semfindr: An R package for Sensitivity Analysis in Structural Equation Modeling

Journal:	*Advances in Methods and Practices in Psychological Science*
Manuscript ID	AMPPS-24-0003
Manuscript Type:	Tutorial
Date Submitted by the Author:	03-Jan-2024
Complete List of Authors:	Cheung, Shu Fai; University of Macau, Department of Psychology Lai, Mark; University of Southern California Department of Psychology, Psychology
Method and Stats :	structural equation modeling (SEM), inferential, R
Substance Keywords:	
Additional Keywords:	

SCHOLARONE™
Manuscripts

AMPPS Questions about Transparency Practices - AMPPS-24-0003

Registered Report

Q: Is your manuscript a Stage-1 Registered Report?

A: No

If yes, provide a link to the project location where the manuscript, materials, and data will eventually be stored if the study is provisionally accepted.

CUST_TRANSPARENCY_RR_TEXT :No data available.

Empirical Work

Q: Does your paper present the results of new studies or analyses of data from human participants or from animals?

A: Not Applicable

If Yes, provide the name of the institution that granted ethical approval and the protocol number (e.g., e.g., Protocol #12345 approved by the University of Illinois IRB). If no ethical approval was required, give a brief explanation of why not.

CUST_TRANSPARENCY_EMPIRICAL_TEXT :No data available.

If your paper presents empirical work with human participants, please indicate whether it adhered to the Declaration of Helsinki.

A: Not Applicable

If your paper presents new empirical work, does it include the following statement: "We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study"?

A: Not Applicable

If your paper presents new empirical work and (a) you did not include this statement in your manuscript, (b) you included a modified version of this statement, or (c) any part of this statement is untrue, please explain why.

CUST_TRANSPARENCY_STATEMENT_TEXT :No data available.

Available Data

Q: Does your paper rely on new or previously unpublished empirical data from your lab?

A: No

Does your paper analyze data from pre-existing datasets or data made available by other researchers?

A: Yes

For Review Only

If you answered Yes to either of these questions, provide a URL (either public or view-only) where the data can be accessed by the editors and reviewers.

https://osf.io/k2xhp/?view_only=9106e5d6376d4f16bdc23426f367e9e6

If your paper relies on existing data that are available via third parties, please indicate who controls access to those data and how other researchers can access them in the same way you have.

CUST_TRANSPARENCY_TEXT_THIRD_PARTY_DATA :No data available.

If needed, add any additional explanation about the data used in your paper.

CUST_TRANSPARENCY_TEXT_DATA_OTHER :No data available.

Available Materials

Q: Have you made available any and all materials necessary to reproduce your experiments, analyses, or other paper contents?

A: CUST_TRANSPARENCY_MATERIALS :No data available.

If No, please explain which materials are unavailable and explain why they are not available.

CUST_TRANSPARENCY_MATERIALS_TEXT :No data available.

Q: Does your paper rely on any materials, code, or other resources that are new to this project (i.e., they were developed or created as part of the research reported in this paper)?

A: Yes

If you answered Yes, provide a URL (either public or view-only) where reviewers and editors can view those materials and resources. Enter “Not Applicable” if your paper does not rely on any such materials.

https://osf.io/k2xhp/?view_only=9106e5d6376d4f16bdc23426f367e9e6

Q: Does your paper rely on any materials, code, or other resources that are in the public domain or previously made available by you or by other researchers?

A: Not Applicable

If Yes, please indicate who controls access to those materials and how other researchers can access them in the same way you have. Enter “Not Applicable” if you are not relying on data from other researchers or third parties.

CUST_TRANSPARENCY_EXISTING_MATERIALS_TEXT :No data available.

Preregistration

Q: Were any of the studies or analyses reported in the manuscript preregistered?

A: Not Applicable â€“ this manuscript does not report any analyses of new data or secondary analyses of existing data

For Review Only

If only a subset of the reported studies were preregistered, indicate which ones were and which ones were not. For any studies that were not preregistered, please indicate why not. If your manuscript does not report new studies, please enter, “Not applicable.”

CUST_TRANSPARENCY_PREREG_STUDIES :No data available.

Please provide a URL for the main project page where reviewers and editors can access the preregistration documentation (leave blank if there are no studies or none was preregistered). This may be an anonymous view-only link for the review process.

CUST_TRANSPARENCY_PREREG_URL :No data available.

Which aspects of your project were preregistered? (check all that apply. Note that if your preregistration includes a complete analysis script that handles coding of measures, missing data, exclusions, analyses, etc., you could check multiple boxes on this list based on preregistering that script.)

CUST_TRANSPARENCY_PREREG_ASPECTS :No data available.

When did you complete the preregistration:

Not Applicable - no preregistered studies

Did you make any changes to the preregistered procedures when completing your study?

A: Not Applicable

If you made changes to your preregistered plans at any stage of the process of completing your research, list all of those changes below.

CUST_TRANSPARENCY_PREREG_CHANGES_TEXT :No data available.

Such changes must also be documented in the manuscript itself. Are all of the changes specified above fully reported in the manuscript text?

CUST_TRANSPARENCY_PREREG_CHANGES_DOCUMENTED :No data available.

If you have other comments or explanations about your preregistration that not covered by the questions above, enter them here (leave blank if no you have no comments/notes or if your paper has no preregistered studies):

CUST_TRANSPARENCY_PREREG_OTHER :No data available.

For Review Only

Q: Authors were asked to select items from the list below to indicate what was preregistered:

- Not Applicable - no preregistered studies
- Theoretical hypotheses
- Tasks/measures used for confirmatory hypothesis tests
- Tasks/measures used for exploratory hypothesis tests
- Tasks/measures that were collected for other purposes
- Data collection stopping rules
- Data source(s) (for preregistration of analyses of pre-existing data)
- Data coding procedures (e.g., how measures would be coded and scored)
- Data exclusion criteria and procedures
- Procedures for handling missing data
- Procedures to handle failures of quality control
- Data analysis plan
- Data analysis scripts/code (e.g., full R scripts for analysis)
- Planned interpretation for different patterns of results (could be part of the "Theoretical hypotheses" if each hypothesis states how different patterns of results would support or disconfirm it).
- Target sample size

The author checked the following boxes:

A: CUST_TRANSPARENCY_PREREG_ASPECTS :No data available.

