waveley_attempt

Waveley Qiu (wq2162)

2022-04-30

How does an athlete's measure of athletic identity affect MHC-SF, as mediated through resilience and a healthy lifestyle?

Latent Variable Construction

Latent Variable 1: Athletic Identity

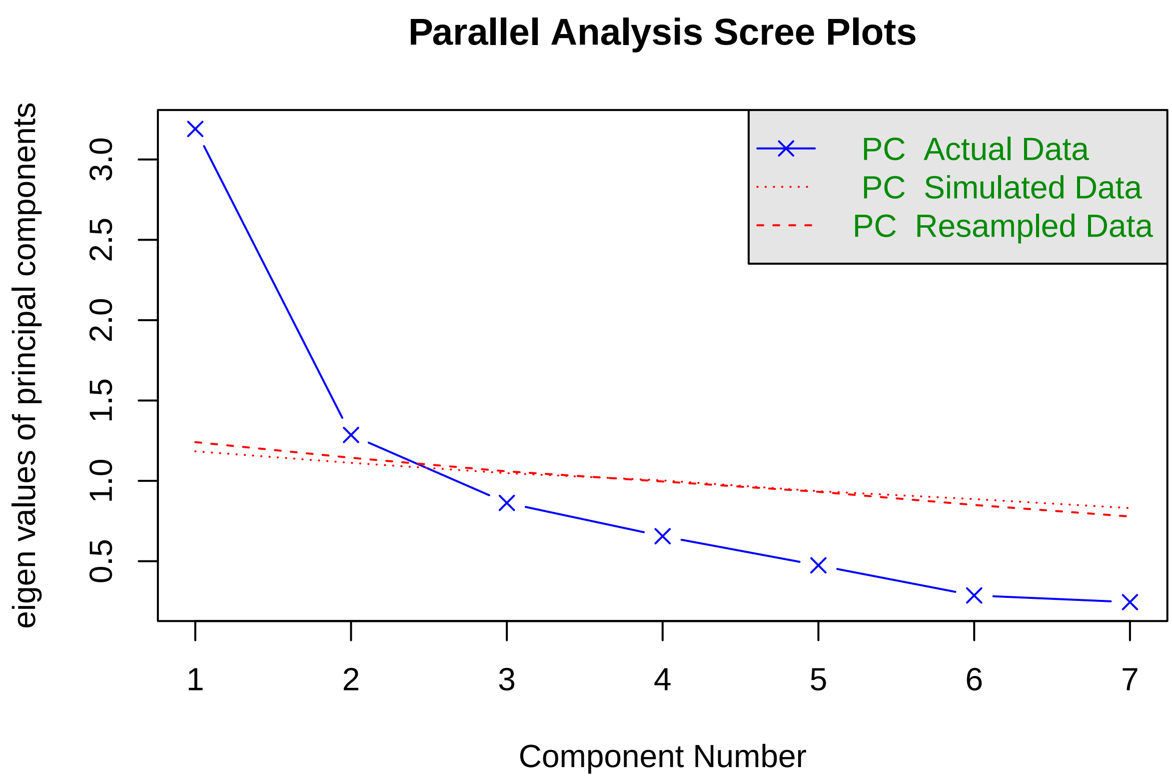
First, let's select the variables we are interested in.

```
athletic_identity <- c("cnsdr_ath", "sprt_goals", "frnds_ath", "sprt_impt", "think_sprt", "bad_sprt", "

athletic_identity_numeric <- athletes[,athletic_identity] %>% map_df(., as.numeric)

athletic_identity_matrix <- athletic_identity_numeric %>% as.matrix()
```

Polychoric Correlations

Now, let us determine the number of factors that might underlie these variables.

```
athlete_parallel <- fa.parallel(athletic_identity_matrix, cor = "poly", fa = "pc")
```




```
## Parallel analysis suggests that the number of factors =  NA  and the number of components =  2
```

```
athlete_parallel$pc.values
```

```
## [1] 3.1901693 1.2855443 0.8626912 0.6554703 0.4741985 0.2866577 0.2452687
```

PCA indicates that two factors underlie these variables.

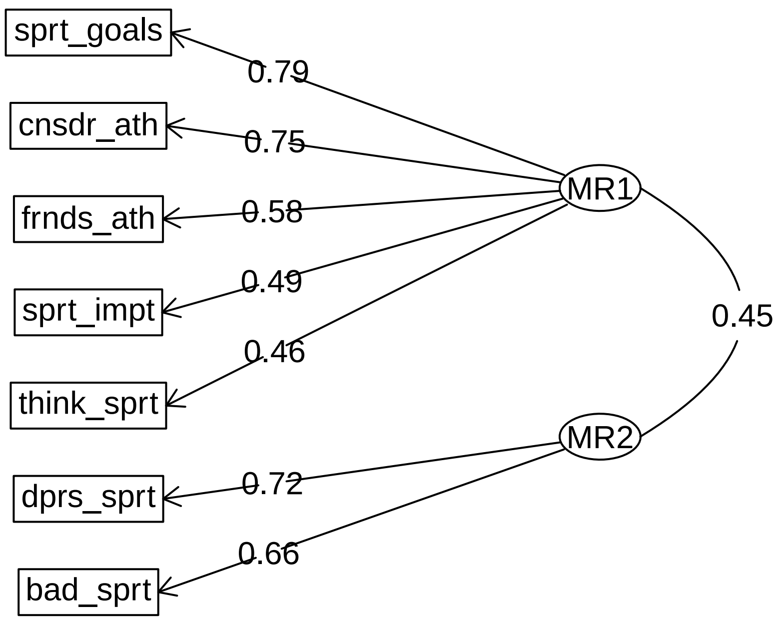
EFA

We now want to examine which variables might load on which factors. To do so, we will perform EFA on a 2-factor model, and also on 1- and 3- factor models.

