LMX_mediation_model

Allama Ikbal Sijan

2023-09-22

Installing and Loading Packages

```
chooseCRANmirror(ind = 1)
install.packages("tidyverse")
install.packages("psych")
install.packages("ggplot2")
install.packages("lavaan")
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.3     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.3     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(psych)
```

```
##
## Attaching package: 'psych'
##
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```
library(ggplot2)
library(lavaan)
```

```
## This is lavaan 0.6-16
## lavaan is FREE software! Please report any bugs.
##
## Attaching package: 'lavaan'
##
## The following object is masked from 'package:psych':
##
##     cor2cov
```

Loading the dataset

```
df <- read_csv("GenderMatchData.csv")
```

```
## Rows: 808 Columns: 59
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (4): otenure_l, ttenure_l, otenure, ttenure
## dbl (55): tmkey, tlkey, office, gender, age, gender_l, office_l, age_l, lc1,...
##
## 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.
```

Data Cleaning

```
df <- df[ , c("tmkey", "tlkey", "gender", "age", "gender_l", "age_l",
              "lc1", "lc2", "lc3", "lc4", "lc5", "lc6", "lc7", "lc8",
              "lmx1", "lmx2", "lmx3", "lmx4", "lmx5", "lmx6", "lmx7",
              "as1", "as2", "as3", "as4", "as5", "as6", "as7")]

df <- df %>%
  mutate(mean_lc = (lc1 + lc2 + lc3 + lc4 + lc5 + lc6 + lc7 + lc8) / 8,
         mean_lmx = (lmx1 + lmx2 + lmx3 + lmx4 + lmx5 + lmx6 + lmx7) / 7,
         mean_as = (as1 + as2 + as3 + as4 + as5 + as6 + as7) / 7)

sum(is.na(df$mean_lmx))
```

```
## [1] 223
```

```
sum(is.na(df$mean_as))
```

```
## [1] 223
```

```
sum(is.na(df$mean_lc))
```

```
## [1] 25
```

```
df <- df %>%
  filter(mean_lc != "NA") %>%
  filter(mean_lmx != "NA") %>%
  filter(mean_as != "NA")

df <- df %>%
  mutate(gender_match = case_when((gender == 3 & gender_l == 1) |
                                    (gender == 4 & gender_l == 2) ~ 1, TRUE ~ 0))

# 565 observation after cleaning (final sample size)
```

Simple modeling (Trial)

```r
model_1 <- lm(mean_as ~ mean_lc, data = df)
summary(model_1)
```

```
##
## Call:
## lm(formula = mean_as ~ mean_lc, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.18574 -0.31644  0.02499  0.31070  1.18711
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.24649    0.12803  25.358  < 2e-16 ***
## mean_lc      0.17428    0.03191   5.461  7.1e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4904 on 563 degrees of freedom
## Multiple R-squared:  0.05031,    Adjusted R-squared:  0.04862
## F-statistic: 29.83 on 1 and 563 DF,  p-value: 7.102e-08
```

```r
model_2 <- lm(mean_as ~ mean_lc + mean_lmx, data = df)
summary(model_2)
```

```
##
## Call:
## lm(formula = mean_as ~ mean_lc + mean_lmx, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.24603 -0.29999  0.02186  0.28623  1.27130
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.67243    0.15148  17.642  < 2e-16 ***
## mean_lc      0.06695    0.03488   1.919   0.0554 .
## mean_lmx     0.24193    0.03696   6.546 1.33e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4731 on 562 degrees of freedom
## Multiple R-squared:  0.1176, Adjusted R-squared:  0.1145
## F-statistic: 37.45 on 2 and 562 DF,  p-value: 5.403e-16
```

```r
model_3 <- lm(mean_as ~ mean_lc + mean_lmx * gender_match, data = df)
summary(model_3)
```

```
##
```

```
## Call:
## lm(formula = mean_as ~ mean_lc + mean_lmx * gender_match, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.22695 -0.28474  0.00477  0.30145  1.25498
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.20185    0.23296  13.744  < 2e-16 ***
## mean_lc                0.07403    0.03465   2.137  0.03307 *
## mean_lmx               0.11603    0.05879   1.974  0.04890 *
## gender_match          -0.89688    0.28315  -3.168  0.00162 **
## mean_lmx:gender_match  0.20099    0.06892   2.916  0.00369 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.469 on 560 degrees of freedom
## Multiple R-squared:  0.1361, Adjusted R-squared:  0.1299
## F-statistic: 22.05 on 4 and 560 DF,  p-value: < 2.2e-16
```

