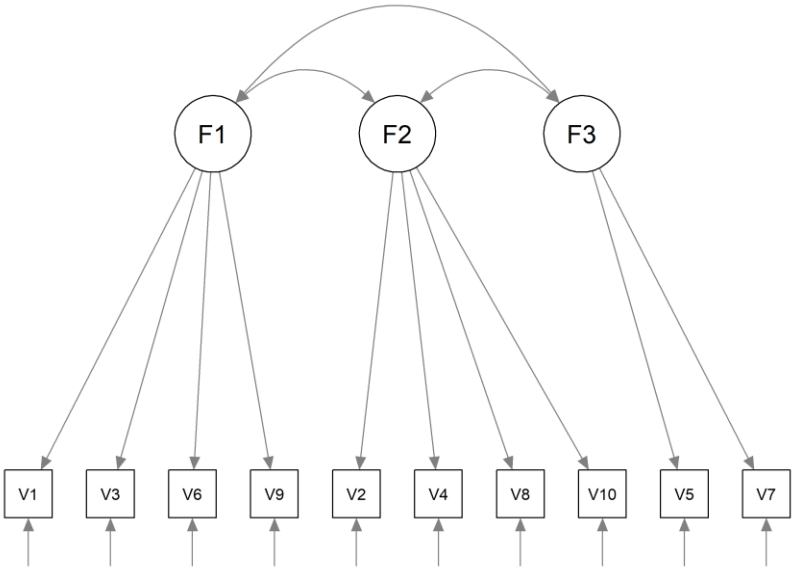
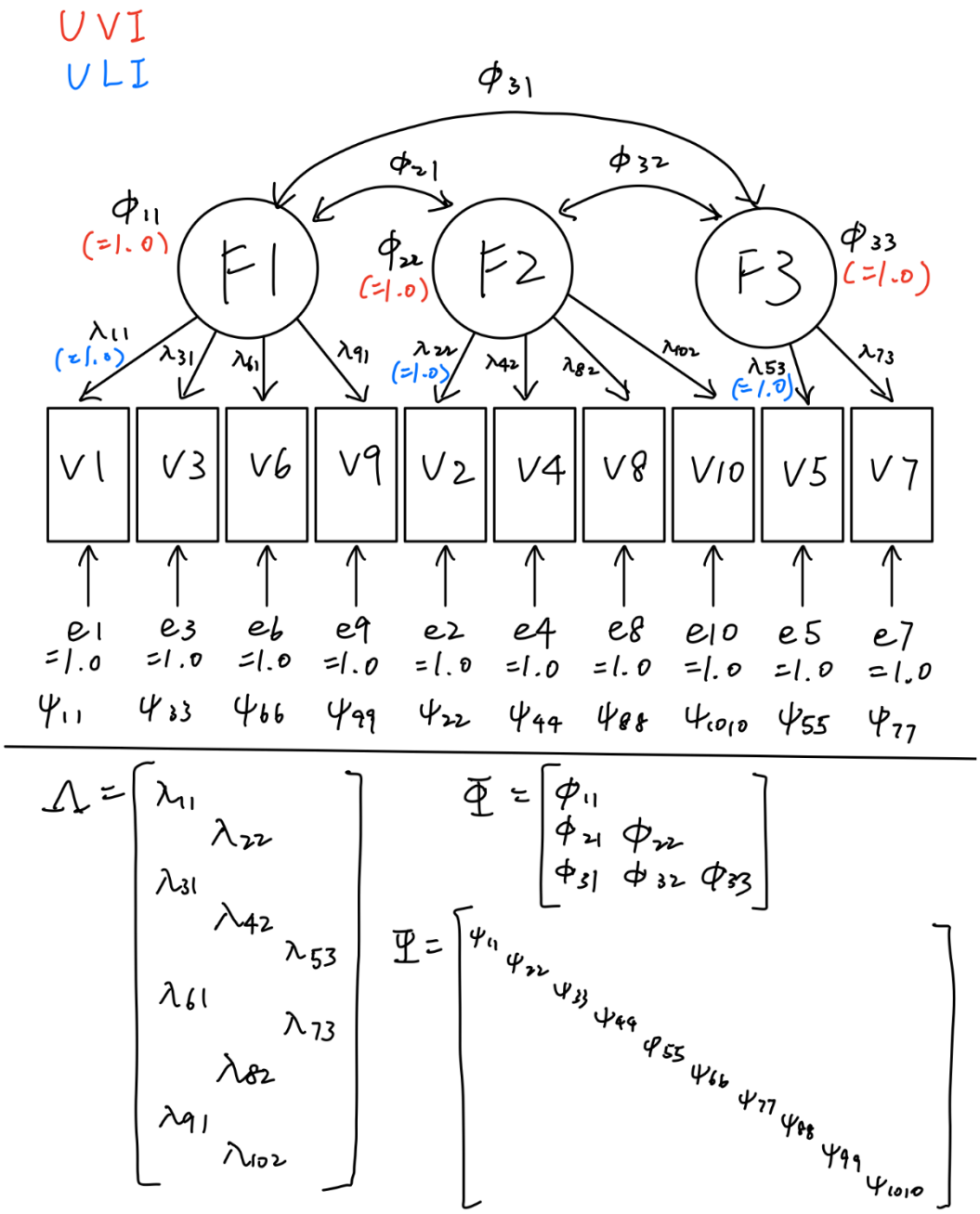
(a) We can generate the path diagram by R, or by hand drawing:

```
# data input
datafile <- read_table("hw4(2020).dat", header=FALSE)

# model specification
CFA.model <-'
  F1 =~ V1 + V3 + V6 + V9
  F2 =~ V2 + V4 + V8 + V10
  F3 =~ V5 + V7
'
#draw diagram (without estimates)
semPaths(fit.UVI,"model", style="lisrel")
```