First, the 2-factor model:

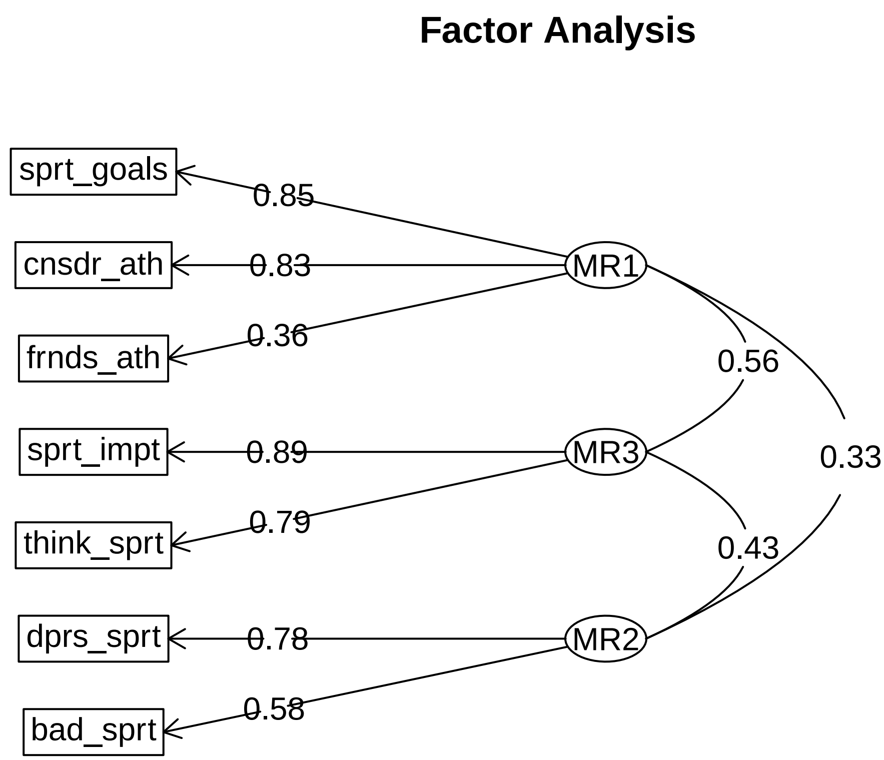
```
athletic_efa2 <- fa(r = athletic_identity_matrix, nfactors = 2, cor = "poly")
fa.diagram(athletic_efa2, digits = 2, simple = TRUE)
```

Factor Analysis



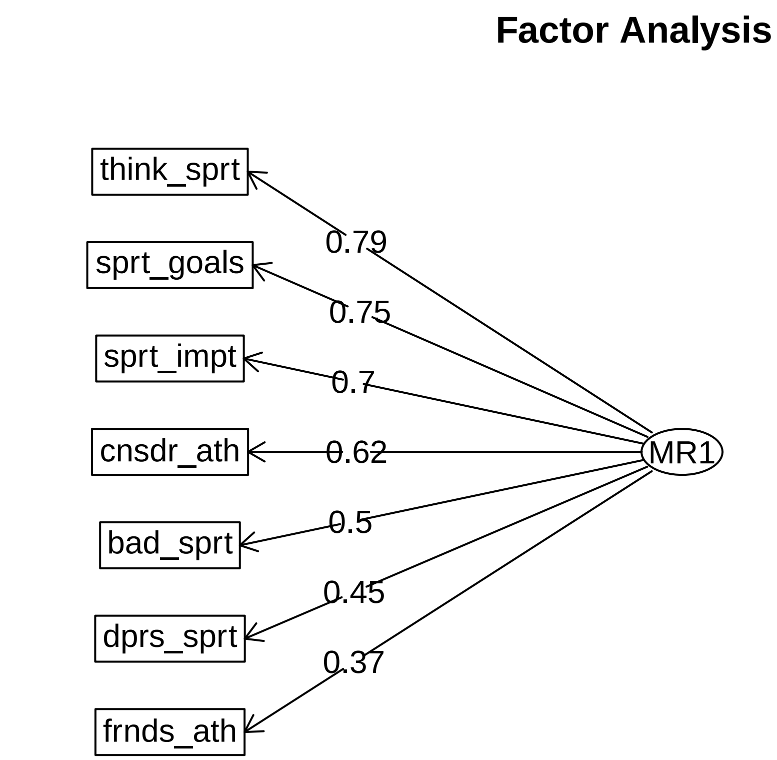
Now, the 3-factor model:

```
athletic_efa3 <- fa(r = athletic_identity_matrix, nfactors = 3, cor = "poly")
fa.diagram(athletic_efa3, digits = 2, simple = TRUE)
```



Now, the 1-factor model:

```
athletic_efa1 <- fa(r = athletic_identity_matrix, nfactors = 1, cor = "poly")
fa.diagram(athletic_efa1, digits = 2, simple = TRUE)
```



The 3-factor model seems to fit the data the best, as it has the smallest BIC. We will proceed by using the 3-factor model for athletic_identity.

Reliability

LV 1: External Identity

```
external_identity <- c("cnsdr_ath", "sprt_goals", "frnds_ath")

external_identity_numeric <- athletes[,external_identity] %>% map_df(., as.numeric)

external_identity_matrix <- external_identity_numeric %>% as.matrix()
psych::alpha(external_identity_matrix)
```

```
##
## Reliability analysis
## Call: psych::alpha(x = external_identity_matrix)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.65      0.68    0.62      0.42 2.1 0.032  5.7 0.89     0.33
##
##  lower alpha upper     95% confidence boundaries
## 0.59 0.65 0.72
##
##  Reliability if an item is dropped:
##            raw_alpha std.alpha G6(smc) average_r  S/N alpha se var.r med.r
## cnsdr_ath       0.47      0.49    0.33      0.33 0.97    0.052    NA  0.33
```

```
## sprt_goals      0.46      0.47    0.31      0.31 0.89    0.055    NA  0.31
## frnds_ath       0.75      0.76    0.61      0.61 3.15    0.025    NA  0.61
##
##  Item statistics
##              n raw.r std.r r.cor r.drop mean   sd
## cnsdr_ath  356  0.79  0.82  0.71   0.53  5.9 1.11
## sprt_goals 356  0.78  0.83  0.72   0.56  6.1 0.96
## frnds_ath  356  0.76  0.70  0.41   0.35  5.2 1.38
##
## Non missing response frequency for each item
##               1    2    3    4    5    6    7 miss
## cnsdr_ath  0.01 0.01 0.03 0.05 0.24 0.33 0.35 0.02
## sprt_goals 0.00 0.00 0.02 0.03 0.20 0.35 0.40 0.02
## frnds_ath  0.01 0.04 0.06 0.10 0.31 0.29 0.17 0.02
```

LV 2: Internal Value

```
internal_value <- c("sprt_impt", "think_sprt")

internal_value_numeric <- athletes[,internal_value] %>% map_df(., as.numeric)

internal_value_matrix <- internal_value_numeric %>% as.matrix()
psych::alpha(internal_value_matrix)
```

```
##
## Reliability analysis
## Call: psych::alpha(x = internal_value_matrix)
##
##   raw_alpha std.alpha G6(smc) average_r S/N  ase mean  sd median_r
##       0.81      0.81    0.67      0.67 4.1 0.02    5 1.4     0.67
##
##  lower alpha upper     95% confidence boundaries
## 0.77 0.81 0.85
##
##  Reliability if an item is dropped:
##            raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
## sprt_impt       0.67      0.67    0.45      0.67 2.1       NA     0  0.67
## think_sprt      0.68      0.67    0.45      0.67 2.1       NA     0  0.67
##
##  Item statistics
##              n raw.r std.r r.cor r.drop mean  sd
## sprt_impt  356  0.91  0.91  0.75   0.67  5.1 1.5
## think_sprt 356  0.92  0.91  0.75   0.67  4.9 1.5
##
## Non missing response frequency for each item
##               1    2    3    4    5    6    7 miss
## sprt_impt  0.03 0.04 0.10 0.10 0.28 0.27 0.18 0.02
## think_sprt 0.02 0.05 0.14 0.12 0.29 0.22 0.16 0.02
```

