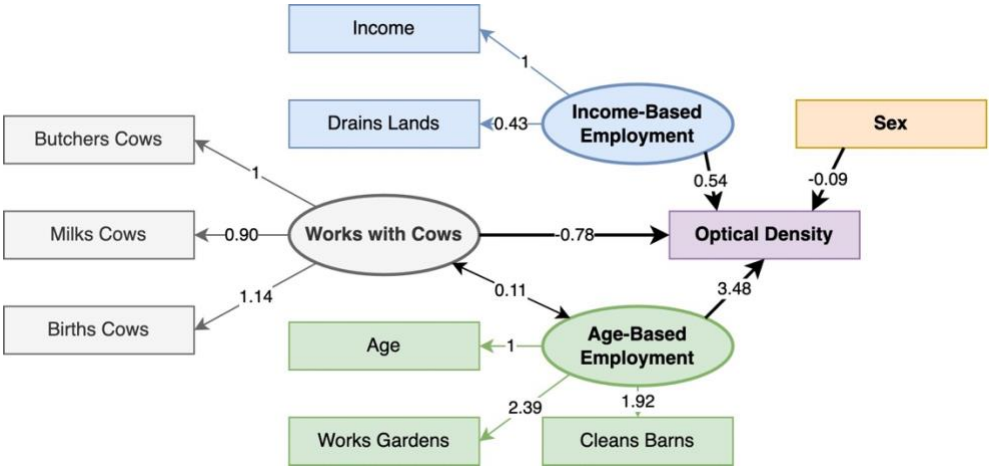
3.5 | Fit Indices for Structural Equations

Now that we have fit the structural equation model to our sample covariance matrix and investigated the significance and magnitude of the corresponding parameter estimates, the final step is to validate how successful this model was at capturing the covariability structure within our original data. Much like the communalities and variances used to validate exploratory factor analysis, there are a wide range of metrics that can be used to determine when a structural equation model provides an adequate fit to the data. Fortunately, if the model is truly effective, then all of these different metrics will lead to the same conclusion.

3.5.1 | Calculating the Chi-Square Statistic

Recall that estimates for each regression coefficient and variance parameter in a structural equation model, which are visualized using a path diagram, can be found by minimizing an objective function Q that quantifies the differences between the reconstructed and observed covariance matrices. As such, one of the simplest means of quantifying how effect the model is at capturing the covariability structure in the original data is to simply use the optimized minimum value produced by this function.

Consider again the path diagram (of the non-hidden constructs and parameters) of the optimized relationships between the set of predictor dimensions and factors and the single response dimension in our example data.



When evaluated at these optimized parameter estimates, the objective function yields a numerical value of $Q = 0.2609$. Recall that actually calculating this value would involve first representing the hidden and non-hidden optimized parameter values via two matrices of regression coefficients and a matrix of variances, then using these parameter matrices in a system of equation to reconstruct the covariance matrix, and finally comparing the difference

between the observed and reconstructed matrices using the objective function defined in Equation 3.12.

If we multiply the minimum function value by $(N-1)$, we arrive at the (hopefully) familiar *chi-square test statistic*:

$$\chi^2 = Q_{min} \times (N-1) \tag{3.13}$$

In R, this chi-square statistic is included as part of the output of the $\mathrm{sem}()$ function.

```
MOD
lavaan 0.6-11 ended normally after 113 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                        24

  Number of observations                           303

Model Test User Model:

  Test statistic                                78.786
  Degrees of freedom                                30
  P-value (Chi-square)                           0.000
```

Using the highlighted in bold value in this output, we can see that the optimized set of parameters in our structural equation model yielded a chi-square test statistic of 78.786. In order to use this test statistic to find the p-value for a hypothesis test of whether or not this model provided an accurate fit to the covariance in the original data matrix, we first need to calculate the corresponding number of *degrees of freedom*.

