STATISTICAL HORIZONS

Scale Construction and Development

Deborah Bandalos, Ph.D.

Upcoming Seminar:

March 10-12, 2022, Remote Seminar

Response Processes (slides by Liz Spratto)

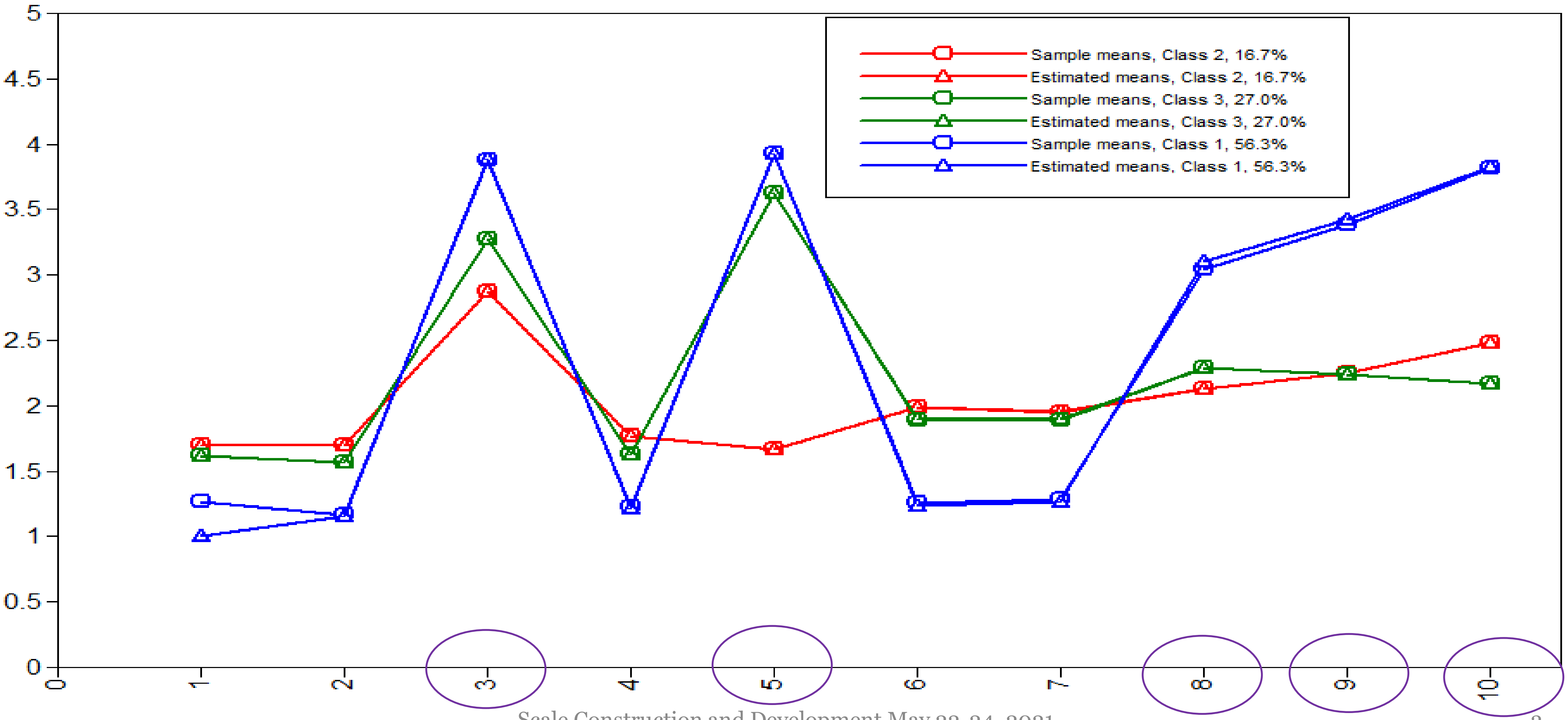
- Tourangeau and Rasinski’s (1988) 4-step process
  1. Interpret the question
  2. Retrieve necessary information
  3. Use the retrieved information to form an opinion
  4. Record the answer
  5. Some theorists include a 5th step – Edit the response
- **Opening Jurassic World after what happened in Jurassic Park was a stupid idea.**