LV 3: Negative Events

```
negative_events <- c("dprs_sprt", "bad_sprt")

negative_events_numeric <- athletes[,negative_events] %>% map_df(., as.numeric)

negative_events_matrix <- negative_events_numeric %>% as.matrix()
```

```r
psych::alpha(negative_events_matrix)
```

```
##
## Reliability analysis
## Call: psych::alpha(x = negative_events_matrix)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean  sd median_r
##       0.63      0.63    0.46      0.46 1.7 0.039  5.6 1.2     0.46
##
##  lower alpha upper     95% confidence boundaries
## 0.55 0.63 0.7
##
##  Reliability if an item is dropped:
##           raw_alpha std.alpha G6(smc) average_r  S/N alpha se var.r med.r
## dprs_sprt      0.40      0.46    0.21      0.46 0.85       NA     0  0.46
## bad_sprt       0.53      0.46    0.21      0.46 0.85       NA     0  0.46
##
##  Item statistics
##             n raw.r std.r r.cor r.drop mean  sd
## dprs_sprt 356  0.83  0.85  0.58   0.46  5.7 1.3
## bad_sprt  356  0.88  0.85  0.58   0.46  5.5 1.4
##
## Non missing response frequency for each item
##              1    2    3    4    5    6    7 miss
## dprs_sprt 0.01 0.02 0.02 0.06 0.30 0.29 0.29 0.02
## bad_sprt  0.02 0.04 0.04 0.08 0.24 0.29 0.28 0.02
```

Our final model for athlete_identity is as follows:

external_identity = sprt_goals + cnsdr_ath + frnds_ath internal_value = sprt_impt + think_sprt negative_events = dprs_sprt + bad_sprt

CFA

```r
athlete_model <-
'external_identity =~ sprt_goals + cnsdr_ath + frnds_ath
 internal_value =~ sprt_impt + think_sprt
 negative_events =~ dprs_sprt + bad_sprt'

athlete_CFA = cfa(athlete_model, data = athletic_identity_matrix,
                  ordered = names(athletic_identity_matrix),
                  std.lv = TRUE)
summary(athlete_CFA, fit.measures = TRUE)
```

```
## lavaan 0.6-10 ended normally after 28 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        17
##
##                                                   Used       Total
##   Number of observations                           356         363
##
## Model Test User Model:
##
```

```
##   Test statistic                                23.840
##   Degrees of freedom                                11
##   P-value (Chi-square)                           0.013
##
## Model Test Baseline Model:
##
##   Test statistic                               683.839
##   Degrees of freedom                                21
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.981
##   Tucker-Lewis Index (TLI)                       0.963
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -3851.304
##   Loglikelihood unrestricted model (H1)      -3839.384
##
##   Akaike (AIC)                                7736.608
##   Bayesian (BIC)                              7802.482
##   Sample-size adjusted Bayesian (BIC)         7748.550
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.057
##   90 Percent confidence interval - lower         0.025
##   90 Percent confidence interval - upper         0.089
##   P-value RMSEA <= 0.05                          0.314
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.039
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                        Estimate  Std.Err  z-value  P(>|z|)
##   external_identity =~
##     sprt_goals            0.831    0.054   15.435    0.000
##     cnsdr_ath             0.785    0.061   12.801    0.000
##     frnds_ath             0.548    0.078    7.011    0.000
##   internal_value =~
##     sprt_impt             1.086    0.082   13.210    0.000
##     think_sprt            1.430    0.085   16.847    0.000
##   negative_events =~
##     dprs_sprt             0.802    0.085    9.424    0.000
##     bad_sprt              1.034    0.103   10.083    0.000
##
```

```
## Covariances:
##                         Estimate  Std.Err  z-value  P(>|z|)
##   external_identity ~~
##     internal_value         0.537    0.052   10.388    0.000
##     negative_evnts         0.387    0.068    5.701    0.000
##   internal_value ~~
##     negative_evnts         0.515    0.061    8.454    0.000
##
## Variances:
##                     Estimate  Std.Err  z-value  P(>|z|)
##    .sprt_goals         0.237    0.062    3.840    0.000
##    .cnsdr_ath          0.615    0.071    8.696    0.000
##    .frnds_ath          1.604    0.126   12.732    0.000
##    .sprt_impt          1.093    0.128    8.564    0.000
##    .think_sprt         0.289    0.171    1.689    0.091
##    .dprs_sprt          0.928    0.120    7.725    0.000
##    .bad_sprt           1.001    0.179    5.581    0.000
##     external_dntty     1.000
##     internal_value     1.000
##     negative_evnts     1.000
```

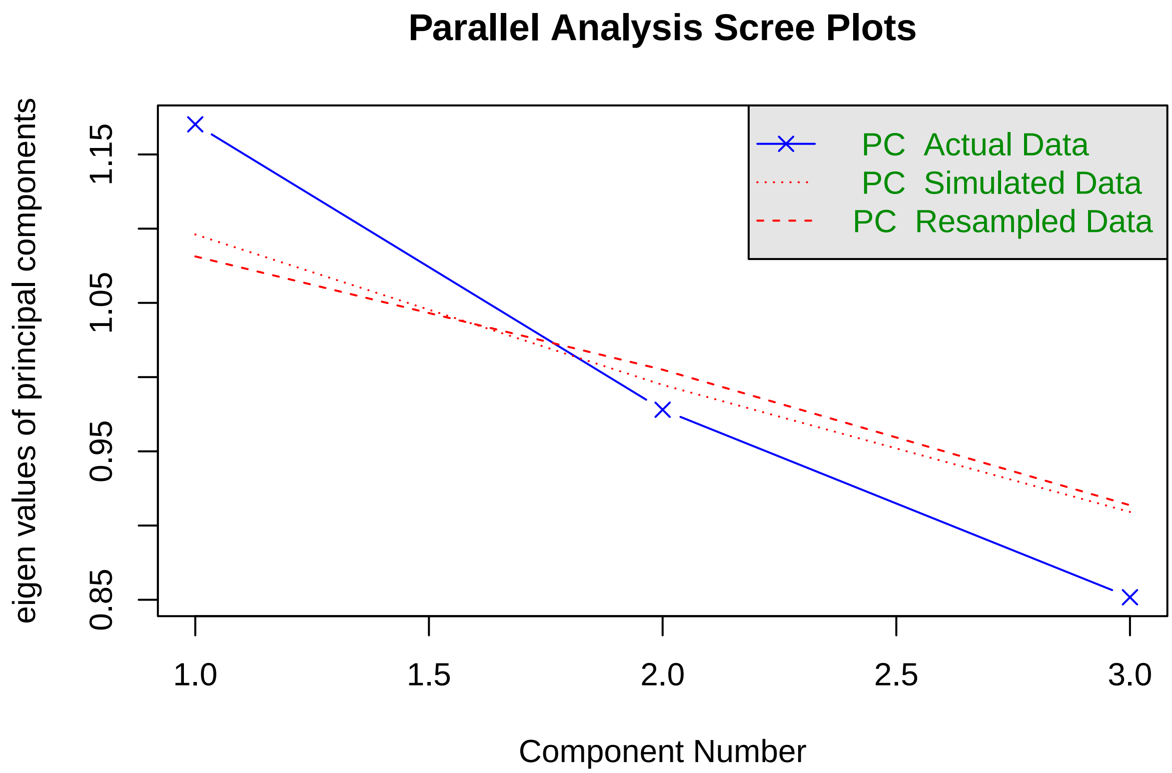
Latent Variable 2: Healthy Lifestyle

```
healthy_lifestyle <- c("hr_sleep", "smoking", "fruit_veg")

healthy_life_numeric <- athletes[,healthy_lifestyle] %>% map_df(., as.numeric)

healthy_life_matrix <- healthy_life_numeric %>% as.matrix()
```