Multiple Group Path Model using LAVAAN (Final Model)

```
df <- df %>%
  mutate(gen_mat4 = case_when(
    gender == 3 & gender_l == 1 ~ "MM",
    gender == 4 & gender_l == 2 ~ "FF",
    gender == 3 & gender_l == 2 ~ "MF",
    gender == 4 & gender_l == 1 ~ "FM"
  ))
df$gender_match <- as.numeric(df$gender_match)

df$gen_mat4 <- factor(df$gen_mat4, ordered = TRUE, levels = c("MM", "FF", "MF", "FM"))

MGmodel1 <- '
  # Direct effects
  mean_lmx ~ c(a1,a2, a3, a4)*mean_lc
  mean_as ~ mean_lc + c(b1,b2,b3,b4)*mean_lmx

  # indirect effects
  indirect1 := a1*b1
  indirect2 := a2*b2
  indirect3 := a3*b3
  indirect4 := a4*b4

  # difference in indirect effect
  indirect_diff1 := indirect1 - indirect2
  indirect_diff2 := indirect1 - indirect3
  indirect_diff3 := indirect1 - indirect4
  indirect_diff4 := indirect2 - indirect3
  indirect_diff5 := indirect2 - indirect4
  indirect_diff6 := indirect3 - indirect4
```

```
'
fitMG1 <- sem(MGmodel1, data = df, group = "gen_mat4", se = "bootstrap", bootstrap = 5000)
```

```
## Warning in lav_data_full(data = data, group = group, cluster = cluster, : lavaan WARNING: group varia
```

```
summary(fitMG1,  standardized = TRUE, rsq = TRUE)
```

```
## lavaan 0.6.16 ended normally after 1 iteration
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        28
##
##   Number of observations per group:
##     FF                                             339
##     MF                                              68
##     FM                                             105
##     MM                                              52
##
## Model Test User Model:
##
##   Test statistic                                 0.000
##   Degrees of freedom                                 0
##   Test statistic for each group:
##     FF                                           0.000
##     MF                                           0.000
##     FM                                           0.000
##     MM                                           0.000
##
## Parameter Estimates:
##
##   Standard errors                            Bootstrap
##   Number of requested bootstrap draws             5000
##   Number of successful bootstrap draws            5000
##
##
## Group 1 [FF]:
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   mean_lmx ~
##     mean_lc   (a1)    0.457    0.050    9.185    0.000    0.457    0.443
##   mean_as ~
##     mean_lc           0.154    0.056    2.767    0.006    0.154    0.178
##     mean_lmx  (b1)    0.273    0.056    4.866    0.000    0.273    0.324
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .mean_lmx          2.340    0.206   11.361    0.000    2.340    3.979
##    .mean_as           2.122    0.227    9.330    0.000    2.122    4.284
##
## Variances:
```

```
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .mean_lmx          0.278    0.023   11.947    0.000    0.278    0.804
##    .mean_as           0.200    0.020   10.198    0.000    0.200    0.813
##
## R-Square:
##                    Estimate
##     mean_lmx          0.196
##     mean_as           0.187
##
##
## Group 2 [MF]:
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   mean_lmx ~
##     mean_lc   (a2)    0.717    0.151    4.755    0.000    0.717    0.577
##   mean_as ~
##     mean_lc            0.082    0.114    0.717    0.474    0.082    0.100
##     mean_lmx  (b2)     0.177    0.082    2.143    0.032    0.177    0.268
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .mean_lmx          1.237    0.603    2.053    0.040    1.237    1.784
##    .mean_as           2.996    0.473    6.334    0.000    2.996    6.556
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .mean_lmx          0.321    0.063    5.125    0.000    0.321    0.667
##    .mean_as           0.185    0.029    6.345    0.000    0.185    0.887
##
## R-Square:
##                    Estimate
##     mean_lmx          0.333
##     mean_as           0.113
##
##
## Group 3 [FM]:
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   mean_lmx ~
##     mean_lc   (a3)    0.332    0.058    5.681    0.000    0.332    0.444
##   mean_as ~
##     mean_lc           -0.009    0.071   -0.120    0.905   -0.009   -0.014
##     mean_lmx  (b3)     0.102    0.102    1.004    0.315    0.102    0.122
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .mean_lmx          2.704    0.236   11.479    0.000    2.704    4.504
##    .mean_as           3.529    0.373    9.468    0.000    3.529    6.991
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .mean_lmx          0.289    0.037    7.764    0.000    0.289    0.803
```

```
##    .mean_as           0.251    0.029    8.558    0.000    0.251    0.986
##
## R-Square:
##                    Estimate
##     mean_lmx          0.197
##     mean_as           0.014
##
##
## Group 4 [MM]:
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   mean_lmx ~
##     mean_lc   (a4)    0.430    0.081    5.305    0.000    0.430    0.564
##   mean_as ~
##     mean_lc           0.136    0.102    1.341    0.180    0.136    0.204
##     mean_lmx  (b4)    0.310    0.127    2.445    0.014    0.310    0.353
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .mean_lmx          2.596    0.335    7.741    0.000    2.596    4.487
##    .mean_as           2.376    0.488    4.867    0.000    2.376    4.681
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .mean_lmx          0.228    0.038    6.068    0.000    0.228    0.682
##    .mean_as           0.194    0.035    5.566    0.000    0.194    0.753
##
## R-Square:
##                    Estimate
##     mean_lmx          0.318
##     mean_as           0.247
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##     indirect1         0.125    0.029    4.356    0.000    0.125    0.143
##     indirect2         0.127    0.055    2.311    0.021    0.127    0.155
##     indirect3         0.034    0.034    0.993    0.321    0.034    0.054
##     indirect4         0.133    0.060    2.211    0.027    0.133    0.199
##     indirect_diff1   -0.002    0.062   -0.035    0.972   -0.002   -0.011
##     indirect_diff2    0.091    0.044    2.045    0.041    0.091    0.089
##     indirect_diff3   -0.008    0.066   -0.127    0.899   -0.008   -0.056
##     indirect_diff4    0.093    0.065    1.431    0.152    0.093    0.101
##     indirect_diff5   -0.006    0.081   -0.078    0.938   -0.006   -0.044
##     indirect_diff6   -0.099    0.069   -1.432    0.152   -0.099   -0.145
```