For a structural equation model, the degrees of freedom are equal to the difference between the number of non-fixed parameters that were estimated and the number of unique data points provided by the observed covariance matrix. Using the output above, we can see that our optimized model had 30 degrees of freedom.

Now that we have a chi-square test statistic and the corresponding number of degrees of freedom, we can use this information to find the p-value. Recall that finding a p-value involves comparing the test statistics to a theoretical or simulated *sampling distribution* derived under the null hypothesis. Using the output above, we can see that the p-value corresponding to our test statistic of 78.786 is approximately zero.

Contrary to many hypothesis tests, non-significant test statistics are preferred in structural equation modeling. Recall that a chi-square test statistic measures the difference between the observed and reconstructed values and that our goal is to create a model that accurately reconstructs the observed covariance structure. As such, a smaller chi-square statistic, which corresponds to a large p-value, indicates a better fit between the model and the original data.

Because the p-value for our model is less than the common alpha-level of 0.05, this would indicate that the optimized structural equation model does not provide an accurate fit to the covariance structure in our original data. However, it is important to note that this chi-square test statistic is highly dependent on sample size. As the sample size increases, even trivial differences between the observed and reconstructed covariance matrices can yield large chi-square test statistics that result in p-values less than the alpha-level.

Given this limitation of using the chi-square statistics and corresponding p-value to validate how successful a structural equation model is at capturing the covariability structure within the data, a number of alternative fit indices have been developed that take into account both the difference between observed and reconstructed covariances, as well as the sample size from which the observed covariance matrix was derived.

3.5.2 | Calculating Comparative Fit Indices

Although there are a wide range of different types of fit indices that can be used to assess model fit, the most common by far are known as *comparative fit indices*. Each of these indices begins by considering the specific structural equation model that was developed based on the corresponding path diagram as just one option along a continuum of possible models, each of which uses the exact same dimensions and factors but hypothesizes different paths among and between them.

On one end of this spectrum is the *independence model*, which hypothesizes absolutely no relationships or paths between the different dimensions. In such a model, each of the dimensions would serve as an independent variable, meaning that the only parameters that need to be estimated are the variances. By extension, this model would have degrees of freedom equal to the number of unique data points provided by the covariance matrix minus the number of variances (which would be the same as the number of dimensions).

On the other end of this spectrum is the *saturated model*, which hypothesizes all possible relationships between dimensions and has the fewest number of degrees of freedom possible.

The *normed fit index* or NFI evaluates the fit of a model by comparing its chi-square statistic to the chi-square statistic of the corresponding independence model:

$$\text{NFI} = \frac{\chi^2_{independence} - \chi^2_{model}}{\chi^2_{independence}} \tag{3.14}$$

Although it is certainly possible to use the `sem()` function to construct this independence model and to derive the corresponding chi-square statistic and degrees of freedom, the importance of these values have led to the development of several functions and packages in R capable of extracting them using only the original structural equation model.

In the lavaan package, we can use the fitMeasures() function to extract the independence chi-square statistic and degrees of freedom, as well as several other fit indices and metrics.

```
indices <- fitMeasures(MOD)
indices
```

```
                npar                fmin               chisq 
              24.000               0.130              78.786 
                  df              pvalue      baseline.chisq 
              30.000               0.000             456.748 
         baseline.df     baseline.pvalue                 cfi 
              45.000               0.000               0.882 
                 tli                nnfi                 rfi 
               0.822               0.822               0.741 
                 nfi                pnfi                 ifi 
               0.828               0.552               0.886 
                 rni                logl   unrestricted.logl 
               0.882           -4005.467           -3966.074 
                 aic                 bic              ntotal 
            8058.934            8148.063             303.000 
                bic2               rmsea      rmsea.ci.lower 
            8071.948               0.073               0.054 
      rmsea.ci.upper        rmsea.pvalue                 rmr 
               0.093               0.026               0.131 
          rmr_nomean                srmr        srmr_bentler 
               0.131               0.059               0.059 
srmr_bentler_nomean                crmr         crmr_nomean 
               0.059               0.065               0.065 
          srmr_mplus   srmr_mplus_nomean              cn_05 
               0.059               0.059             169.345 
               cn_01                 gfi                agfi 
             196.724               0.945               0.900 
                pgfi                 mfi                ecvi 
               0.516               0.923               0.418 
```