Polychoric Correlations

Now, let us determine the number of factors that might underlie these variables.

```
health_parallel <- fa.parallel(healthy_life_matrix, fa = "pc")
```




```
## Parallel analysis suggests that the number of factors =  NA  and the number of components =  0
```

```
health_parallel$pc.values
```

```
## [1] 1.1702721 0.9780215 0.8517064
```

1 latent factor appears to underlie these variables. *** need to check if this is the correct way to assess formative LV's ***

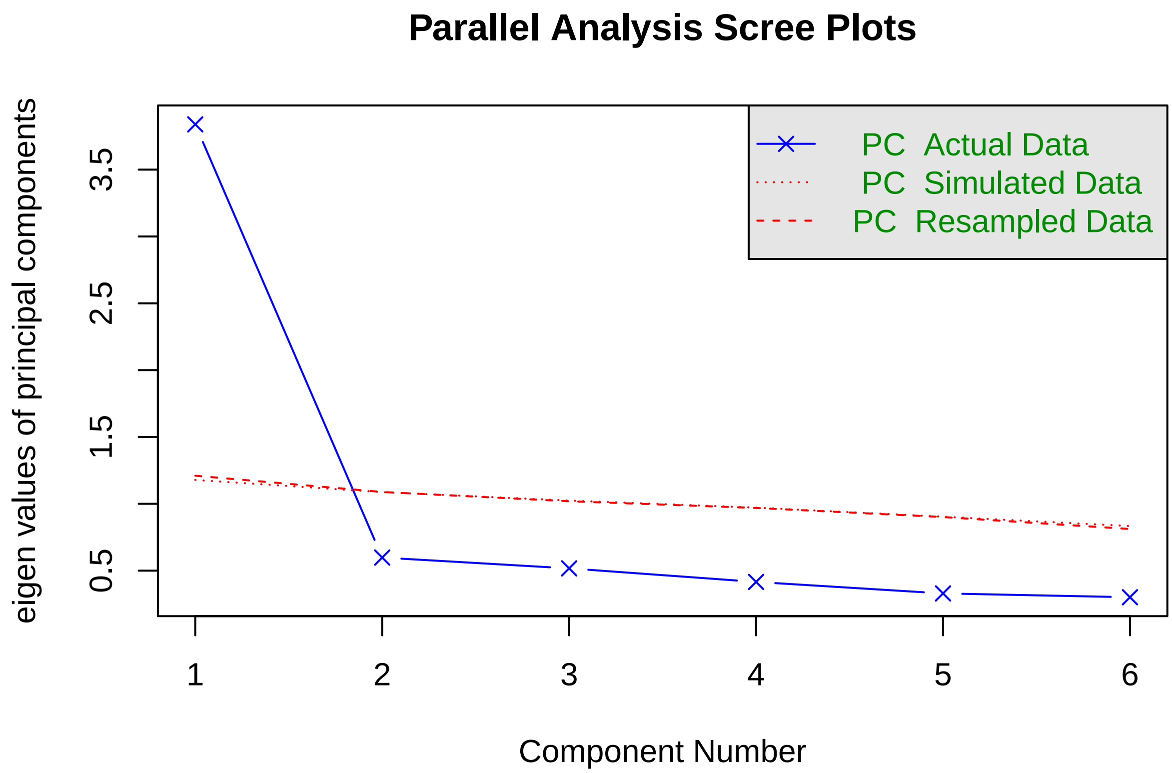
Latent Variable 3: Resilience

Finally, we will look at resilience.

```
resilience <- c("bounce", "strs_evnt", "strs_rcvr", "snap_back", "difficult", "setbacks")

resilience_numeric <- athletes[,resilience] %>% map_df(., as.numeric)

resilience_matrix <- resilience_numeric %>% as.matrix()
```

Polychoric Correlations

```
resilience_parallel <- fa.parallel(resilience_matrix, fa = "pc")
```




```
## Parallel analysis suggests that the number of factors =  NA  and the number of components =  1
```

```r
resilience_parallel$pc.values
```

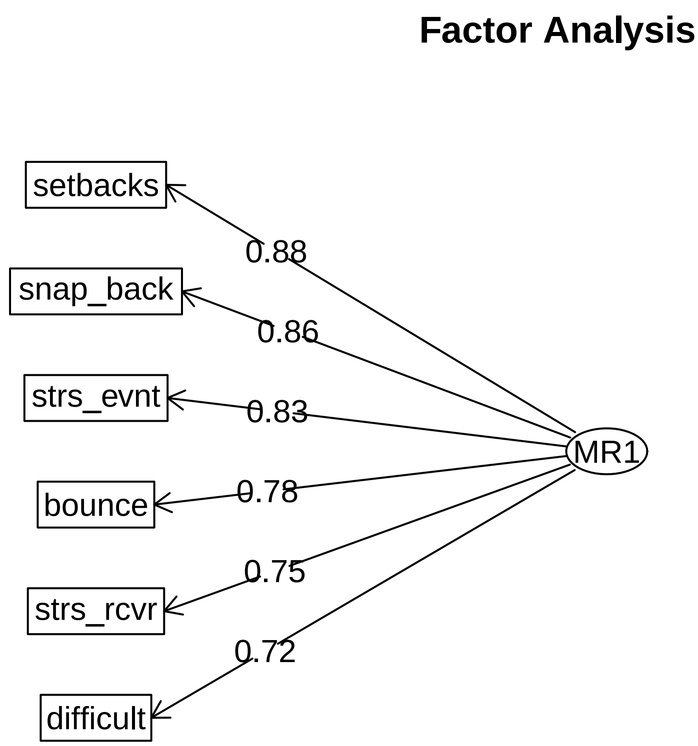
```
## [1] 3.8386245 0.5979661 0.5170194 0.4156602 0.3296245 0.3011053
```

It appears that a 1-factor model will sufficiently explain the variability across these variables.