Checking Model Fit

```
summary(fitMG1, fit.measures = TRUE)
```

```
## lavaan 0.6.16 ended normally after 1 iteration
##
```

```
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        28
##
##   Number of observations per group:
##     FF                                             339
##     MF                                              68
##     FM                                             105
##     MM                                              52
##
## Model Test User Model:
##
##   Test statistic                                 0.000
##   Degrees of freedom                                 0
##   Test statistic for each group:
##     FF                                           0.000
##     MF                                           0.000
##     FM                                           0.000
##     MM                                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                               239.099
##   Degrees of freedom                                12
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    1.000
##   Tucker-Lewis Index (TLI)                       1.000
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)               -795.762
##   Loglikelihood unrestricted model (H1)       -795.762
##
##   Akaike (AIC)                                1647.525
##   Bayesian (BIC)                              1768.906
##   Sample-size adjusted Bayesian (SABIC)       1680.020
##
## 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
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.000
##
## Parameter Estimates:
##
```

```
##   Standard errors                             Bootstrap
##   Number of requested bootstrap draws              5000
##   Number of successful bootstrap draws             5000
##
##
## Group 1 [FF]:
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   mean_lmx ~
##     mean_lc   (a1)    0.457    0.050    9.185    0.000
##   mean_as ~
##     mean_lc           0.154    0.056    2.767    0.006
##     mean_lmx  (b1)    0.273    0.056    4.866    0.000
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .mean_lmx          2.340    0.206   11.361    0.000
##    .mean_as           2.122    0.227    9.330    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .mean_lmx          0.278    0.023   11.947    0.000
##    .mean_as           0.200    0.020   10.198    0.000
##
##
## Group 2 [MF]:
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   mean_lmx ~
##     mean_lc   (a2)    0.717    0.151    4.755    0.000
##   mean_as ~
##     mean_lc           0.082    0.114    0.717    0.474
##     mean_lmx  (b2)    0.177    0.082    2.143    0.032
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .mean_lmx          1.237    0.603    2.053    0.040
##    .mean_as           2.996    0.473    6.334    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .mean_lmx          0.321    0.063    5.125    0.000
##    .mean_as           0.185    0.029    6.345    0.000
##
##
## Group 3 [FM]:
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   mean_lmx ~
##     mean_lc   (a3)    0.332    0.058    5.681    0.000
##   mean_as ~
```

```
##     mean_lc          -0.009    0.071   -0.120    0.905
##     mean_lmx  (b3)    0.102    0.102    1.004    0.315
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .mean_lmx          2.704    0.236   11.479    0.000
##    .mean_as           3.529    0.373    9.468    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .mean_lmx          0.289    0.037    7.764    0.000
##    .mean_as           0.251    0.029    8.558    0.000
##
##
## Group 4 [MM]:
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   mean_lmx ~
##     mean_lc   (a4)    0.430    0.081    5.305    0.000
##   mean_as ~
##     mean_lc            0.136    0.102    1.341    0.180
##     mean_lmx  (b4)    0.310    0.127    2.445    0.014
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .mean_lmx          2.596    0.335    7.741    0.000
##    .mean_as           2.376    0.488    4.867    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .mean_lmx          0.228    0.038    6.068    0.000
##    .mean_as           0.194    0.035    5.566    0.000
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|)
##     indirect1         0.125    0.029    4.356    0.000
##     indirect2         0.127    0.055    2.311    0.021
##     indirect3         0.034    0.034    0.993    0.321
##     indirect4         0.133    0.060    2.211    0.027
##     indirect_diff1   -0.002    0.062   -0.035    0.972
##     indirect_diff2    0.091    0.044    2.045    0.041
##     indirect_diff3   -0.008    0.066   -0.127    0.899
##     indirect_diff4    0.093    0.065    1.431    0.152
##     indirect_diff5   -0.006    0.081   -0.078    0.938
##     indirect_diff6   -0.099    0.069   -1.432    0.152
```