Hand drawing:



$$\Lambda = \begin{bmatrix} \lambda_{11} & & \\ & \lambda_{22} & \\ \lambda_{31} & & \\ & \lambda_{42} & \\ & & \lambda_{53} \\ \lambda_{61} & & \\ & & \lambda_{73} \\ & \lambda_{82} & \\ \lambda_{91} & & \\ & \lambda_{102} & \end{bmatrix} \qquad \Phi = \begin{bmatrix} \phi_{11} & & \\ \phi_{21} & \phi_{22} & \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}$$

$$\Psi = \begin{bmatrix} \psi_{11} & & & & & & & & & \\ & \psi_{22} & & & & & & & & \\ & & \psi_{33} & & & & & & & \\ & & & \psi_{44} & & & & & & \\ & & & & \psi_{55} & & & & & \\ & & & & & \psi_{66} & & & & \\ & & & & & & \psi_{77} & & & \\ & & & & & & & \psi_{88} & & \\ & & & & & & & & \psi_{99} & \\ & & & & & & & & & \psi_{1010} \end{bmatrix}$$

Fixed parameters:

- ULI: first factor loadings under each factor are fixed at 1.0
- UVI: factor variances are fixed at 1.0

The rest are free parameters.

- ULI: factor variances, factor covariances, factor loadings other than the first factor and error variances

- UVI: factor covariances, all factor loadings, error variances

The number of available pieces of information (p*=(10)(11)/2=55) in the above CFA model is larger than the number of free parameters (q=23). As a result, the necessary condition (the t-rule) for model identification is satisfied.

We have 3 factors with at least 2 indicators per factor in the model; the factor correlation is free; each indicator loads on one factor and the error variances are uncorrelated. Hence, the two-indicator rule is satisfied.

Hence, the proposed model can be identified.