EFA

First, we can do a 1-factor EFA model:

```r
resilience_efa1 <- fa(r = resilience_matrix, nfactors = 1, cor = "poly")
fa.diagram(resilience_efa1, digits = 2, simple = TRUE)
```

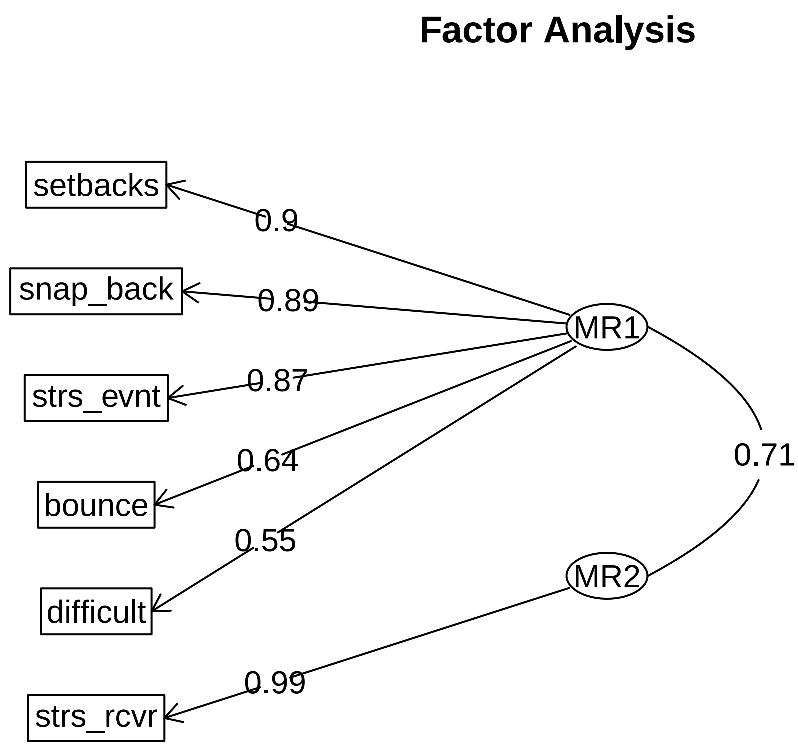



```
summary(resilience_efa1)
```

```
##
## Factor analysis with Call: fa(r = resilience_matrix, nfactors = 1, cor = "poly")
##
## Test of the hypothesis that 1 factor is sufficient.
## The degrees of freedom for the model is 9  and the objective function was  0.1
## The number of observations was  363  with Chi Square =  35.64  with prob <  4.6e-05
##
## The root mean square of the residuals (RMSA) is  0.03
## The df corrected root mean square of the residuals is  0.04
##
## Tucker Lewis Index of factoring reliability =  0.969
## RMSEA index =  0.09  and the 10 % confidence intervals are  0.06 0.123
## BIC =  -17.41
```

Now, we can try a 2-factor EFA model:

```
resilience_efa2 <- fa(r = resilience_matrix, nfactors = 2, cor = "poly")
fa.diagram(resilience_efa2, digits = 2, simple = TRUE)
```



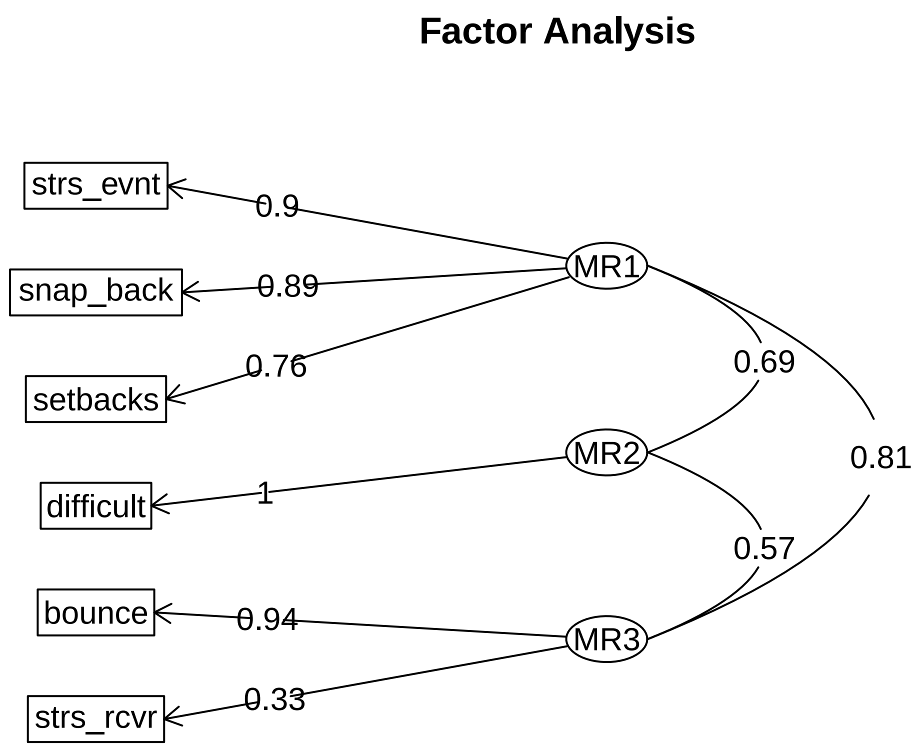

```
resilience_efa2
```

```
## Factor Analysis using method =  minres
## Call: fa(r = resilience_matrix, nfactors = 2, cor = "poly")
## Standardized loadings (pattern matrix) based upon correlation matrix
##            MR1   MR2   h2     u2 com
## bounce    0.64  0.18 0.60 0.4044 1.2
## strs_evnt 0.87 -0.04 0.71 0.2918 1.0
## strs_rcvr 0.01  0.99 1.00 0.0034 1.0
## snap_back 0.89 -0.02 0.76 0.2373 1.0
## difficult 0.55  0.20 0.50 0.5004 1.3
## setbacks  0.90 -0.02 0.79 0.2059 1.0
##
##                        MR1  MR2
## SS loadings           3.19 1.17
## Proportion Var        0.53 0.19
## Cumulative Var        0.53 0.73
## Proportion Explained  0.73 0.27
## Cumulative Proportion 0.73 1.00
##
##  With factor correlations of
##      MR1  MR2
## MR1 1.00 0.71
## MR2 0.71 1.00
##
## Mean item complexity =  1.1
```

```
## Test of the hypothesis that 2 factors are sufficient.
##
## The degrees of freedom for the null model are  15  and the objective function was  3.98 with Chi Squa
## The degrees of freedom for the model are 4  and the objective function was  0.02
##
## The root mean square of the residuals (RMSR) is  0.01
## The df corrected root mean square of the residuals is  0.02
##
## The harmonic number of observations is  322 with the empirical chi square  1.5  with prob <  0.83
## The total number of observations was  363  with Likelihood Chi Square =  8.56  with prob <  0.073
##
## Tucker Lewis Index of factoring reliability =  0.988
## RMSEA index =  0.056  and the 90 % confidence intervals are  0 0.109
## BIC =  -15.02
## Fit based upon off diagonal values = 1
## Measures of factor score adequacy
##                                                    MR1  MR2
## Correlation of (regression) scores with factors   0.96 1.00
## Multiple R square of scores with factors          0.93 1.00
## Minimum correlation of possible factor scores     0.85 0.99
```