Using the highlighted in bold values in this output, we can see that the independence model, also referred to as the *baseline model*, has a chi-square statistic of 456.748 with 45 degrees of freedom. Recall that with ten dimensions, the corresponding covariance matrix provides 55 unique data points, and that the independence path diagram would only include ten variances, one for each of the dimensions. The difference between these two values results in the 45 degrees of freedom seen in this output.

Now that we have these values, we can either use them to manually calculate the NFI or reference the corresponding values from the output provided by the fitMeasures() function using simple vector notation.

```
NFI <- (indices["baseline.chisq"] - indices["chisq"]) /
        indices["baseline.chisq"]
NFI
```

```
baseline.chisq 
     0.8275064 
```

Ignoring the residual name above this value, we can see that our optimized structural equation model has an NFI of 0.828. Note that this is the exact same value as appears under the `nfi` heading of the `fitMeasures()` function output above.

The NFI is, by definition, bounded between zero and one, with values above 0.90 or 0.95 generally being considered as indicative of a good fit to the data.

Unfortunately, the NFI may underestimate the fit of a good model if the sample size is low. The *non-normed fit index* or NNFI accounts for this by incorporating the degrees of freedom of both the actual and the independence model:

$$\text{NNFI} = \frac{\chi^2_{independence} - \frac{\text{df}_{independence}}{\text{df}_{model}}\chi^2_{model}}{\chi^2_{independence} - \text{df}_{independence}} \tag{3.15}$$

Unfortunately, the NNFI may occasionally produce values that are outside the bound of zero to one, as well as also often underestimating the fit of a good model if the sample size is low. This issue is addressed by the *incremental fit index* or IFI:

$$\text{IFI} = \frac{\chi^2_{independence} - \chi^2_{model}}{\chi^2_{independence} - \text{df}_{model}} \tag{3.16}$$

The *comparative fit index* or CFI also assesses the fit of the model relative to that of the independence model, but takes a somewhat different approach:

$$\text{CFI} = 1 - \frac{\chi^2_{model} - \text{df}_{model}}{\chi^2_{independence} - \text{df}_{independence}} \tag{3.17}$$

Much like the NFI discussed earlier, values of NNFI, IFI, and CFI above 0.95 indicate that the structural equation model provides a good fit to the covariance structure in our original data. Using the output of the `fitMeasures()` function, we can see that these indices have values of 0.822, 0.886, and 0.882, respectively, indicating that the model does not provide an adequate fit to the data (at least when compared to a fully independent model).

Whereas these four comparative fit indices measure the improvement that the actual model makes when compared to the independence model, the *root mean square error of approximation* or RMSEA measures the detriment of using the actual model when compared to the saturated model:

$$\text{RMSEA} = \sqrt{\frac{\left(\frac{\chi^2_{model} - \text{df}_{model}}{N}\right)}{\text{df}_{model}}} \tag{3.18}$$

Unlike other comparative fit indices, lower values of RMSEA are indicative of a well-fitting model. Although there is no set limit for what value indicates that the structural equation model provides a good fit to the covariance structure in a data matrix, a common threshold is any value below 0.08.

Using the output of the `fitMeasures()` function, we can see that the optimized model for our example data has an RMSEA value of 0.073, indicating that the model does provide an adequate fit to the data (at least when compared to a fully saturated model).

Although this conclusion is contrary to that indicated by the other comparative fit indices, it is important to recognize that they use comparisons that are on opposite ends of the modeling spectrum. As such, it does occasionally occur that they result in contradictory conclusions.