(b)

Number of free parameters:

- ULI:
  7 factor loadings + 10 error variances + 3 factor variances + 3 factor covariances = 23
- UVI:
  10 factor loadings + 10 error variances + 3 factor covariances = 23

Degrees of freedom = p*-q = 10(10+1)/2 – 23 =55-23 = 32

(c)

(i) UVI

```
# fit model using UVI

fit.UVI <- cfa(CFA.model, data=datafile, std.lv=TRUE)
summary(fit.UVI, standardized=TRUE, fit.measures=TRUE)
## lavaan 0.6-5 ended normally after 25 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         23
##
##   Number of observations                           222
##
## Model Test User Model:
##
##   Test statistic                               232.378
##   Degrees of freedom                                32
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
```

```
##
##   Test statistic                                956.825
##   Degrees of freedom                             45
##   P-value                                       0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.780
##   Tucker-Lewis Index (TLI)                       0.691
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -2398.615
##   Loglikelihood unrestricted model (H1)      -2282.426
##
##   Akaike (AIC)                               4843.230
##   Bayesian (BIC)                             4921.492
##   Sample-size adjusted Bayesian (BIC)        4848.603
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                         0.168
##   90 Percent confidence interval - lower        0.148
##   90 Percent confidence interval - upper        0.189
##   P-value RMSEA <= 0.05                         0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                          0.092
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard errors                             Standard
##
## Latent Variables:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   F1 =~
##     V1               0.534    0.043   12.500    0.000    0.534    0.758
##     V3               0.749    0.064   11.636    0.000    0.749    0.718
##     V6               0.528    0.053   10.026    0.000    0.528    0.639
##     V9               0.818    0.052   15.619    0.000    0.818    0.892
##   F2 =~
##     V2               0.665    0.059   11.203    0.000    0.665    0.729
##     V4               0.658    0.047   14.078    0.000    0.658    0.889
##     V8               0.330    0.045    7.356    0.000    0.330    0.503
##     V10              0.410    0.056    7.293    0.000    0.410    0.499
##   F3 =~
##     V5               0.702    0.068   10.322    0.000    0.702    0.828
##     V7               0.695    0.077    9.016    0.000    0.695    0.689
##
## Covariances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
```

```
##   F1 ~~
##     F2                0.081    0.078    1.040    0.298    0.081    0.081
##     F3                0.384    0.074    5.182    0.000    0.384    0.384
##   F2 ~~
##     F3                0.449    0.073    6.178    0.000    0.449    0.449
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .V1                0.211    0.026    8.190    0.000    0.211    0.426
##    .V3                0.528    0.060    8.759    0.000    0.528    0.485
##    .V6                0.404    0.043    9.433    0.000    0.404    0.592
##    .V9                0.172    0.039    4.454    0.000    0.172    0.205
##    .V2                0.391    0.052    7.479    0.000    0.391    0.469
##    .V4                0.114    0.036    3.150    0.002    0.114    0.209
##    .V8                0.321    0.033    9.837    0.000    0.321    0.747
##    .V10               0.508    0.052    9.853    0.000    0.508    0.751
##    .V5                0.225    0.073    3.072    0.002    0.225    0.314
##    .V7                0.534    0.086    6.246    0.000    0.534    0.525
##     F1                1.000                              1.000    1.000
##     F2                1.000                              1.000    1.000
##     F3                1.000                              1.000    1.000
```