For Review Only

semfindr: An R package for Sensitivity Analysis in Structural Equation Modeling

Shu Fai Cheung[1] and Mark H. C. Lai[2]

[1]Department of Psychology, University of Macau

[2]Department of Psychology, University of Southern California

For Review Only

Author Note

Shu Fai Cheung https://orcid.org/0000-0002-9871-9448, Department of Psychology; Mark H. C. Lai https://orcid.org/0000-0002-9196-7406, Department of Psychology. Correspondence concerning this article should be addressed to Shu Fai Cheung, Department of Psychology, Faculty of Social Sciences, University of Macau, Avenida da Universidade, Taipa, Macao SAR, China or by email (sfcheung@um.edu.mo).

Abstract

Measuring case influence on parameter estimates and model fit measures (casewise sensitivity analysis) is important for assessing the robustness of findings in structural equation modeling (SEM). However, it was rarely reported clearly (Wulff et al., 2023), or was conducted inappropriately, ignoring the model-specific influence. One possible reason is the need to refit a model once for each case (Pek & MacCallum, 2011), implemented in some existing tools, which is time consuming when a model is not fast to fit and/or the sample size is large. We developed an easy-to-use R package, *semfindr*, for casewise sensitivity analysis in SEM using the leave-one-out method. It reduces the computational cost by separating the refitting step from case influence computation step. It also have various plot functions for effective sensitivity analysis in complicated models. Last, it supports multiple-group models and missing data. This tutorial illustrates how to use *semfindr* to do casewise sensitivity analysis efficiently, with publication-ready results and plots.

Keywords: structural equation modeling, sensitivity analysis, outliers, influential cases

For Review Only

semfindr: An R package for Sensitivity Analysis in Structural Equation Modeling

Checking for influential cases, cases that affect results substantially if removed, has been an important topic for decades in multiple linear regression. However, in structural equation modeling (SEM), it received little attention in applied research. Sometimes it was conducted but not reported transparently (see Wulff et al., 2023, on outliers in general), or sometimes it was done incorrectly, focusing on univariate or multivariate outliers instead of case influential on model-specific results, such as parameter estimates (Pek & MacCallum, 2011). Identifying potentially influential cases allows researchers to evaluate the robustness of the results. Whether a case is influential depends on the model being fitted. A case is influential if including this case substantially influences one or more crucial aspects of the SEM results, such as substantially decreasing the estimate of a path coefficient or the model fit substantially worse (Pek & MacCallum, 2011). Note that an influential case needs *not* be an outlier, and an outlier needs *not* be an influential case (Flora et al., 2012). Therefore, having done outlier screening cannot justify not checking for influential cases.

We believe one reason for focusing on outliers instead of influential cases is the scarcity of easy-to-use efficient tools. Some tools are available, such as *faoutlier* (Chalmers & Flora, 2015) and *influence.SEM* (Pastore & Altoe, 2018) for R and *FINDOUT* (Cheung & Pesigan, 2023) for AMOS (Arbuckle, 2021). *Mplus* (Muthén & Muthén, 2017) can also compute some measures of case influence. However, there are cases in which researchers may be deterred from using these techniques. In this tutorial, we introduce *semfindr*, an R package, for identifying influential cases in SEM analysis efficiently, with easy-to-use diagnostic plots. It supports multiple-group models, models with missing data, and any estimation method supported by *lavaan* (Rosseel, 2012), a popular R package for SEM. For models that take a long time to fit, or for large samples, it also supports doing the search on a selected subset of cases.

For Review Only

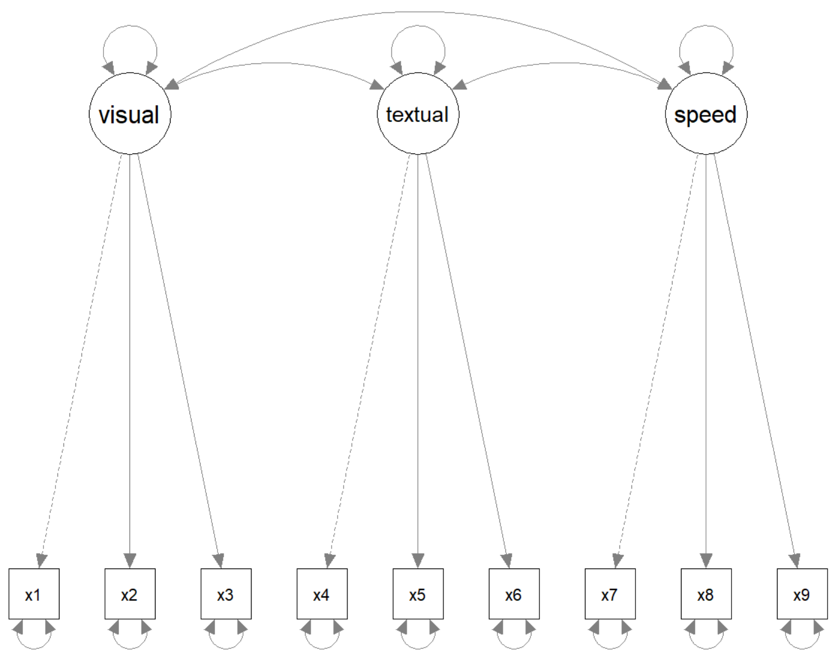
Leave-One-Out (LOO) By *semfindr*

We present below one simple-and-exact method to measure case influence: the leave-one-out (LOO) method, and illustrate how to do it by *semfindr*. Basic knowledge in using *lavaan* and R is assumed.

For illustration, we adopted the approach used by Pek and MacCallum (2011) to create a dataset based on the real dataset by Holzinger and Swineford (1939, cited in Pek & MacCallum, 2011). We used the version supplied in Rosseel (2012) and use nine variables, `x1` to `x9`, to measure three factors, `speech`, `textual`, and `speed`. The model is shown in Figure 1.

Figure 1

The Model In The Illustration.



We randomly selected 99 cases from the dataset. We then added a case, Case 100, with scores on `x1`, `x2`, `x3`, `x7`, `x8`, and `x9` two standard deviations (*SD*s) above means, and `x4`, `x5`, and `x6` two *SD*s below means, *SD*s and means computed from the 99 cases selected. This case is not an outlier based on the commonly used criterion of 3 *SD*s. Moreover, the patterns of scores are consistent with the factor structure if considered separately for each factor. However, if all nine scores are considered together, the pattern is inconsistent with the structure because the factor covariances are all positive if estimated using the 99 cases. This case will be used to illustrate the importance of identifying influential cases based on the model being fitted. The dataset for illustration can be downloaded from the OSF page for this manuscript (https://osf.io/k2xhp/?view_only=9106e5d6376d4f16bdc23426f367e9e6).