Finally, we can try a 3-factor EFA model:

```
resilience_efa3 <- fa(r = resilience_matrix, nfactors = 3, cor = "poly")
fa.diagram(resilience_efa3, digits = 2, simple = TRUE)
```



```r
summary(resilience_efa3)
```

```
##
## Factor analysis with Call: fa(r = resilience_matrix, nfactors = 3, cor = "poly")
##
## Test of the hypothesis that 3 factors are sufficient.
## The degrees of freedom for the model is 0  and the objective function was  0
## The number of observations was  363  with Chi Square =  0.18  with prob <  NA
##
## The root mean square of the residuals (RMSA) is  0
## The df corrected root mean square of the residuals is  NA
##
## Tucker Lewis Index of factoring reliability =  -Inf
##  With factor correlations of
##      MR1  MR2  MR3
## MR1 1.00 0.69 0.81
## MR2 0.69 1.00 0.57
## MR3 0.81 0.57 1.00
```

Reliability

```r
psych::alpha(resilience_matrix)
```

```
##
## Reliability analysis
## Call: psych::alpha(x = resilience_matrix)
##
##   raw_alpha std.alpha G6(smc) average_r S/N    ase mean   sd median_r
##       0.89      0.89    0.87      0.57 7.8 0.0092  3.6 0.79     0.55
##
##  lower alpha upper     95% confidence boundaries
## 0.87 0.89 0.9
##
##  Reliability if an item is dropped:
##           raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r med.r
## bounce         0.87      0.87    0.85      0.57 6.7   0.0109 0.0054  0.54
## strs_evnt      0.86      0.86    0.84      0.56 6.3   0.0115 0.0043  0.55
## strs_rcvr      0.87      0.87    0.85      0.58 6.9   0.0104 0.0066  0.58
## snap_back      0.86      0.86    0.84      0.55 6.1   0.0117 0.0036  0.55
## difficult      0.88      0.88    0.86      0.59 7.3   0.0099 0.0038  0.58
## setbacks       0.85      0.85    0.83      0.54 5.9   0.0121 0.0036  0.53
##
##  Item statistics
##             n raw.r std.r r.cor r.drop mean   sd
## bounce    322  0.78  0.79  0.73   0.69  4.0 0.90
## strs_evnt 322  0.83  0.82  0.78   0.73  3.4 1.08
## strs_rcvr 322  0.77  0.77  0.70   0.66  3.6 0.99
## snap_back 322  0.83  0.83  0.79   0.74  3.5 1.01
## difficult 322  0.74  0.74  0.66   0.62  3.4 0.99
## setbacks  322  0.85  0.85  0.83   0.78  3.6 0.99
##
## Non missing response frequency for each item
##              1    2    3    4    5 miss
## bounce    0.02 0.06 0.15 0.49 0.28 0.11
```

```
## strs_evnt 0.05 0.18 0.29 0.34 0.14 0.11
## strs_rcvr 0.02 0.14 0.25 0.42 0.16 0.11
## snap_back 0.02 0.18 0.19 0.48 0.13 0.11
## difficult 0.02 0.19 0.28 0.39 0.11 0.11
## setbacks  0.03 0.13 0.23 0.47 0.15 0.11
```

Chronbach's alpha is 0.89 (0.88, 0.91). No items can be dropped to improve this measure, so we will keep all of them in this latent variable.

Final CFA

```r
final_model <-
'external_identity =~ sprt_goals + cnsdr_ath + frnds_ath
 internal_value =~ sprt_impt + think_sprt
 negative_events =~ dprs_sprt + bad_sprt

 athlete_identity =~ external_identity + internal_value + negative_events

 healthy_lifestyle =~ hr_sleep + smoking + fruit_veg

 resilience =~ bounce + strs_evnt + strs_rcvr + snap_back + difficult + setbacks
'
fin_df <- athletes %>% map_df(., as.numeric)

final_cfa = cfa(final_model,
                data = fin_df,
                std.lv = TRUE)
summary(final_cfa, fit.measures = TRUE)
```

```
## lavaan 0.6-10 ended normally after 50 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        38
##
##                                                   Used       Total
##   Number of observations                           322         363
##
## Model Test User Model:
##
##   Test statistic                               160.621
##   Degrees of freedom                                98
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              1723.967
##   Degrees of freedom                               120
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.961
##   Tucker-Lewis Index (TLI)                       0.952
```

```
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -6889.306
##   Loglikelihood unrestricted model (H1)      -6808.996
##
##   Akaike (AIC)                               13854.612
##   Bayesian (BIC)                             13998.045
##   Sample-size adjusted Bayesian (BIC)        13877.514
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.045
##   90 Percent confidence interval - lower         0.032
##   90 Percent confidence interval - upper         0.057
##   P-value RMSEA <= 0.05                          0.758
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.061
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                        Estimate  Std.Err  z-value  P(>|z|)
##   external_identity =~
##     sprt_goals           0.614    0.065    9.502    0.000
##     cnsdr_ath            0.597    0.062    9.619    0.000
##     frnds_ath            0.402    0.066    6.058    0.000
##   internal_value =~
##     sprt_impt            0.590    0.116    5.097    0.000
##     think_sprt           0.790    0.174    4.531    0.000
##   negative_events =~
##     dprs_sprt            0.634    0.082    7.714    0.000
##     bad_sprt             0.811    0.109    7.445    0.000
##   athlete_identity =~
##     external_dntty       0.895    0.162    5.515    0.000
##     internal_value       1.478    0.429    3.441    0.001
##     negative_evnts       0.770    0.148    5.211    0.000
##   healthy_lifestyle =~
##     hr_sleep             0.469    0.132    3.563    0.000
##     smoking             -0.274    0.103   -2.669    0.008
##     fruit_veg            0.117    0.047    2.471    0.013
##   resilience =~
##     bounce               0.661    0.045   14.723    0.000
##     strs_evnt            0.852    0.052   16.417    0.000
##     strs_rcvr            0.676    0.051   13.350    0.000
##     snap_back            0.812    0.048   16.971    0.000
##     difficult            0.644    0.051   12.575    0.000
##     setbacks             0.832    0.046   18.099    0.000
```

```
##
## Covariances:
##                      Estimate  Std.Err  z-value  P(>|z|)
##   athlete_identity ~~
##     healthy_lfstyl       0.114    0.124    0.915    0.360
##     resilience          -0.140    0.069   -2.026    0.043
##   healthy_lifestyle ~~
##     resilience           0.398    0.119    3.328    0.001
##
## Variances:
##                  Estimate  Std.Err  z-value  P(>|z|)
##    .sprt_goals       0.230    0.061    3.749    0.000
##    .cnsdr_ath        0.618    0.074    8.305    0.000
##    .frnds_ath        1.574    0.130   12.135    0.000
##    .sprt_impt        1.178    0.137    8.591    0.000
##    .think_sprt       0.281    0.182    1.538    0.124
##    .dprs_sprt        0.963    0.130    7.417    0.000
##    .bad_sprt         1.010    0.190    5.330    0.000
##    .hr_sleep         0.809    0.130    6.236    0.000
##    .smoking          1.058    0.095   11.099    0.000
##    .fruit_veg        0.234    0.021   11.384    0.000
##    .bounce           0.377    0.034   10.998    0.000
##    .strs_evnt        0.434    0.042   10.234    0.000
##    .strs_rcvr        0.527    0.046   11.430    0.000
##    .snap_back        0.348    0.035    9.900    0.000
##    .difficult        0.564    0.049   11.624    0.000
##    .setbacks         0.279    0.031    9.041    0.000
##    .external_dntty   1.000
##    .internal_value   1.000
##    .negative_evnts   1.000
##     athlete_idntty   1.000
##     healthy_lfstyl   1.000
##     resilience       1.000
```