(ii) ULI

```
# fit model using ULI
#
fit.ULI <- cfa(CFA.model, data=datafile)
summary(fit.ULI, standardized=TRUE, fit.measures=TRUE)
## lavaan 0.6-5 ended normally after 35 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         23
##
##   Number of observations                           222
##
## Model Test User Model:
##
##   Test statistic                               232.378
##   Degrees of freedom                                32
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                               956.825
##   Degrees of freedom                                45
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.780
##   Tucker-Lewis Index (TLI)                       0.691
##
## Loglikelihood and Information Criteria:
```

```
##
##   Loglikelihood user model (H0)              -2398.615
##   Loglikelihood unrestricted model (H1)      -2282.426
##
##   Akaike (AIC)                               4843.230
##   Bayesian (BIC)                             4921.492
##   Sample-size adjusted Bayesian (BIC)        4848.603
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.168
##   90 Percent confidence interval - lower         0.148
##   90 Percent confidence interval - upper         0.189
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.092
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard errors                             Standard
##
## Latent Variables:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   F1 =~
##     V1               1.000                               0.534    0.758
##     V3               1.402    0.134   10.434    0.000    0.749    0.718
##     V6               0.989    0.107    9.221    0.000    0.528    0.639
##     V9               1.533    0.126   12.196    0.000    0.818    0.892
##   F2 =~
##     V2               1.000                               0.665    0.729
##     V4               0.989    0.103    9.607    0.000    0.658    0.889
##     V8               0.496    0.072    6.881    0.000    0.330    0.503
##     V10              0.617    0.090    6.829    0.000    0.410    0.499
##   F3 =~
##     V5               1.000                             0.702    0.828
##     V7               0.991    0.157    6.318    0.000    0.695    0.689
##
## Covariances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   F1 ~~
##     F2               0.029    0.028    1.027    0.305    0.081    0.081
##     F3               0.144    0.034    4.204    0.000    0.384    0.384
##   F2 ~~
##     F3               0.210    0.046    4.586    0.000    0.449    0.449
##
## Variances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .V1               0.211    0.026    8.190    0.000    0.211    0.426
##    .V3               0.528    0.060    8.759    0.000    0.528    0.485
##    .V6               0.404    0.043    9.433    0.000    0.404    0.592
```

```
##    .V9               0.172    0.039    4.454    0.000    0.172    0.205
##    .V2               0.391    0.052    7.479    0.000    0.391    0.469
##    .V4               0.114    0.036    3.150    0.002    0.114    0.209
##    .V8               0.321    0.033    9.837    0.000    0.321    0.747
##    .V10              0.508    0.052    9.853    0.000    0.508    0.751
##    .V5               0.225    0.073    3.072    0.002    0.225    0.314
##    .V7               0.534    0.086    6.246    0.000    0.534    0.525
##     F1               0.285    0.046    6.250    0.000    1.000    1.000
##     F2               0.442    0.079    5.601    0.000    1.000    1.000
##     F3               0.492    0.095    5.161    0.000    1.000    1.000
```

Fit measures using UVI:

```
> fitMeasures(fit.UVI,fit.measures=c("chisq", "df", "pvalue", "nfi","nnfi","cfi
","gfi","agfi","srmr","rmsea"))
  chisq      df  pvalue   nfi    nnfi     cfi     gfi    agfi    srmr   rmsea
232.378  32.000   0.000  0.757   0.691   0.780   0.849   0.741   0.092   0.168
```

Fit measures using ULI:

```
> fitMeasures(fit.ULI,fit.measures=c("chisq", "df", "pvalue","nfi","nnfi","cfi",
"gfi","agfi","srmr","rmsea"))
  chisq      df  pvalue   nfi    nnfi     cfi     gfi    agfi    srmr   rmsea
232.378  32.000   0.000  0.757   0.691   0.780   0.849   0.741   0.092   0.168
```

UVI and ULI differ in parameter estimations but do not differ in terms of goodness-of-fit as they are fitting essentially the same model.

Parameter estimations:

- ULI: the estimates of the first variables under each factor are set at 1.0 while other loadings are free to vary. Factor variances and covariances are free to vary.
- UVI: Factor variances are set at 1.0. All factor loadings and factor covariances are free to vary.
- Factor loadings and factor covariances are different across the two methods.
- Error variances are the same across the two methods.

(d)

The Chi-square test is significant, $X^2(32)=232.38$, p=.00, implying that the reproduced covariance matrix is significantly different form the ideal covariance matrix, indicating unsatisfactory goodness-of-fit. However, as the Chi-square is very stringent, rejection is likely to happen. Goodness-of-fit indices would be provided as complementary indicators of the model goodness-of-fit.