The LOO method involves three steps:

1. Fit the model.

2. Remove a case and fit the model again.

3. Explore case influence by comparing the results with and without this case.

Steps 2 and 3 are usually repeated once for each case to compute case influence measures for all cases.

Step 1: Fit the Model

Fit the model in *lavaan* as usual:

```
library(lavaan)
mod <- "visual  =~ x1 + x2 + x3
        textual =~ x4 + x5 + x6
        speed   =~ x7 + x8 + x9"
fit <- cfa(model = mod, data = dat)
```

No special treatment in this step in using *semfindr*.

Step 2: Refit the Model

This step is conducted by `lavaan_rerun()` on the output of `sem()` and `cfa()` in *lavaan*. It retrieves the stored information and fits the model once for each case, each time with this case removed, and stores the results in a `lavaan_rerun`-class object. Parallel processing can be enabled to speed up the process.

```
library(semfindr)
fit_rerun <- lavaan_rerun(fit,
                          parallel = TRUE,
                          makeCluster_args = list(4))
```

For Review Only

Setting `parallel` to `TRUE` enables parallel processing. Setting `makeCluster_args` to `list(4)` uses four CPU cores. [1]

Separating this step from case influence computation is much more efficient than the approaches in some other tools because researchers can do as many case influence computations as they want without refitting the models.

Printing the output gives a summary of this step:

```
> fit_rerun
=== lavaan_rerun Output ===
Call:
lavaan_rerun(fit = fit, parallel = TRUE, makeCluster_args = list(4))
Number of reruns: 100
Number of reruns that converged (solution found): 100
Number of reruns that failed to converge (solution not found): 0
Number of reruns that passed post.check of lavaan: 100
Number of reruns that failed post.check of lavaan: 0
Number of reruns that both converged and passed post.check: 100
Number of reruns that either did not converge or failed post.check: 0
```

Step 3: Explore Case Influence

We first introduce an all-in-one function, `influence_stat()`, for computing major case influence measures (presented later) using one function. Using the default options, researchers only need to pass the output of `lavaan_rerun()` to `influence_stat()`:

```
fit_influence <- influence_stat(fit_rerun)
```

It will:

- compute the case influence of each case on model χ^2, CFI, TLI, and RMSEA by calling `fit_measures_change()`,
- compute the standardized changes of all parameters and the *gCD* based on these parameters by calling `est_change()`, and

[1] Advanced users can use `makeCluster_args` to pass arguments to `makeCluster()` from the package *parallel* to customize the cluster.

For Review Only

- compute the Mahalanobis distance (Mahalanobis, 1936) of each case on all variables by calling `mahalanobis_rerun()`.

The output is a numeric matrix of the class `influence_stat`, when printed, shows the three types of influence measures in the above list. By default, only the top ten influential cases will be printed in each section, though only selected cases are shown below to save space. In most cases, just printing the output is enough. For illustration, we use the argument `what` below to control what is printed.[2]

Case Influence on Fit Measures

To assess case influence on model fit, we can compute the change of model χ^2 when a case is included (Pek & MacCallum, 2011):

$$\Delta\chi^2_i = \chi^2 - \chi^2_{(-i)}, \tag{1}$$

where χ^2 and $\chi^2_{(-i)}$ are the model χ^2 in the full sample and without the ith case, respectively. Therefore, $\Delta\chi^2_i$ is the influence of the ith case on model χ^2. $\Delta\chi^2$ can also be used for any estimation method that yields a χ^2-like measure, such as MLR in *M*plus and *lavaan*, which report a scaled χ^2.[3]

The idea can be applied to most other measures of fit. For example, researchers may assess how a case influences fit measures such as CFI and RMSEA:

$$\Delta M_i = M - M_{(-i)}, \tag{2}$$

where M and $M_{(-1)}$ are the fit measure in the full sample and without the ith case,

[2] By default, case influence measures on model χ^2, CFI, TLI, and RMSEA will be computed. This can be changed by the argument `fit_measures`, default to `c("chisq", "cfi", "rmsea", "tli")`, to support any fit measures reported by *lavaan*, including measures such as scaled χ^2 and robust CFI when robust estimator such as MLR is used.

[3] Note that $\Delta\chi^2_i$ is a measure of case influence. It is not a measure of the difference in fit between two models. It cannot be used to conduct a χ^2 difference test.

For Review Only

respectively. ΔM_i is the case influence of the ith case on the fit measure M.

To print case influence on fit measures, call `print()` and set `what` to `"fit_measures"`. To sort cases by absolute influence on, say, model χ^2, add `sort_fit_measures_by = "chisq"`, `"chisq"` being the column name:

```
print(fit_influence,
      what = "fit_measures",
      sort_fit_measures_by = "chisq")
```

This is an excerpt of the output:

```
-- Case Influence on Fit Measures --
     chisq    cfi  rmsea    tli
100 -3.694  0.017 -0.013  0.026
99   3.481 -0.014  0.016 -0.020
85   2.973 -0.012  0.013 -0.018
14  -2.586  0.010 -0.010  0.015
8   -2.172  0.009 -0.008  0.013
```

Case 100 case has a value of -3.694. Including this case in the analysis decreases the model χ^2 by 3.694 (better fit). Case 99 has a value of 3.481. That is, including this case increases the model χ^2 by 3.481 (worse fit). The largest absolute value is 3.694, indicating the largest absolute change among all cases.