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

Negative Keying: Effects on Factor Structure

- Inclusion of both negatively and positively keyed items often results in extra dimensionality (factors).
  - This occurs because positively and negatively keyed items are more highly correlated within sets than across sets, resulting in a correlational pattern that cannot be accounted for by a single factor.
  - This can introduce confusion when studying scale dimensionality.
- Numerous studies have shown this (here are just a few):
  - Coleman, 2013; Corwyn, 2000; DiStefano, & Motl, 2003; Hazlett-Stevens, Ullman, & Craske, 2004; Magazine, Williams, & Williams, 1996; Marsh, 1996; Motl, Conroy, & Horan, 2000; Tomás & Oliver, 1999
- Although not all:
  - Bernstein & Garbin, 1985

We did find a class (in red, below) in which respondents appeared to answer in the same way to both positively and negatively keyed items



Order Effects

- For the two items below, does it matter whether item 1 or item 2 comes first in the questionnaire?
  - In Lord of the Rings, do you think Gollum’s obsession with the Ring was beyond his control?
  - Do you think Gollum’s death in Return of the Ring was justly deserved?

- Some studies have found that item order matters, although this depends on the specific items used.

Order Effects

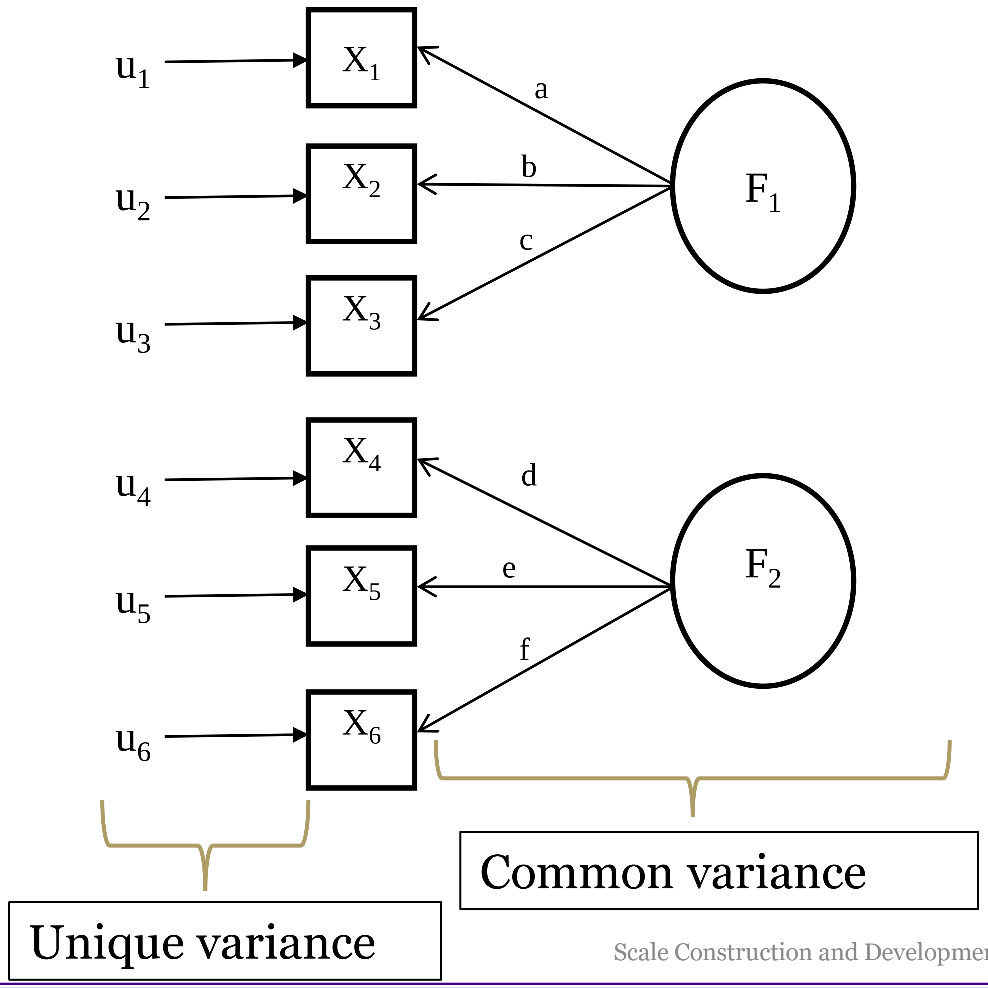
- There are two basic categories of item order effects: context or priming effects and serial order effects.
- Context effects are those induced by exposure to material earlier in the questionnaire – the item on the previous slide is an example of this.
  - As another example, suppose respondents were asked the question “Do you have fond memories of your grandparents?” and were then asked
  - “Do you support or oppose expansion of Medicare benefits for seniors?”
  - Respondents’ answers to the first question will likely influence their responses to the second.
    - This is sometimes called a *priming effect.*
  - Context effects can also result as a cumulative effect of a series of previous items.