It is also important to note that if the degrees of freedom exceed the chi-square test statistic of a model, as sometimes occurs if the model fit is extremely good, the RMSEA cannot be calculated because the square root of a negative number is undefined. In such situations, it is common practice to instead set the RMSEA to zero.

3.5.3 | Calculating Absolute Fit Indices

One issue with comparative indices, as the name implies, is that they only provide a comparative or relative measure of model fit by comparing the actual model to either an independence or saturated model. An alternative approach, which does not require a comparison, is provided by the *McDonald* (*and Marsh*) *Fit Index* or MFI:

$$\mathrm{MFI} = e^{-\frac{1}{2}\times\left(\frac{\chi^2_{model}-\mathrm{df}_{model}}{N}\right)} \tag{3.19}$$

Much like many of the comparative fit indices introduced earlier, the MFI is bounded between 0 and 1, with values above 0.90 or 0.95 indicating a good fit to the original covariance structure. The MFI for our example model, for instance, is 0.923, indicating that the model does provide an adequate fit to the data, at least when using a threshold of 0.90.

Another absolute fit index is the *goodness-of-fit index* or GFI:

$$\mathrm{GFI} = \frac{\mathrm{tr}\left(\widehat{\boldsymbol{\Sigma}}^{\boldsymbol{T}} \times \widehat{\boldsymbol{\Sigma}}^{\boldsymbol{-1}} \times \widehat{\boldsymbol{\Sigma}}\right)}{\mathrm{tr}\left(\boldsymbol{\Sigma}^{\boldsymbol{T}} \times \widehat{\boldsymbol{\Sigma}}^{\boldsymbol{-1}} \times \boldsymbol{\Sigma}\right)} \tag{3.20}$$

It has been proposed that the GFI may be interpreted analogously to the *coefficient of determination* or R^2 used in linear regression. For our example data, the GFI is equal to 0.945, indicating that approximately 94.5% of the variance and covariance in the original data matrix can be explained by the structural equation model.

It's important to note that unlike most fit indices used with structural equation modeling, the calculation of GFI actually changes depending on which estimation method is used to derive model parameters. The calculations in Equation 3.20, for example, are only applicable to models whose parameters were estimated under the maximum likelihood principle. Although this principle is the most common approach, several other approaches have also been proposed and are used throughout various disciplines; as such, special care should be taken when calculating this index to ensure that the proper equation is used.

Just like the coefficient of determination used in multiple linear regression, the GFI will always increase as more parameters are added to the model, reflecting the fact that adding additional relationships into a modeling framework can only increase the predictive power of the model. Much like the *adjusted coefficient of determination* or R^2_{adj}, the *adjusted goodness-of-fit index* or AGFI accounts for this by adjusting the index by the number of parameters in the model:

$$\text{AGFI} = 1 - \frac{1 - \text{GFI}}{1 - \dfrac{\text{number of parameters}}{\text{number of data points}}} \tag{3.21}$$

The AGFI follows what is known as the *parsimony principle* of constructing statistical models. Under this principle, the increased benefit (in terms of predictive power) of adding an extra parameter to a model has to outweigh the detriment of making the model more complex. The goal then becomes not just to find the model with the most predictive model, but a simpler model that still explain a high amount of the covariability structure in the original data.

3.5.4 | Calculating Parsimonious Fit Indices

There are several additional fit indices that follow the parsimony principle, the first of which is the *parsimonious goodness-of-fit index* or GFI:

$$\text{PGFI} = \left(1 - \frac{\text{number of parameters}}{\text{number of data points}}\right) \times \text{GFI} \tag{3.22}$$

As with the GFI and AGFI, higher values are indicative of a better fitting model. The PGFI, however, includes a much more stringent penalty for including additional parameters.

The PGFI for our example model, for instance, is only a moderate value of 0.516, indicating that the number of included parameters is not warranted given the total amount of covariability that the model is able to explain.

