

Independent Errors

The Durbin-Watson test should be consulted in order to evaluate evidence of non-independence of errors. We want the Durbin-Watson statistic to be within the bounds of 1 and 3 (as shown below).

- 0-1 indicates potential problems with positive auto-correlation, leading to increased Type I errors for individual predictors in our model.
- DW values of 3-4 indicate potential for Type II errors.

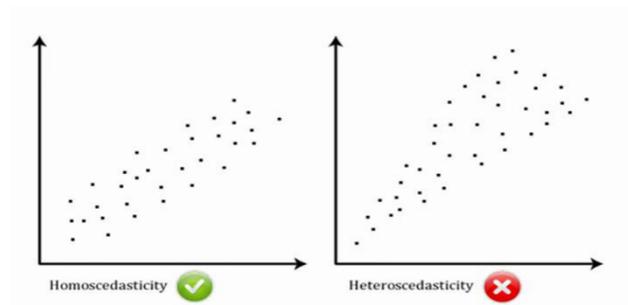
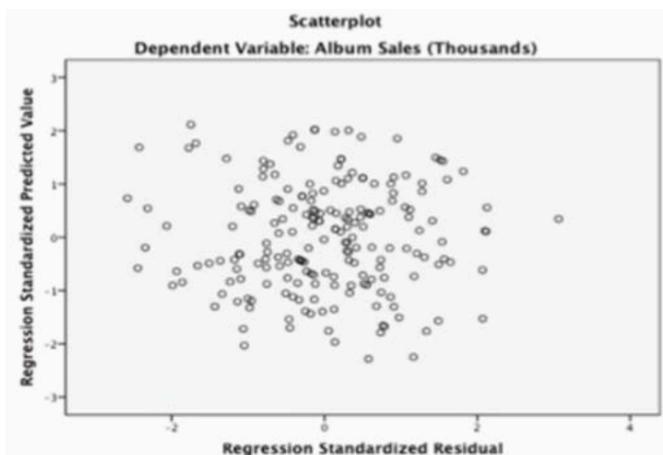
Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.787 ^a	.619	.575	.63674	2.063

Homoscedasticity

Variance of residual should be constant at each level of the predictor(s).

To check for homoscedasticity ask SPSS for a scatterplot of ZResid and ZPred which should be evenly spread



Singularity/multicollinearity

Singularity reflects a perfect correlation between two variables which means one of the variables is redundant ($r = 1$, or $r = -1$)

Multicollinearity indicates a strong, though imperfect relationship between two variables → one variable is not quite redundant, but very close.

In MR, we are concerned with these issues between IVs only.

Statistically S/MC wreak havoc:

- Results become less stable
- Standard errors become inflated
- Consider removing offending variable(s)

Theoretically, these variables are redundant:

- Variable 2 won't add much beyond variable 1

Tucker Lewis Index (Non-normative Fit Index, NNFI)

- The covariance matrix of the null model subtracted by the covariance matrix generated by the hypothesised model
- Penalises model complexity → effects *df*

$$TLI = \frac{X^2/df(\text{Null Model}) - X^2/df(\text{Hypothesised Model})}{X^2/df(\text{Null Model}) - 1}$$

Interpreting Fit Indices

CFI and TLI values range from 0-1.

- Above .90 = an acceptable fit
- > .95 = good fit
- Eg. A CFI or TLI of .93 would mean that the researchers hypothesis model is 93% better in its fit of the sample data compared to the null model