Model-fit

The CFI value of 1 and TL value of 1 suggests that model provides a perfect fit to the data compared to baseline model. The RMSEA value of 0 (which was 0.079 in the previous model I sent) and SRMR value of 0 (previously 0.042) suggests excellent fit to the data.

Mediation effect

In Group 1 (FF), characterized by female coaches and female employees, as well as in Group 4 (MM), where both coach and employee genders are male, our findings demonstrate statistically significant mediation effects ($p < 0.05$). LMX serves as a mediator in the relationship between leadership coaching and adaptive selling for these gender-matched groups. We also notice a significant mediation effect for Group 3 (FM), where the employee is female and the manager is male but not for the Group 4 (MF), where the employee is male and the manager is female. To summarize, there is mediation effect of LMX on all the groups except for group 4 (MF).

I also calculated mediation differences between all the groups (MM, FF, FM, MF). Except for the difference between Group 1 (FF) and Group 3 (FM) {indirect_dff2}, all the other differences were non-significant. This suggests the mediating process through LMX operates similarly acorss almost all groups.

Group differences in mediation (Moderation effect)

By looking at the regression coefficients and corresponding p-values for each group, a moderation effect is noticed. This effect is particularly significant for Group 1 (FF) and Group 4 (MM) as the $p< 0.02$ for the mediator variable in both the groups. Surprisingly. for Group 2 (MF), female coaches and male employee, the effect is also significant ($p = 0.037$). For Group 3, the p-value suggests insignificant effect which supports the hypothesis.