A (hopefully) more familiar parsimonious fit index is the *Akaike Information Criterion* or AIC:

$$\text{AIC} = \chi^2_{model} - 2\text{df}_{model} \tag{3.23}$$

Much like the RMSEA, smaller values of AIC are indicative of a good-fitting model. However, unlike the RMSEA or other fit indices, whether parsimonious, absolute, or comparative, the value of the AIC cannot actually be interpreted or used to determine if the model itself provides a good fit to the data. Instead, AIC is only used to determine which of a set of structural equation models provides a better fit to the data.

Our example model, for instance, has an AIC of 8058.934. If we had an alternative structural equation model, perhaps one with different relationships between factors or a different set of factors altogether (but based on the same data and covariability structure), we could compare that model's AIC to that of our example model to determine which model provides a better fit to the data. Models with lower AIC values (again, that are based off of the same data) provide a better fit to the data, although it is common practice to consider any two models that differ by less than two AIC units as providing identically good fit to said data.

Like many fit indices, AIC is sensitive to the common issue of a small sample size. The *Conditional* or *Consistent Akaike's Information Criterion* or CAIC accounts for this issue by incorporating the sample size directly into the calculations themselves:

$$\text{CAIC} = \chi^2_{model} - (\ln(N) + 1) \times \text{df}_{model} \tag{3.24}$$

Much like the AIC, the CAIC can only be used to compare two alternative or competing structural equation models, with the lower value again indicating which of the two models provides a better fit to the data. It is important to recognize, however, that just because one model provides a better fit to the data than an alternative model, it doesn't necessarily mean that the model provides a good fit to the data. To determine that, one or several of the other fit indices should be investigated.

3.5.5 | The Heuristics of Fit Indices

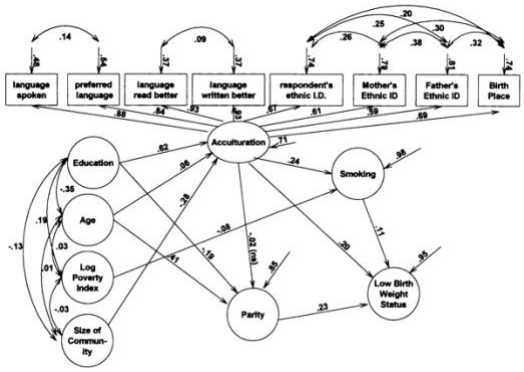
It is important to note that the different fit indices described above, as well as numerous others that exist, are heuristic in nature, which means that one is not inherently better than any other.

In an ideal scenario, multiple fit indices will yield the same conclusion, indicating that the structural equation model either does or does not provide a good fit to the covariance structure of the original data matrix. When this does not occur, as is the case with our example model, the general advice is to either report multiple indices or to report those most commonly used in the discipline or academic journal that best aligns with the analysis itself.

Remember, different disciplines may have different standards when it comes to statistical modeling and may even disagree on the role that structural equation models should play in an analysis. Part of being a good data analyst is being aware of these differences and accounting for them throughout all the stages of any analysis or modeling study you may be conducting.

Exercises for Section 3.5

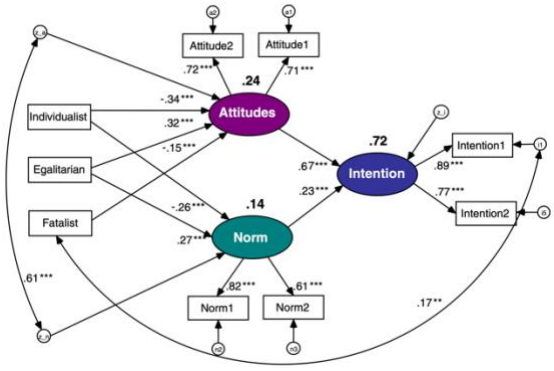
3.17 Acculturation and Low Birthweight in Infants. Acculturation, which refers to the process in which an individual adopts to a new cultural environment, has been associated with negative birth outcomes among mothers in numerous immigrant populations, including Latinas. An article published in 1996 in the *American Journal of Public Health* reanalyzed data from an earlier study to demonstrate the effect of acculturation on mothers' low birthweight status.