It is difficult to identify influential cases using the numerical displays. Moreover, the *distribution* of values is more useful than the absolute values in judging whether a case is unusual in case influence (Aguinis et al., 2013). Therefore, *semfindr* emphasizes inspecting case influence graphically. The function `index_plot()` is for visualizing case influence:

```
index_plot(fit_influence,
           "chisq",
           absolute = TRUE,
           x_label = "Influence on Chi-Square",
           largest_x = 3)
```

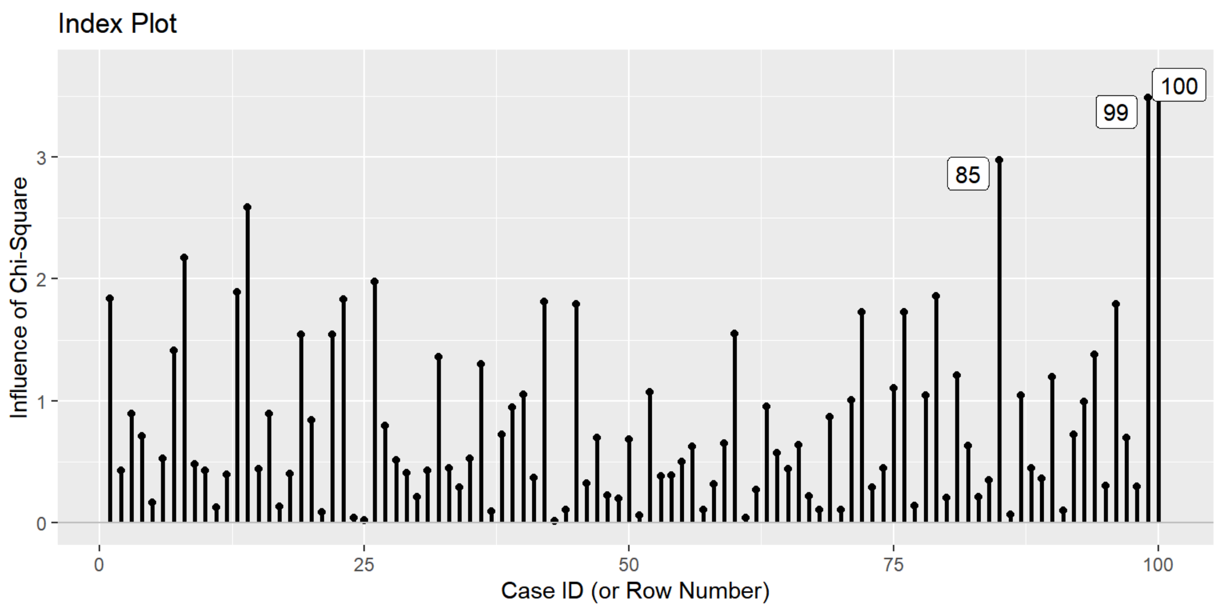
The first argument is a matrix-like object, which can be the output of `influence_stat()`. The second argument is the name of the column to be plotted. If the magnitude is to be compared, adding `absolute = TRUE` will plot the absolute

For Review Only

values. The label of the vertical axis is set by `x_label`, and `largest_x` controls the number of cases to be labelled based on the magnitude of influence.[4]

Figure 2

The Index Plot of Case Influence on Model Chi-Square.



As shown in Figure 2, no case has an unusually large influence on model χ^2.

If desired, users can set `sort_fit_measures_by` to another column, such as `cfi`, to sort cases by this fit measure. The function `index_plot()` can also be used on any other columns.

Case Influence on Parameter Estimates

Two types of measures are usually used in LOO for case influence on parameter estimates. The first is the raw influence on a parameter estimate. For the jth parameter, θ_j, it is given by

$$\Delta\hat{\theta}_{ji} = \hat{\theta}_j - \hat{\theta}_{j(-i)}, \tag{3}$$

where $\hat{\theta}_j$ and $\hat{\theta}_{j(-i)}$ are the parameter estimates of $\hat{\theta}_j$ in the full sample and with the ith

[4] The plot can be customized in many ways. See the help page for other available options.

case removed, respectively. Because this measure is similar to DFBETA in multiple regression, We call $\Delta\hat{\theta}_{ji}$ the *DFTHETA*.

DFTHETA is useful when the units of variables are meaningful, for example, when the parameter is a regression coefficient from one variable to another, both measured in units such as seconds or kilograms. However, if the units are not interpretable, DFTHETA is difficult to interpret. There are two alternatives: raw change in standardized parameter and standardized change.

The former is simply the raw change of a parameter estimate in the standardized solution, which we call *DFZTHETA*, Z for standardization. For example, if the parameter is the covariance between two latent factors, the raw change in the standardized solution is the change in their correlation. If the standardized solution is interpretable, so is the DFZTHETA.

The latter, proposed by Pek and MacCallum (2011), divides the raw change of a case by the standard error of the parameter estimate, estimated without this case:

$$\Delta\hat{\theta}^*_{ji} = \frac{\hat{\theta}_j - \hat{\theta}_{j(-i)}}{\hat{\sigma}_{\theta_{j(-i)}}} = \frac{\Delta\hat{\theta}_{ji}}{\hat{\sigma}_{\theta_{j(-i)}}}, \tag{4}$$

where $\hat{\sigma}_{\theta_{j(-i)}}$ is the estimated standard error of $\hat{\theta}_{j(-i)}$. This measure is invariant to scale changes in the observed variables and takes into account the sampling variance. Conceptually, this measure is similar to DFBETAS in multiple regression, S for standardized by the standard error. Therefore, we call it *DFTHETAS*.

When a model has many parameters, a summary statistic measuring the overall influence on a set of parameters would be useful. Pek and MacCallum (2011) proposed to use generalized Cook's distance, gCD, (Cook, 1977):

$$gCD_i = (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{(-i)})'\hat{V}_{\hat{\boldsymbol{\theta}}_{(-i)}}(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{(-i)}) \tag{5}$$

where $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}_{(-i)}$ are the $k \times 1$ vectors of the parameter estimates in the full sample and with the ith case removed, respectively, and $\hat{V}_{\hat{\boldsymbol{\theta}}_{(-i)}}$ is the estimated sampling variance-covariance of the estimates without the ith case.

The gCD measures the total influence of a case on the parameters. Although

For Review Only

usually computed for all free parameters, *gCD* can also be computed using a subset of parameters (Pek & MacCallum, 2011). For example, if the main interest is in the factor loadings, then using only the factor loadings to compute *gCD* can assess case influence on factor loadings only.

