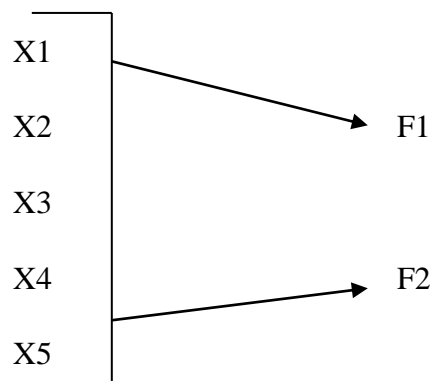


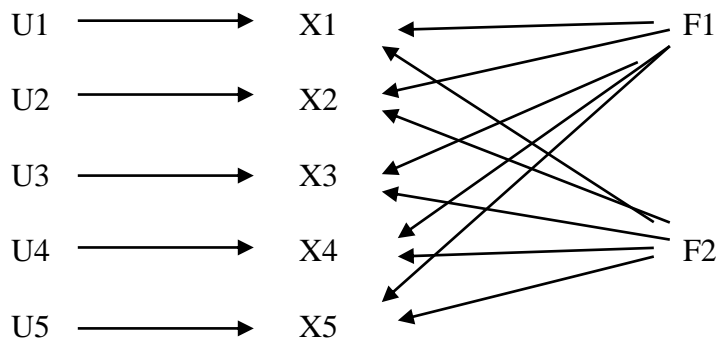
Neuendorf
Factor Analysis

Assumptions:

1. Metric (interval/ratio) data
2. Linearity (in the relationships among the variables)
 - Factors are linear constructions of the set of variables (see #8 under Statistics)
 - The critical source of info for the factor analysis is typically the correlation matrix among all variables...and correlations are linear tests
3. Univariate and multivariate normal distributions
4. There are no dependent variables specified. . .
 Instead, the model is:



Or, in terms of basic measurement theory, we could model it as:



Meaning that each variable is composed of the common variance it shares with other variables (F1 and F2) and its own unique variance that is not shared (e.g., U1). The proportion of a variable's variance that is common is called its communality. (This is the same as $1 - U$.)

Decisions to make:

1. Intent: Exploratory vs. Confirmatory (this is essentially SPSS vs. AMOS)

2. Factor extraction (how factors represent variables):

Principal Component Factoring (the most typical):

- factors derived on the basis of total variance for each variable (both the variance accounted for by F's and U's (unique variances) above).
- uses the "regular" correlation matrix.
- more widely used than common factoring.

Common Factoring:

- factors derived on basis on common variance only (variance accounted for by F's, but not the U's above).
- uses a correlation matrix with estimated communalities, not 1.0s, in diagonals.

There are several options for common factor extraction--principal axis factoring (most typical), unweighted least squares, generalized least squares, maximum-likelihood, alpha, and image (see an SPSS manual or SPSS "help" for more info.).

3. Rotation:

Rotating the factors leads to greater interpretability of factors; since there is no true DV or criterion, there is no reason not to rotate in factor analysis. Rotation will maximize "simple structure"--where each item hopefully has a high loading on one factor and low loadings on all others. There are two rotation options:

Orthogonal Factors

- factors will be completely uncorrelated (orthogonal=at right angles to each other).
- SPSS provides 3 types--varimax (by far the most typical), quartimax, and equamax (your book calls this equimax). See Hair p. 115 for comparisons.

Oblique Factors

- factors are allowed to correlate (but don't have to)
- SPSS provides two types--direct oblimin (most typical) and promax.

4. How many factors?

A. Theory (a priori).

B. Latent root criterion--eigenvalues of 1+ (a factor must account for at least the amount of variance associated with one hypothetical variable).

C. Scree test--plots eigenvalues sequentially. Look for a "natural break" between the "mountain" and the "scree".

Statistics:

1. Remember that the interitem correlation matrix is the basis, the "raw material," for the typical factor analysis. Two statistics help assess whether the items are indeed correlated enough to proceed with factor analysis. Be reminded of Hair et al.'s and others' caution that large samples and a large number of items will contribute to strong and significant statistics here:
 - A. Bartlett's test of sphericity--indicates whether the correlation matrix is significantly different from an identity matrix (a matrix with all 1's on the diagonal, 0's everywhere else). Bartlett's calculation is based on a chi-square transformation of the determinant of the correlation matrix. Statistical significance is good here (note that this is the only statistical significance test in the Factor Analysis output!). It tells us that this set of measures is intercorrelated enough so as to be appropriate for factor analysis.
 - B. KMO (Kaiser-Meyer-Olkin) measure of sampling adequacy (MSA)--assesses whether your sample of items (not people) is adequate by comparing r's with pr's. We look for large values for the KMO (approaching 1.0). Note that Hair et al. call the statistic "MSA" while SPSS calls it "KMO." Although there is no significance test, Hair (p. 104) lists some pretty fun rules of thumb for the stat (i.e., .80 or above is "meritorious"; .70 or above is "middling"; .60 or above is "mediocre"; .50 or above is "miserable"; below .50 is "unacceptable").
2. Coefficients predicting variables from factors--found in the rotated component matrix in SPSS for orthogonal rotation, and in the pattern matrix in SPSS for oblique rotation. The β 's as below:

$$X1 = \beta_1 F1 + \beta_2 F2 + \beta_3 F3$$

$$X2 = \beta_4 F1 + \beta_5 F2 + \beta_6 F3$$

$$X3 = \beta_7 F1 + \beta_8 F2 + \beta_9 F3$$
 etc.
3. Factor loadings, i.e., correlations between the variables and the factors--in orthogonal rotation, these are the same as the coefficients predicting variables from factors, in oblique they are not, and are contained in the structure matrix in SPSS. In order for an item to have "loaded" on a factor, Hair says a .30 coefficient is minimal (others use a .50 cutoff). Hair (p. 117) provides a chart for assessing the statistical significance of each loading, something that is *not* provided by SPSS.
4. Community (h^2)--the total amount of variance a variable shares with all factors (and, therefore, the amount it shares with all other variables in the factor analysis). In an orthogonal rotation, the community is the sum of all squared loadings for one variable. Rules of thumb are actually few, but Costello and Osborne (2005) do indicate a .40 as a minimum.

h^2 = how much (%) of the variable's variance is explained by all the factors;
 $1-h^2$ = what is unique to the variable.

e.g., $\beta_1^2 + \beta_2^2 + \beta_3^2$ from #2 above would be the communality of X1, assuming orthogonal rotation. Here, the communality is comparable to $R^2_{X1 \cdot F1F2F3}$.

5. Eigenvalue--the sum of all squared loadings for one factor.
 Eigen = the amount (not expressed as a %) of the variables' total variance that is accounted for by one factor.

e.g., $\beta_1^2 + \beta_4^2 + \beta_7^2$, etc. from #2 above would be the eigenvalue for F1.

6. Proportion of total variance accounted for by F1 =

$$\frac{\text{eigenvalue for F1}}{k}$$

where k=# of variables

7. Proportion of common variance accounted for by F1 =

$$\frac{\text{eigenvalue for F1}}{\text{SUM of eigenvalues}}$$

8. Factor score coefficients (betas predicting factors from variables-- β 's as below; these are not the same as loadings or coefficients predicting variables from factors, shown in #2 above). The factor score coefficients are found in the component score coefficient matrix in SPSS. These values are used by SPSS to create "factor scores"--values on new scales created for each factor. These new scales must be saved in the FACTOR procedure if you want to use them for further analysis. There are three options--we typically use the "regression method." SPSS will name the new scales automatically and place them at the end of the data set.

$$F1 = \beta_{1a} X1_z + \beta_{2a} X2_z + \beta_{3a} X3_z + \dots + \beta_{ka} Xk_z$$

$$F2 = \beta_{1b} X1_z + \beta_{2b} X2_z + \beta_{3b} X3_z + \dots + \beta_{kb} Xk_z$$

$$F3 = \beta_{1c} X1_z + \beta_{2c} X2_z + \beta_{3c} X3_z + \dots + \beta_{kc} Xk_z$$

9. Correlations among factors--for oblique rotation only! Found in the component correlation matrix, this indicates how much the various factors (and therefore their scales, if created) relate to one another in linear fashion.

References

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9. Accessed from: <http://pareonline.net/getvn.asp?v=10&n=7>

[Kim, J.-O., & Mueller, C. W.](#) (1978). *Factor analysis: Statistical methods and practical issues*. Newbury Park, CA: Sage Publications.

WORKSHEET--Making Sense of Factor Loadings

For this exercise:

- Assume all measures are 0-10 indicators of “enjoyment” of different TV/film genres.
- Assume an orthogonal rotation.
- Don't forget that loadings must be squared before adding to get communalities and eigenvalues.
- Notice that the loadings are *not* ordered by size. Such an ordering makes it easier to interpret the matrix. SPSS *will* do this if you click “Sorted by size” under Factor → Options.

Label:	Factor Loadings:			Communalities	Hypothetical
	FACTOR 1	FACTOR 2	FACTOR 3	at 3	at 10
	_____	_____	_____	Factors:	Factors:
X1--News	.25691	.24989	.80751	.78	1.0
X2--Westerns	.61936	.18150	.10120	.43	1.0
X3--Action	.61937	-.01767	.32450	.49	1.0
X4--Weepies	.01612	.75095	-.10217	.57	1.0
X5--Game shows	-.62921	.09360	.24968	.47	1.0
X6--Soaps	-.42346	.62972	-.00855	.58	1.0
X7--Talk shows	-.54741	.41574	.01161	.47	1.0
X8--Comedies	.48948	.31404	-.45143	.54	1.0
X9--Horror	.74784	.18525	-.05745	.60	1.0
X10--Cop shows	.68488	.12134	-.06675	_____?	1.0
Eigenvalues:	2.98	1.39	1.05	sum=_____?	sum=10.0
% of Total					
Variance:	_____%?	_____%?	_____%?	_____%?	
% of Common					
Variance:	_____%?	_____%?	_____%?	100%	