The path diagram above represents the structural framework for this analysis, which yielded a chi-square statistic of 199.65 with 62 degrees of freedom based on $N = 1426$ respondents.

a. Use the chi-square statistic and degrees of freedom to find the MFI for this model
b. Use the MFI for this model to explain if the model provides a good fit to the data
c. Use the chi-square statistics and degrees of freedom to find the AIC for this model
d. Use the AIC for this model to explain if the model provides a good fit to the data

3.18 Climate Change, Theory of Planned Behavior, and Values (continued). Exercise 3.05 introduced data on climate-friendly behavioral intentions and the underlying psychological processes that were published in 2011 in *Climate Change*, which are recreated below:



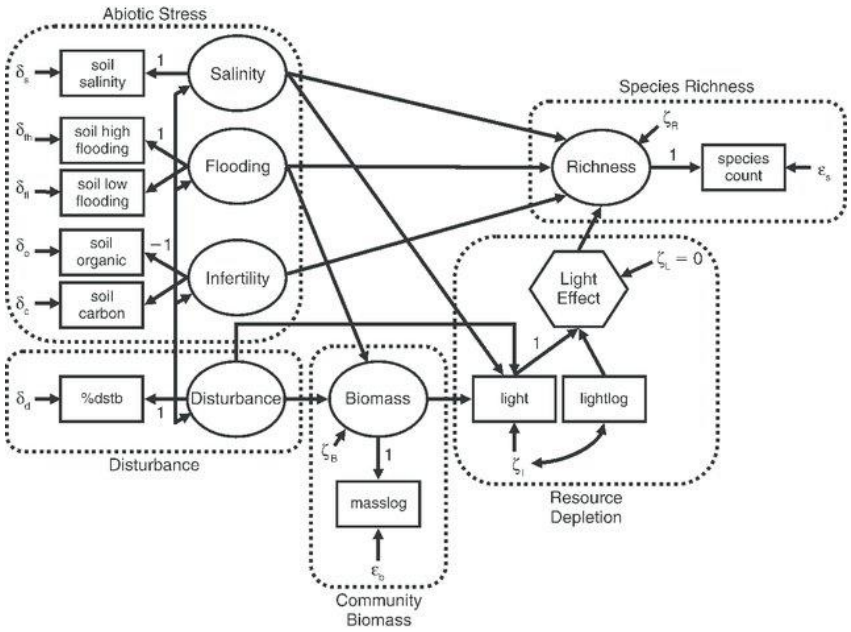
The path diagram above represents the structural framework for this analysis, which yielded a chi-square statistic of 57.646 ($DF = 21$) for the actual model and 1964.737 ($DF = 36$) for the independence model, based on $N = 3541$ staff and students at a Swiss technical university.

a. Use the chi-square statistic and degrees of freedom to find the NFI for this model
b. Use the chi-square statistic and degrees of freedom to find the NNFI for this model
c. Use the chi-square statistic and degrees of freedom to find the IFI for this model
d. Use the chi-square statistic and degrees of freedom to find the CFI for this model
e. Use these indices to explain if the model provides a good fit to the data
f. Use the chi-square statistics and degrees of freedom to find the RMSEA for this model

3.19 Stress, Disturbance, and Species Richness (continued). Exercise 3.14 introduced data on the relative importance of abiotic conditions, disturbance, and community biomass on plant species richness in a coastal wetland that were published in 1997 in *The American Naturalist*, which are recreated below:

	1	2	3	4	5	6	7	8	9	10
1. light log	1.00	.858	.667	-.251	-.699	.060	.012	.552	.547	.327
2. light	.858	1.00	.776	-.404	-.794	.157	.120	.439	.462	.321
3. %dstb	.667	.776	1.00	-.228	-.686	.218	.186	.249	.290	.216
4. species count	-.251	-.404	-.228	1.00	.291	.119	.132	-.374	-.406	-.292
5. mass log	-.699	-.794	-.686	.291	1.00	-.096	-.071	-.426	-.466	-.138
6. soil carbon	.060	.157	.218	.119	-.096	1.00	.973	-.170	-.150	.249
7. soil organic	.012	.120	.186	.132	-.071	.973	1.00	-.211	-.188	.244
8. soil low flooding	.552	.439	.249	-.374	-.426	-.170	-.211	1.00	.959	.073
9. soil high flooding	.547	.462	.290	-.406	-.466	-.150	-.188	.959	1.00	.052
10. soil salinity	.327	.321	.216	-.292	-.138	.249	.244	.073	.052	1.00

The standardized covariance matrix above was calculated using data on $N = 190$ field plots throughout coastal marsh communities of the northern Gulf of Mexico.



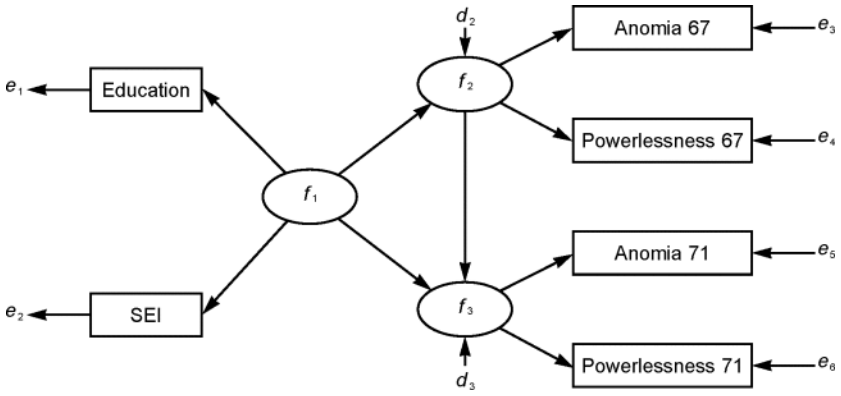
The path diagram above represents the direct effects of four predictive factors on a single response factor, as illustrated in an article published in 2010 in *Ecological Monographs*.

a. Construct this covariance matrix in R with the proper row and column names
b. Use the `sem()` function to define a structural equation model based on this covariance matrix and the non-hidden elements in the corresponding path diagram (note that the arrows from `LightEffect` to `light` and `lightlog` are incorrectly reversed)
c. Use the `fitMeasures()` function to find the MFI for this model
d. Use the MFI for this model to explain if the model provides a good fit to the data
e. Use the `fitMeasures()` function to find the GFI for this model
f. Interpret the GFI for this model in the context of the covariability structure in the data

3.20 The Stability of Alienation (continued). Exercise 3.13 introduced data on how feelings of alienation change over time and how this change is influenced by socioeconomic factors that were published in 1977 in *Sociological Methodology*, which are recreated below:

	Anomia 67	Powerlessness 67	Anomia 71	Powerlessness 71	Education	SEI
Anomia 67	11.834	6.947	6.819	4.783	-3.839	-2.190
Powerlessness 67	6.947	9.364	5.090	5.028	-3.889	-1.883
Anomia 71	6.819	5.090	12.532	7.495	-3.841	-2.175
Powerlessness 71	4.783	5.028	7.495	9.986	-3.625	-1.878
Education	-3.839	-3.889	-3.841	-3.625	9.610	3.552
SEI	-2.190	-1.883	-2.175	-1.878	3.552	4.503

The covariance matrix above was calculated using data on $N = 932$ people from Illinois.



The path diagram above represents the structural framework for this analysis.

a. Construct this covariance matrix in R with the proper row and column names
b. Use the `sem()` function to define a structural equation model based on this covariance matrix and the non-hidden elements in the corresponding path diagram
c. Use the `fitMeasures()` function to find the RMSEA for this model
d. Use the RMSEA for this model to explain if the model provides a good fit to the data