To print DFTHETASs, call `print()` and set `what` to `"parameters"`. By default, cases are sorted by *gCD*:

```
print(fit_influence,
      what = "parameters")
```

This is an abridged version of the output:

```
-- Standardized Case Influence on Parameter Estimates --
    visual=~x2 ... visual~~textual ... textual~~speed   gcd
100      0.433 ...          -0.879 ...          -0.519 3.303
85       0.190 ...           0.144 ...          -0.006 2.223
87       0.346 ...          -0.704 ...          -0.121 1.626
40       0.738 ...          -0.363 ...           0.348 1.284
32       0.025 ...           0.466 ...           0.173 1.226
```

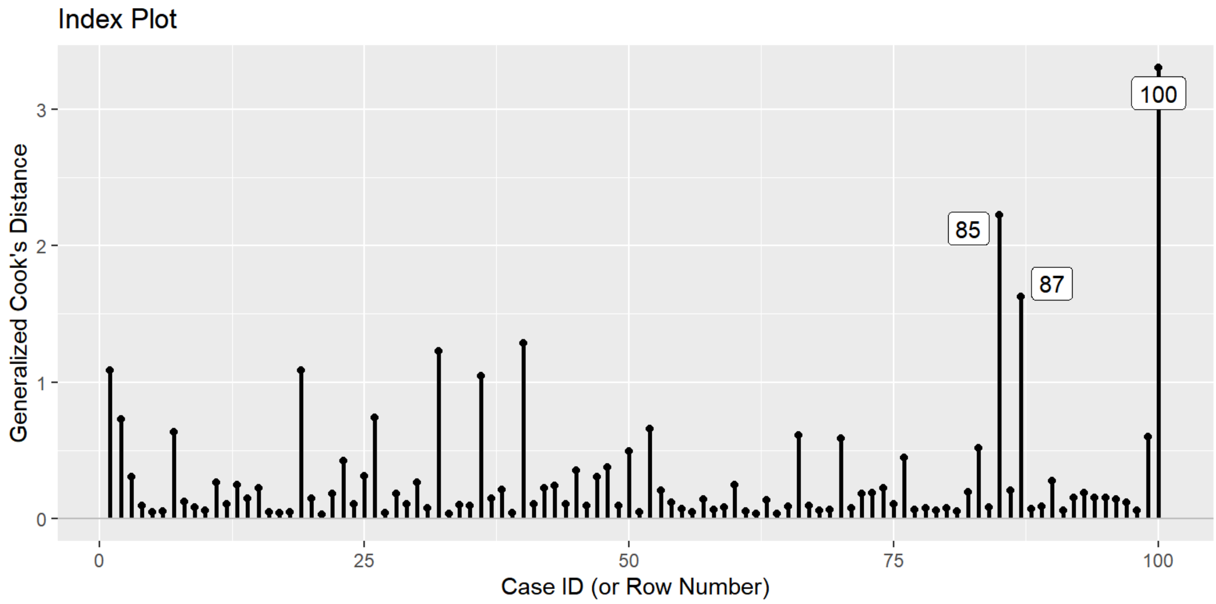
Case 100 has the largest *gCD* (3.303). As shown in Figure 3, this case is the only case having unusually large *gCD*:

```
index_plot(fit_influence,
           "gcd",
           x_label = "Generalized Cook's Distance",
           largest_x = 3)
```

For Review Only

Figure 3

The Index Plot of gCD.



Review Only

The DFTHETASs also suggest that Case 100 has a large influence on the estimated covariance between `visual` and `textual` (`visual~~textual`).

The DFTHETAS -0.519 for Case 100 on `visual~~textual` is difficult to interpret. The function `est_change_raw()` can be used to compute DFTHETAs, using `parameters` to specify the parameters to be included. Adding `standardized = TRUE` gives DFZTHETAs. This is an example of computing the DFZTHETA for all three factor correlations:

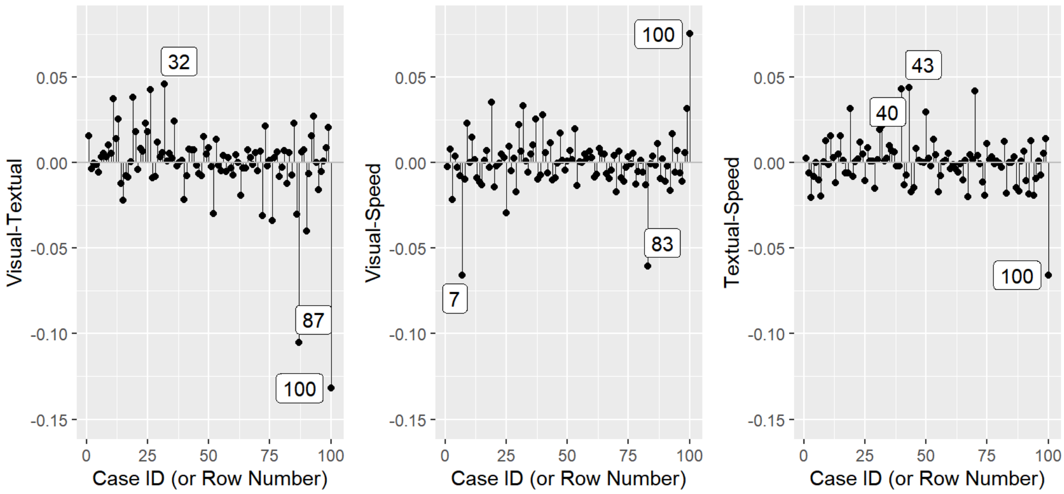
```
# est_change_raw(fit_rerun,
#                parameters = c("visual~~textual",
#                               "visual~~speed",
#                               "textual~~speed"),
#                standardized = TRUE)
-- Case Influence on Standardized Parameter Estimates --
    id textual~~speed  id visual~~speed  id visual~~textual
1  100         -0.066 100         0.075 100          -0.132
2   43          0.044   7        -0.066  87          -0.105
3   40          0.043  83        -0.060  32           0.046
4   70          0.042  19         0.035  26           0.042
5   19          0.031  32         0.033  90          -0.040
```

Adding Case 100 decreases the correlation between `visual` and `textual` by

0.132, which is a substantial change for a correlation. The index plots for the three columns of DFZTHETA also confirmed that this case is unusually influential, as shown in Figure 4

Figure 4

The Index Plots of DFZTHETAs.