NFI=.76<.90, NNFI=.69<.90, CFI=.78<.90, AGFI=.74<.90, SRMR=.092>.05, RMSEA=.17>.08, suggesting that the model does not provide an adequate fit for the data.

(e)

```
# MI
mi <- modindices(fit.UVI, sort. = TRUE)

# show useful paths
mi[mi$mi >= 3.841, ]
##    lhs op rhs     mi    epc sepc.lv sepc.all sepc.nox
## 48  V1 ~~  V6 56.721  0.195   0.195    0.668    0.668
## 57  V3 ~~  V9 45.750  0.360   0.360    1.193    1.193
## 89 V10 ~~  V5 25.264  0.165   0.165    0.488    0.488
## 90 V10 ~~  V7 23.206 -0.194  -0.194   -0.373   -0.373
## 65  V6 ~~  V2 22.734  0.146   0.146    0.367    0.367
## 69  V6 ~~  V5 19.658 -0.132  -0.132   -0.438   -0.438
## 47  V1 ~~  V3 19.427 -0.151  -0.151   -0.452   -0.452
## 64  V6 ~~  V9 18.229 -0.158  -0.158   -0.599   -0.599
## 43  F3 =~  V2 16.078  0.270   0.270    0.296    0.296
## 49  V1 ~~  V9 11.879 -0.135  -0.135   -0.708   -0.708
## 27  F1 =~  V2 11.292  0.170   0.170    0.186    0.186
## 31  F1 =~  V5  8.728 -0.255  -0.255   -0.301   -0.301
## 38  F2 =~  V7  8.728 -0.304  -0.304   -0.302   -0.302
## 32  F1 =~  V7  8.728  0.253   0.253    0.251    0.251
## 37  F2 =~  V5  8.728  0.307   0.307    0.362    0.362
## 53  V1 ~~ V10  8.560  0.072   0.072    0.221    0.221
## 41  F3 =~  V6  7.622 -0.158  -0.158   -0.192   -0.192
## 74  V9 ~~ V10  7.186 -0.076  -0.076   -0.256   -0.256
## 62  V3 ~~  V5  7.169 -0.094  -0.094   -0.273   -0.273
## 75  V9 ~~  V5  7.070  0.073   0.073    0.369    0.369
## 44  F3 =~  V4  6.781 -0.155  -0.155   -0.210   -0.210
## 29  F1 =~  V8  6.384 -0.106  -0.106   -0.162   -0.162
## 68  V6 ~~ V10  5.361 -0.075  -0.075   -0.167   -0.167
## 70  V6 ~~  V7  5.262  0.084   0.084    0.181    0.181
## 60  V3 ~~  V8  4.927 -0.068  -0.068   -0.164   -0.164
## 30  F1 =~ V10  4.570 -0.113  -0.113   -0.138   -0.138
## 73  V9 ~~  V8  3.891  0.044   0.044    0.188    0.188
## 56  V3 ~~  V6  3.848 -0.075  -0.075   -0.163   -0.163
```