Modification Indices

```
modindices(final_cfa, power = TRUE, sort = TRUE, minimum.value = 10)
```

```
##  [1] lhs      op       rhs      mi       epc      sepc.all delta    ncp
##  [9] power    decision
## <0 rows> (or 0-length row.names)
```

Structural Equation Modeling

```
athlete_sem <- '
  # measurement model
    external_identity =~ sprt_goals + cnsdr_ath + frnds_ath
    internal_value =~ sprt_impt + think_sprt
    negative_events =~ dprs_sprt + bad_sprt

    athlete_identity =~ external_identity + internal_value + negative_events

    healthy_lifestyle =~ hr_sleep + smoking + fruit_veg
```

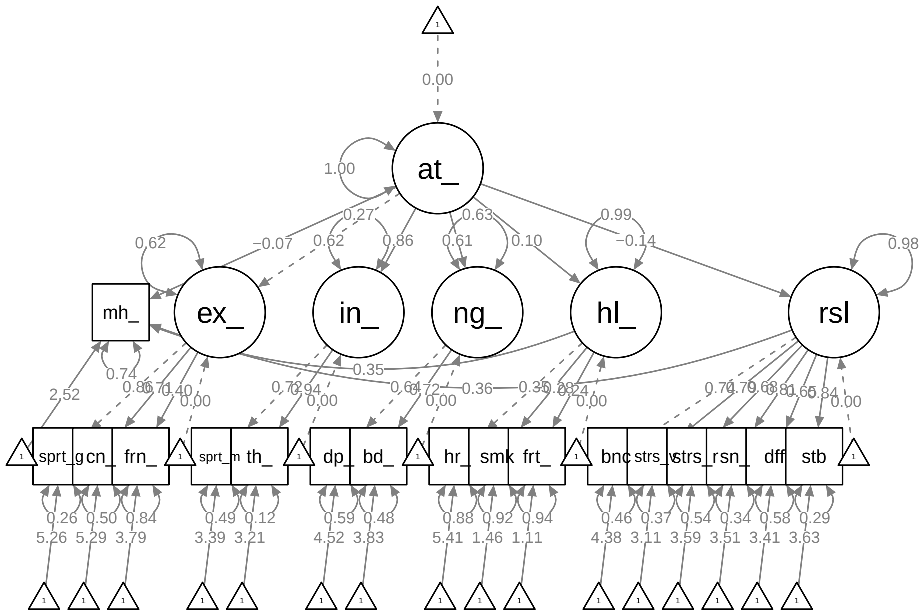
```r
    resilience =~ bounce + strs_evnt + strs_rcvr + snap_back + difficult + setbacks

  # structural model - direct effects
    mhc_sf ~ a*athlete_identity + b*healthy_lifestyle + c*resilience
    resilience ~ d*athlete_identity
    healthy_lifestyle ~ e*athlete_identity

  # indirect
    indirect_athlete_identity := d*c

  # total
    total_athlete_identity:= d*c + a
'

athlete_sem_fit <- sem(athlete_sem,
                       data = athletes,
                       sample.cov = TRUE,
                       missing = "ML")
summary(athlete_sem_fit)
```

```
## lavaan 0.6-10 ended normally after 139 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        58
##
##   Number of observations                           363
##   Number of missing patterns                         4
##
## Model Test User Model:
##
##   Test statistic                               202.637
##   Degrees of freedom                               112
##   P-value (Chi-square)                           0.000
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Observed
##   Observed information based on                Hessian
##
## Latent Variables:
##                        Estimate  Std.Err  z-value  P(>|z|)
##   external_identity =~
##     sprt_goals            1.000
##     cnsdr_ath             0.947    0.096    9.848    0.000
##     frnds_ath             0.661    0.105    6.270    0.000
##   internal_value =~
##     sprt_impt             1.000
##     think_sprt            1.330    0.132   10.110    0.000
##   negative_events =~
##     dprs_sprt             1.000
##     bad_sprt              1.296    0.208    6.240    0.000
```