Mahalanobis Distance

Readers may wonder whether this case can also be identified using outlier measures such as Mahalanobis distance, a multivariate measure of how far away a case is from the means of several variables, available in some SEM programs (e.g., *M*plus and AMOS). The Mahalanobis distance can be printed by setting `what` to `"mahalanobis"`:

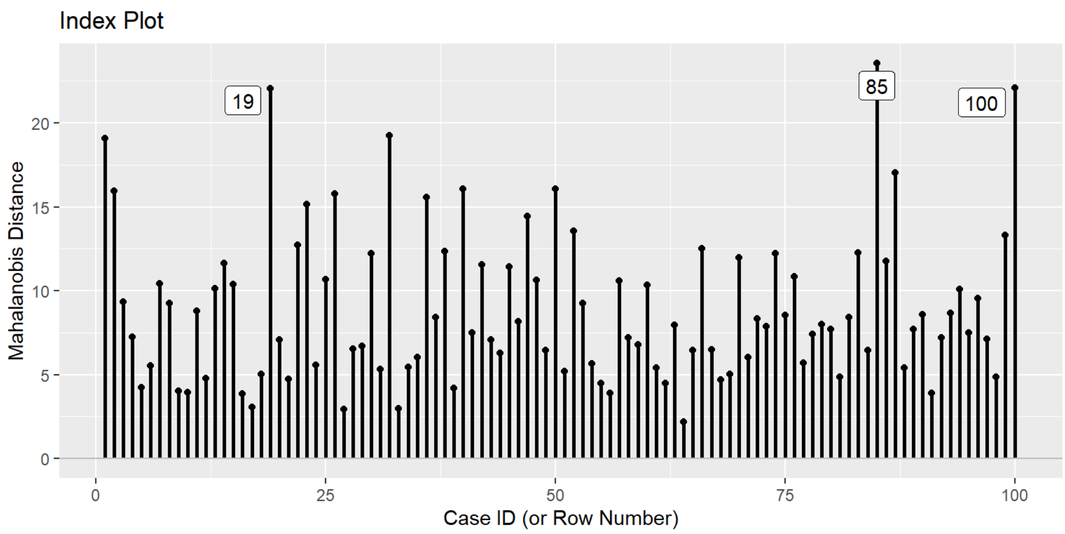
```
print(fit_influence,
      what = "mahalanobis")
```

This is an excerpt of the printout, and Figure 5 is the index plot of the Mahalanobis distance for all cases.

```
-- Mahalanobis Distance --
         md
85   23.558
100  22.078
19   22.044
32   19.233
1    19.090
```

Figure 5

The Index Plots of Mahalanobis Distance.



r Review Only

This example shows that a case can be influential on some aspects of a model even if it is not extreme on the observed variables. Actually, none of the cases show unusually large Mahalanobis distance, and Case 100 is not even the case with the largest Mahalanobis distance. If researchers rely on a measure of extremeness, they will fail to find the case that is influential on the estimates of some parameters.

gCD *for Selected Parameters*

By default, *gCD* is computed using all parameters. To assess case influence on specific parameters, use the function `est_change()`:

```
gcd_loadings <- est_change(fit_rerun,
                           parameters = "=~")
```

The first argument is the output of `lavaan_rerun()`. The argument `parameters` is used to select the parameters used to compute *gCD*. If we only need to select parameters based on the operators in *lavaan* model syntax, we can include the operator. For example, in the above example, `"=~"` is used to select all factor loadings.

This is an excerpt of the output:

```
-- Standardized Case Influence on Parameter Estimates --
    visual=~x2 visual=~x3 textual=~x5 textual=~x6 speed=~x8 speed=~x9   gcd
40       0.738      0.786       0.073       0.000     0.189     0.319 0.898
```

```
85        0.190    -0.045     0.606     0.631     0.158     0.109 0.591
26       -0.119    -0.487    -0.064    -0.343    -0.117     0.026 0.408
100       0.433     0.527     0.074    -0.031    -0.098     0.094 0.365
19       -0.338    -0.288    -0.120    -0.392    -0.007     0.215 0.358
```

By default, cases were sorted by *gCD* in descending order. The columns for each parameter are the same DFTHETASs presented before.[5] The column `gcd` shows the *gCD* values computed using only these columns. The results show that, if computed only on the free factor loadings. Case 100 does not have unusual influence on the factor loadings.

We can also assess case influence on factor covariance:

```
> gcd_fcov <- est_change(fit_rerun,
+                        parameters = c("visual~~textual",
+                                       "visual~~speed",
+                                       "textual~~speed"))
> gcd_fcov
-- Standardized Case Influence on Parameter Estimates --
    visual~~textual visual~~speed textual~~speed   gcd
100          -0.879         0.539         -0.519 1.661
87           -0.704        -0.008         -0.121 0.504
19            0.496         0.295          0.264 0.348
40           -0.363         0.006          0.348 0.285
32            0.466         0.268          0.173 0.280
```

The results confirm that Case 100 substantially influences factor covariances, with *gCD* much larger than those of other cases relatively.

Diagnostics Plots

There are several other plot functions for visualizing case influence.

gcd_gof_md_plot(). To visualize three measures in one single plot: case influence on a fit measure, Mahalanobis distance, and *gCD*, use `gcd_gof_md_plot()`. This is an example:

```
gcd_gof_md_plot(fit_influence,
                fit_measure = "chisq",
```

[5] Note that the first factor loadings are fixed to one by default, for identification. Therefore, the model only has six free factor loadings.

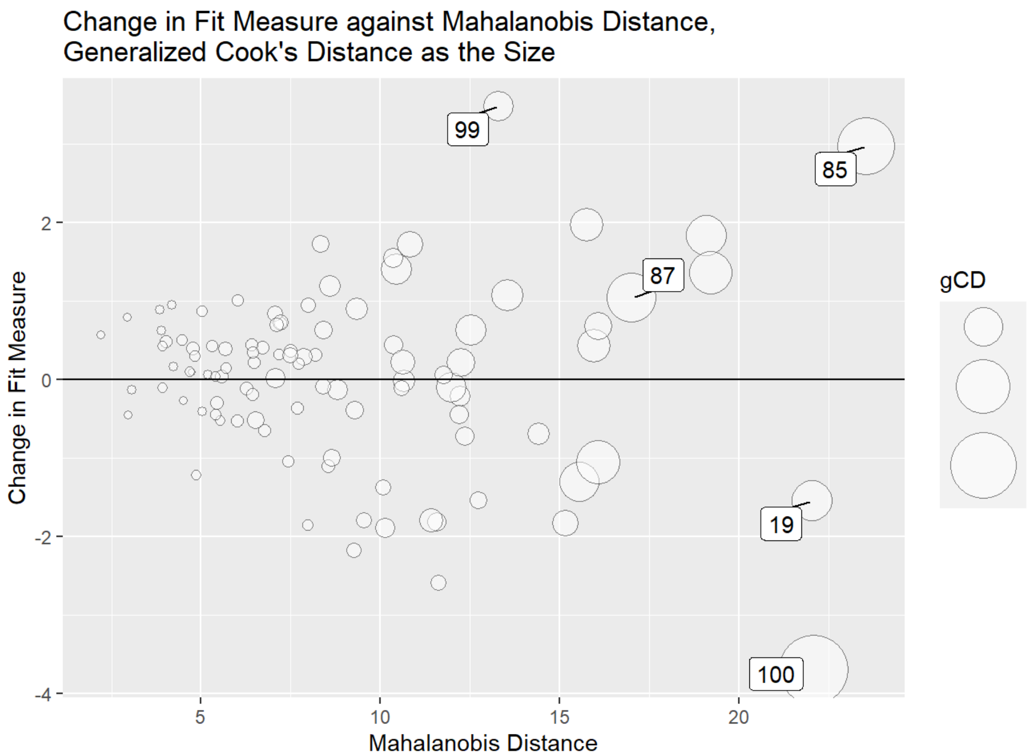
For Review Only

```
largest_gcd = 3,
largest_fit_measure = 3,
largest_md = 3,
circle_size = 25)
```

