistics:
Chi-square=0, d.f.=0, RMSEA=0, CFI=1

d. Similarity: Both models test causal effect of BMI and unhealthy weight control behaviors as well as causal effect of tease tendency and unhealthy weight control behaviors, viewing body satisfaction as a mediator. Difference: Model in b considers latent variable. It considers tendency to tease of both friends and family, and consider diet and other specific behaviors as a whole. Model in c only considers tendency of friends to tease participants about their weights, and considers two outcomes separately.