The error covariance path between V1 and V6 will be added as the mi value is the highest. Since V1 and V6 load on the same factor, it may be reasonable to allow their error terms to correlate.

```
# modified CFA model (V1~~V6)
CFA.model.V1V6 <-'
  F1 =~ V1 + V3 + V6 + V9
  F2 =~ V2 + V4 + V8 + V10
  F3 =~ V5 + V7
  V1~~V6
'
fit.V1V6 <- cfa(CFA.model.V1V6, data=datafile)
summary(fit.V1V6, standardized=TRUE, fit.measures=TRUE)
## lavaan 0.6-5 ended normally after 36 iterations
##
##   Estimator                                         ML
```

```
##   Optimization method                           NLMINB
##   Number of free parameters                        24
##
##   Number of observations                          222
##
## Model Test User Model:
##
##   Test statistic                               177.096
##   Degrees of freedom                            31
##   P-value (Chi-square)                          0.000
##
## Model Test Baseline Model:
##
##   Test statistic                               956.825
##   Degrees of freedom                            45
##   P-value                                       0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.840
##   Tucker-Lewis Index (TLI)                       0.767
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -2370.974
##   Loglikelihood unrestricted model (H1)      -2282.426
##
##   Akaike (AIC)                               4789.948
##   Bayesian (BIC)                             4871.613
##   Sample-size adjusted Bayesian (BIC)         4795.555
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.146
##   90 Percent confidence interval - lower         0.125
##   90 Percent confidence interval - upper         0.167
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.087
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard errors                             Standard
##
## Latent Variables:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   F1 =~
##     V1               1.000                               0.464    0.658
##     V3               1.572    0.163    9.644    0.000    0.729    0.699
##     V6               0.937    0.090   10.422    0.000    0.434    0.526
```

```
##     V9                1.966    0.203    9.680    0.000    0.911    0.993
##   F2 =~
##     V2                1.000                               0.668    0.732
##     V4                0.979    0.101    9.666    0.000    0.654    0.884
##     V8                0.493    0.072    6.864    0.000    0.329    0.502
##     V10               0.618    0.090    6.869    0.000    0.413    0.502
##   F3 =~
##     V5                1.000                               0.740    0.873
##     V7                0.892    0.146    6.123    0.000    0.660    0.654
##
## Covariances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##  .V1 ~~
##    .V6               0.192    0.031    6.111    0.000    0.192    0.516
##   F1 ~~
##     F2               0.025    0.023    1.076    0.282    0.080    0.080
##     F3               0.122    0.030    4.061    0.000    0.357    0.357
##   F2 ~~
##     F3               0.224    0.047    4.773    0.000    0.453    0.453
##
## Variances:
##                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .V1               0.281    0.031    9.194    0.000    0.281    0.567
##    .V3               0.556    0.064    8.643    0.000    0.556    0.512
##    .V6               0.494    0.049   10.123    0.000    0.494    0.724
##    .V9               0.011    0.057    0.197    0.844    0.011    0.013
##    .V2               0.387    0.052    7.431    0.000    0.387    0.464
##    .V4               0.119    0.036    3.324    0.001    0.119    0.218
##    .V8               0.322    0.033    9.832    0.000    0.322    0.748
##    .V10              0.506    0.051    9.831    0.000    0.506    0.748
##    .V5               0.170    0.081    2.101    0.036    0.170    0.238
##    .V7               0.583    0.084    6.937    0.000    0.583    0.573
##     F1               0.215    0.042    5.171    0.000    1.000    1.000
##     F2               0.446    0.079    5.636    0.000    1.000    1.000
##     F3               0.547    0.103    5.290    0.000    1.000    1.000
```

The fit of the new model can be compared to that of the old model using an LR test as they are nested models.

```
# LR test
lavTestLRT(fit.ULI,fit.V1V6)
## Chi-Squared Difference Test
##
##         Df    AIC    BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## fit.V1V6 31 4789.9 4871.6 177.10
## fit.ULI  32 4843.2 4921.5 232.38     55.282       1  1.044e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Chi-square change = 55.28, df change =1, p<.001, indicating that the new model provides a better fit compared to the original model. The LR test shows how the revised model significantly improves the model fit.

Goodness-of-fit of the new model:

```
> fitMeasures(fit.V1V6,fit.measures=c("chisq", "df","pvalue", "nfi","nnfi","cfi
","gfi","agfi","srmr","rmsea"))
  chisq      df  pvalue     nfi    nnfi     cfi     gfi    agfi    srmr   rmsea
177.096  31.000   0.000   0.815   0.767   0.840   0.873   0.774   0.087   0.146
```

After adding the path V1~~V6, the Chi-square test is still significant, $X^2(31)=177.10$, p=.00. Goodness-of-fit indices are slightly improved but still suggest an inadequate fit.