RMSEA values approaching 0 are desired

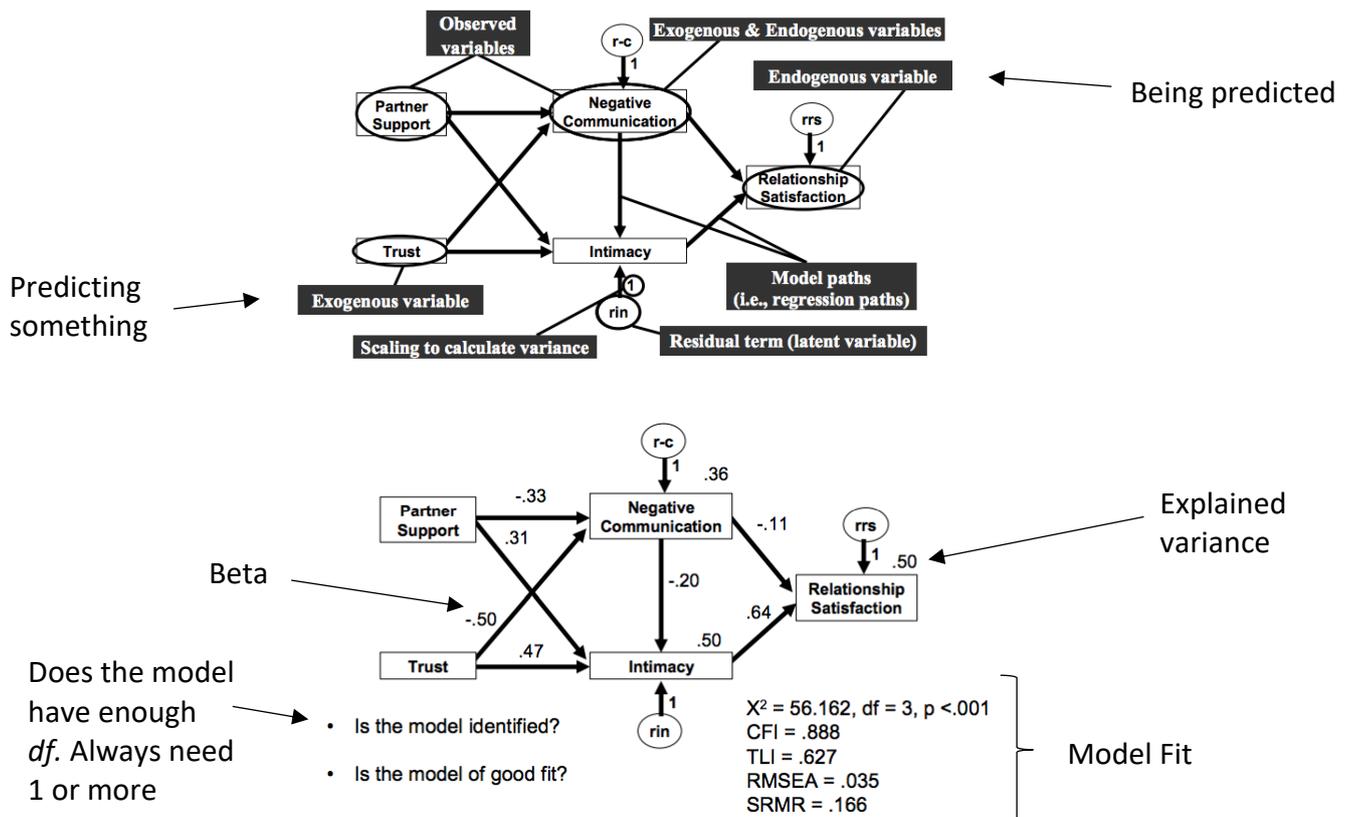
- Values .08 or less suggest reasonable approximation of error

SRMR values approaching 0 are desired

- Standardised values .06 and less represent a well-fitting model with little error

Smaller values = better fitting model

*For models of small sample size, values up to RMSEA of .11 and SRMR .08 are regarded as acceptable, however, **it is best to avoid working with small sample sizes.**



Why We Need to Know Factor Analysis

- 37.5% of Australian Journal of Psychology rely on factor analysis or multivariate designs (FA on steroids)

What Is Factor Analysis

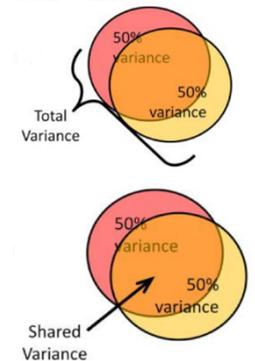
There are two broad kinds of factor analysis, FA and PCA aim to reduce a set of variables into a smaller set of dimensions (called factors in EFA and components in PCA).

PCA – estimates components that account for 100% of **TOTAL** variance in variables.

- Does not account for error/mistakes

Factor Techniques – estimate components that account for 100% of **SHARED** variance between variables

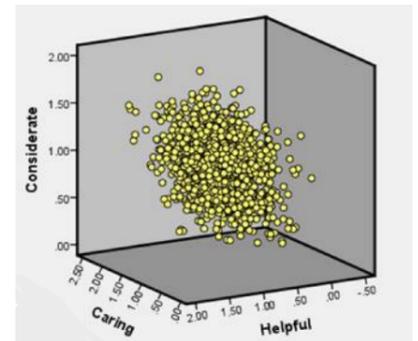
- Least Squares → minimises the difference between the data and the FA model
- Maximum likelihood → finds the most probable FA



What are Dimensions, Eigenvectors and Eigenvalues?

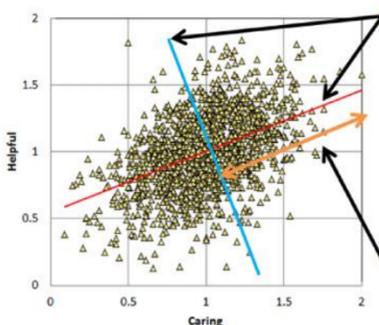
Dimensionality

- Each variable in a set of data contributes a dimension to the analysis
- Dimensionality = number of variables (k)
- Each dimension is uncorrelated to each other dimension
- Three variables = 3 dimensions
- Dimensions tell us how many variables contribute to the data set



Eigenvectors and Eigenvalues

- Factors are the relationship between two or more variables (a correlation over multiple variables)
- Each factor has a direction = an Eigenvector
- An Eigenvector has an amount of variance



Eigenvector is the direction of a relationship

$$\text{Eigenvalue} = \sqrt{\text{Radius}}$$



Radius is the distance from the centre to the edge

Eigenvalue is the variance
Total Eigenvalue = number of variables

- The sum of all Eigenvalues (Ev) = number of variables (k)
- The average Eigenvalue always = 1
- Eigenvalues bigger than 1 are better than average