The first argument is the output of `influence_stat()`; `fit_measure` specifies the fit measure to be used for case influence (`"chisq"`, model χ^2, in this example); `largest_gcd`, `largest_fit_mesures`, and `largest_md` specifies the numbers of cases with the largest absolute value on these values to be labelled (default is 1). The size of the largest circle is controlled by `circle_size`. Increase this number to make the differences in circle sizes larger for readability. The output is shown in Figure 6.

Figure 6

The Output of gcd_gof_md_plot().



We recommend inspecting this plot first because it gives an overview of three different aspects in one graph, inspired by `influencePlot()` from the *car* package (Fox & Weisberg, 2018). The cases with large "bubbles" are cases with large *gCD*. If there are bubbles that are relatively large, cases influential on parameter estimates are present. Cases unusually high or low vertically are cases that are influential on the selected fit measures. Although outliers are not necessarily influential in a model, they are still cases that may need to be examined. These outliers can be identified by finding cases unusually far to the right (large on Mahalanobis distance).

The sample plot shows that (a) no case is unusually extreme when all variables are considered together (no case is unusually far away to the right), (b) no case unusually affects model χ^2, and (c) there is a case, Case 100, that unusually affects parameter estimates (unusually large in size). Using this plot is much easier than reading the tables of the three measures presented before.

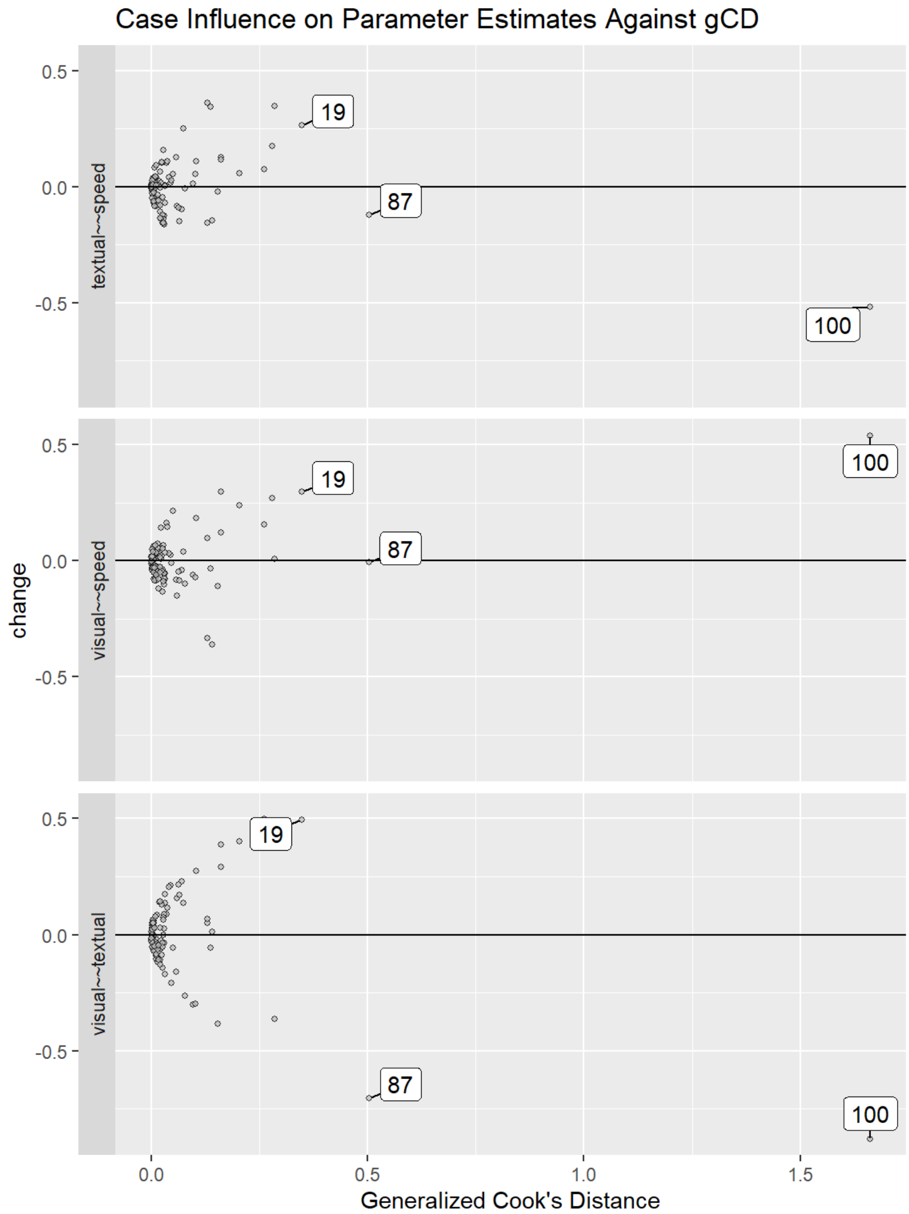
est_change_gcd_plot(). To visualize DFTHETAS and *gCD* in one plot, use `est_change_gcd_plot()`. The following is a sample call, using `est_change()` to compute *gCD* on the factor covariances, and using `est_change_gcd_plot()` to plot the results. The number of cases to be labelled based on *gCD* is set by `lartgest_gcd`. The argument `wrap_aes` is for customizing the call to `facet_wrap()` from the *ggplot2* package. Adding `list(scales = "fixed")` will use the same ranges of the horizontal and vertical axes, making it easier to compare the plots.

```
gcd_fcov <- est_change(fit_rerun,
                       parameters = c("visual~~textual",
                                      "visual~~speed",
                                      "textual~~speed"))
est_change_gcd_plot(gcd_fcov,
                    largest_gcd = 3,
                    wrap_aes = list(scales = "fixed"))
```

For Review Only

Figure 7

The Output of est_change_gcd_plot().