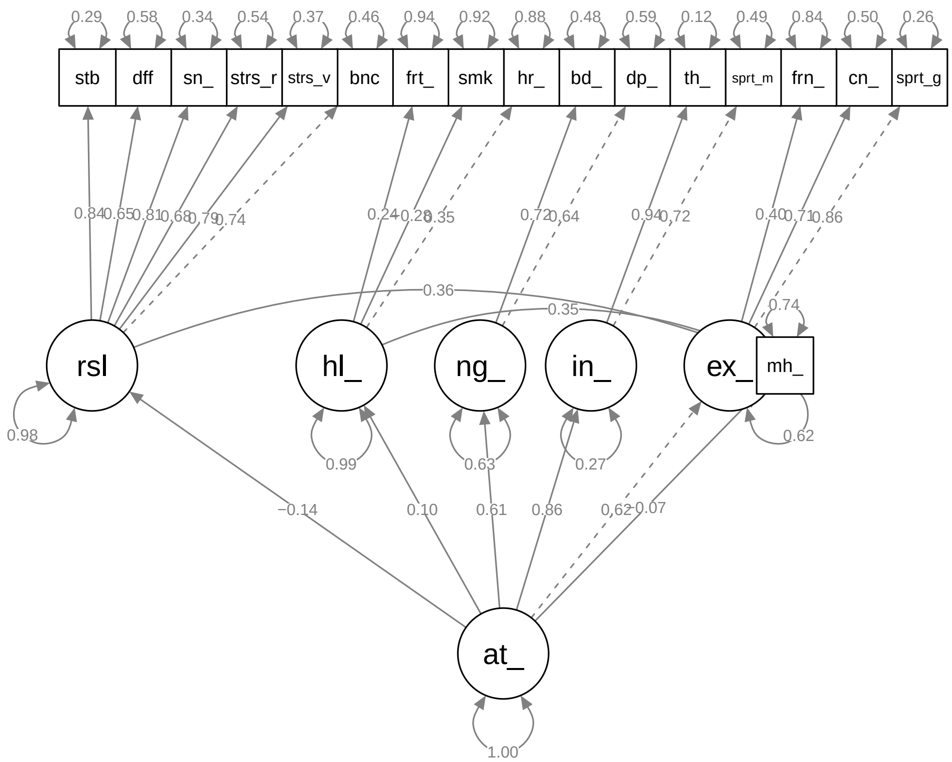
```
##   athlete_identity =~
##     external_dntty       1.000
##     internal_value       1.799    0.367    4.897    0.000
##     negative_evnts       0.951    0.176    5.415    0.000
##   healthy_lifestyle =~
##     hr_sleep             1.000
##     smoking             -0.408    0.212   -1.927    0.054
##     fruit_veg            0.159    0.107    1.484    0.138
##   resilience =~
##     bounce               1.000
##     strs_evnt            1.286    0.094   13.740    0.000
##     strs_rcvr            1.017    0.085   11.918    0.000
##     snap_back            1.227    0.087   14.137    0.000
##     difficult            0.968    0.086   11.208    0.000
##     setbacks             1.251    0.085   14.704    0.000
##
## Regressions:
##                      Estimate  Std.Err  z-value  P(>|z|)
##   mhc_sf ~
##     athlt_dntt (a)     -1.679    1.822   -0.921    0.357
##     hlthy_lfst (b)      5.910    3.576    1.653    0.098
##     resilience (c)      6.963    1.140    6.105    0.000
##   resilience ~
##     athlt_dntt (d)     -0.177    0.096   -1.848    0.065
##   healthy_lifestyle ~
##     athlt_dntt (e)      0.144    0.211    0.683    0.495
##
## Intercepts:
##                   Estimate  Std.Err  z-value  P(>|z|)
##    .sprt_goals       5.067    0.051   99.244    0.000
##    .cnsdr_ath        5.876    0.059   99.891    0.000
##    .frnds_ath        5.225    0.073   71.430    0.000
##    .sprt_impt        5.109    0.080   63.958    0.000
##    .think_sprt       4.904    0.081   60.569    0.000
##    .dprs_sprt        5.666    0.066   85.290    0.000
##    .bad_sprt         5.505    0.076   72.194    0.000
##    .hr_sleep        11.543    0.112  103.085    0.000
##    .smoking          1.565    0.056   27.741    0.000
##    .fruit_veg        0.554    0.026   21.222    0.000
##    .bounce           3.953    0.050   78.679    0.000
##    .strs_evnt        3.351    0.060   55.906    0.000
##    .strs_rcvr        3.565    0.055   64.519    0.000
##    .snap_back        3.522    0.056   63.042    0.000
##    .difficult        3.373    0.055   61.209    0.000
##    .setbacks         3.578    0.055   65.245    0.000
##    .mhc_sf          32.084    0.696   46.094    0.000
##    .external_dntty   0.000
##    .internal_value   0.000
##    .negative_evnts   0.000
##     athlete_idntty   0.000
##    .healthy_lfstyl   0.000
##    .resilience       0.000
##
## Variances:
```

```
##                     Estimate  Std.Err  z-value  P(>|z|)
##    .sprt_goals         0.239    0.062    3.832    0.000
##    .cnsdr_ath          0.614    0.071    8.684    0.000
##    .frnds_ath          1.603    0.127   12.616    0.000
##    .sprt_impt          1.104    0.129    8.584    0.000
##    .think_sprt         0.269    0.175    1.537    0.124
##    .dprs_sprt          0.931    0.119    7.815    0.000
##    .bad_sprt           0.996    0.179    5.578    0.000
##    .hr_sleep           3.989    0.472    8.448    0.000
##    .smoking            1.061    0.104   10.174    0.000
##    .fruit_veg          0.233    0.021   11.175    0.000
##    .bounce             0.375    0.034   10.969    0.000
##    .strs_evnt          0.432    0.042   10.232    0.000
##    .strs_rcvr          0.530    0.046   11.404    0.000
##    .snap_back          0.346    0.035    9.898    0.000
##    .difficult          0.567    0.049   11.624    0.000
##    .setbacks           0.282    0.031    9.115    0.000
##    .mhc_sf           120.073   16.487    7.283    0.000
##    .external_dntty     0.425    0.076    5.569    0.000
##    .internal_value     0.312    0.155    2.009    0.045
##    .negative_evnts     0.401    0.099    4.064    0.000
##     athlete_idntty     0.264    0.068    3.887    0.000
##    .healthy_lfstyl     0.557    0.399    1.397    0.162
##    .resilience         0.432    0.059    7.334    0.000
##
## Defined Parameters:
##                     Estimate  Std.Err  z-value  P(>|z|)
##     indrct_thlt_dn    -1.233    0.686   -1.797    0.072
##     ttl_thlt_dntty    -2.912    1.887   -1.543    0.123
```

```
semPaths(athlete_sem_fit,
         whatLabels = "std",
         reorder = FALSE,
         layout = "tree2")
```




```r
# graph looks cleaner
semPaths(athlete_sem_fit,
         what = "paths",
         whatLabels = "std",
         reorder = FALSE,
         layout = "tree2",
         rotation = 3,
         intercepts = FALSE)
```



Double headed arrow between healthy lifestyle and athlete identity means that they may share a common cause and their direct causal relationship is unspecified. We hypothesize that athlete identity causes healthy lifestyle. We can change the double headed arrow to direct causal arrow.

Interpretation

From our SEM, we observe that athlete identity *negatively* affects emotional well being, as defined by the MHC-SF scale. Resilience is associated with a positive effect on MHC-SF. Healthy lifestyle habits are also associated with a positive effect on MHC-SF.

Athletic identity is positively associated with resilience.