The plot (Figure 7) has one panel for each parameter, DFTHETAS plotted against *gCD*. It shows that Case 100 has unusually large *gCD*, and substantially influences the factor covariance estimates.

What to Do With Influential Cases

It is beyond the scope of this paper to give a comprehensive discussion on handling influential cases. Our focus is to introduce an accessible and efficient tool to check for influential cases in SEM. Readers are referred to Aguinis et al. (2013) for a comprehensive discussion for SEM. They proposed that the first step is to identify *error outliers*, cases that are not valid observations (e.g., error in data, or not from the target population, etc.). This can be done by other tools that compute standardized scores (z scores) and Mahalanobis distance. These cases should be handled before fitting a model in SEM. For example, correct the data error or remove cases not from the target population. The package *semfindr* is for the second stage they proposed, a decision tree for SEM (Aguinis et al., 2013, Figure 2), summarized briefly below.

First, identify cases influential on model fit and/or parameter estimates[6]. Aguinis et al. (2013) suggested using graphical tools such as index plots due to the lack of commonly agreed-upon cutoff values. This can be done by `gcd_gof_md_plot()`. It is convenient to use *semfindr* because just three function calls (`lavaan_rerun()`, `influence_stat()`, and `gcd_gof_md_plot()`) can visualize all the information needed in this step. For cases large on gCD, researchers can pinpoint the influence by computing gCD for selected sets of parameters using `est_change()`, and then visualize the influence using index plots by `est_change_gcd_plot()`.

Second, handle the influential cases. There is no simple one-size-fits-all solution in this step. We recommend first checking whether the influential cases are actually error outliers not identified in the initial screening. If yes, then they can be handled as error outliers. If not, then inspect them to identify potential reasons for being influential. They may be what Aguinis et al. (2013) called *interesting outliers*, cases suggesting phenomena or situations which deserve further investigation. For example, an influential case may reveal an unmodelled curvilinear relation, or undetected contamination such as a procedural error in data collection Cohen et al. (2003). The probable nature of the influential case helps determining how to handle them.

Cohen et al. (2003) suggested three remedial actions for multiple regression,

[6] Aguinis et al. (2013) named cases influential on model fit as *model fit outliers* and cases influential on parameter estimates as *prediction outliers*. We do not use *outliers*, to prevent confusing them with cases extreme univariately or multivariately, regardless of the model fitted.

For Review Only

which are also applicable to SEM (Aguinis et al., 2013). First, remove the influential cases. Aguinis et al. (2013) recommended reporting both the results with and without the influential cases, for transparency. Second, respecify the model, such as modelling the nonlinear relation and including moderating effect. Third, use robust estimation methods such as robust SEM using M-estimator (Yuan & Zhang, 2012). Which action to take depends on the probable source of the influence. We agree with Aguinis et al. (2013) that transparency is the major concern and recommend researchers to report in details actions adopted in identifying and handling influential cases, including the rationales.

Features in *semfindr*

The *semfindr* package has other features to facilitate sensitivity analysis:

- Subset data in Step 1 manually or using Mahalanobis distance.
- Select fit measures used in computing case influence on model fit.
- Use a version of the one-step approximation method discussed by Tanaka et al. (1991) to approximate case influence measures. This saves the computational cost for refitting. The case influence is not exact but can be used for selecting cases for refitting the model. Details can be found at (https://sfcheung.github.io/semfindr/articles/casewise_scores.html).
- Support samples with missing data. LOO can be used with methods such as full-information maximum likelihood (Arbuckle, 1996).
- Support multiple-group models.

Researchers can visit https://sfcheung.github.io/semfindr/ for illustrations of these options.

Conclusion

Sensitivity analysis is important for interpreting results in published studies. However, it has not received the attention it deserves in the applications of SEM. We

For Review Only

hope *semfindr* can make it more accessible for researchers to do casewise sensitivity analysis.

For Review Only

References

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270–301. https://doi.org/10.1177/1094428112470848

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. Marcoulides & R. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Lawrence Erlbaum Associates.

Arbuckle, J. L. (2021). *IBM® SPSS® Amos™ 28 user's guide.*

Chalmers, R. P., & Flora, D. B. (2015). Faoutlier: An R package for detecting influential cases in exploratory and confirmatory factor analysis. *Applied Psychological Measurement*, *39*(7), 573–574. https://doi.org/10.1177/0146621615597894

Cheung, S. F., & Pesigan, I. J. A. (2023). FINDOUT: Using either SPSS commands or graphical user interface to identify influential cases in structural equation modeling in AMOS. *Multivariate Behavioral Research*, *0*. https://doi.org/10.1080/00273171.2022.2148089

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences, 3rd ed.* Lawrence Erlbaum Associates Publishers.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, *19*(1), 15–18. http://www.tandfonline.com/doi/abs/10.1080/00401706.2000.10485981

Flora, D. B., LaBrish, C., & Chalmers, R. P. (2012). Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor Analysis. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00055

Fox, J., & Weisberg, S. (2018, October 16). *An R companion to applied regression* (3rd edition). SAGE Publications, Inc.

Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Science of India*, *2*, 49–55.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8.10).

For Review Only

Pastore, M., & Altoe, G. (2018). Package 'influence.SEM'.

Pek, J., & MacCallum, R. (2011). Sensitivity analysis in structural equation models: Cases and their influence. *Multivariate Behavioral Research*, *46*(2), 202–228. https://doi.org/10.1080/00273171.2011.561068

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2). http://www.jstatsoft.org/v48/i02/paper

Tanaka, Y., Watadani, S., & Ho Moon, S. (1991). Influence in covariance structure analysis: With an application to confirmatory factor analysis. *Communications in Statistics - Theory and Methods*, *20*(12), 3805–3821. https://doi.org/10.1080/03610929108830742

Wulff, J. N., Sajons, G. B., Pogrebna, G., Lonati, S., Bastardoz, N., Banks, G. C., & Antonakis, J. (2023). Common methodological mistakes. *The Leadership Quarterly*. https://doi.org/10.1016/j.leaqua.2023.101677

Yuan, K.-H., & Zhang, Z. (2012). Robust structural equation modeling with missing data and auxiliary variables. *Psychometrika*, *77*(4), 803–826. https://doi.org/10.1007/s11336-012-9282-4

For Review Only