Context Effects: Interpretation phase

- Context effects affect question interpretation in (at least) 2 ways (Torangeau & Rasinski, 1988):
- By providing an interpretive framework (Knowles & Byers, 1996 refer to this as the construct-awareness hypothesis)
  - Most applicable to questions about unfamiliar or unclear topics.
  - For example, if questions about an unfamiliar government policy are preceded by questions about inflation, respondents will likely assume that the unfamiliar policy is related to inflation.
  - This is especially true if the questions are all blocked together (Ostrom, Betz, & Skowronski, 1992; Weinberger et al., 2006).

EFA model

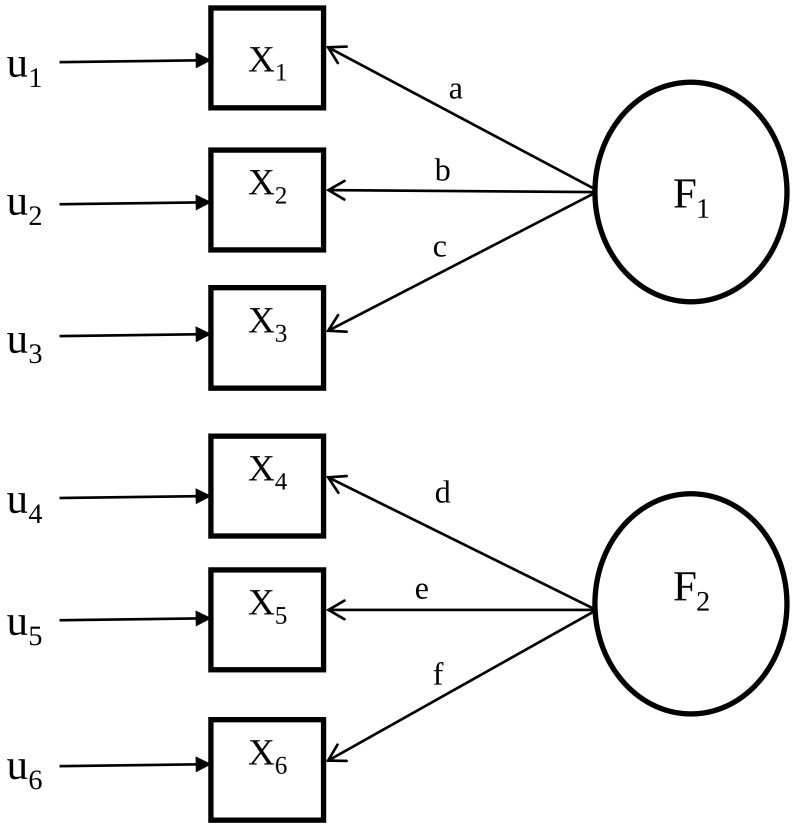
- The correlations in the previous slide can be represented by the following ***path diagram***:



F_1 and F_2 represent the two factors. These two factors are thought to account for the covariance among the variables $X_1 - X_6$.

The letters a - f label the factor coefficients or loadings; these are measures of the relation between a factor and a variable.
The variance explained by the factors is called the *common variance.*

$u_1 - u_6$ are the *uniquenesses*. These are the parts of each variables' variance that is *not* accounted for by the factor (the unique variance).

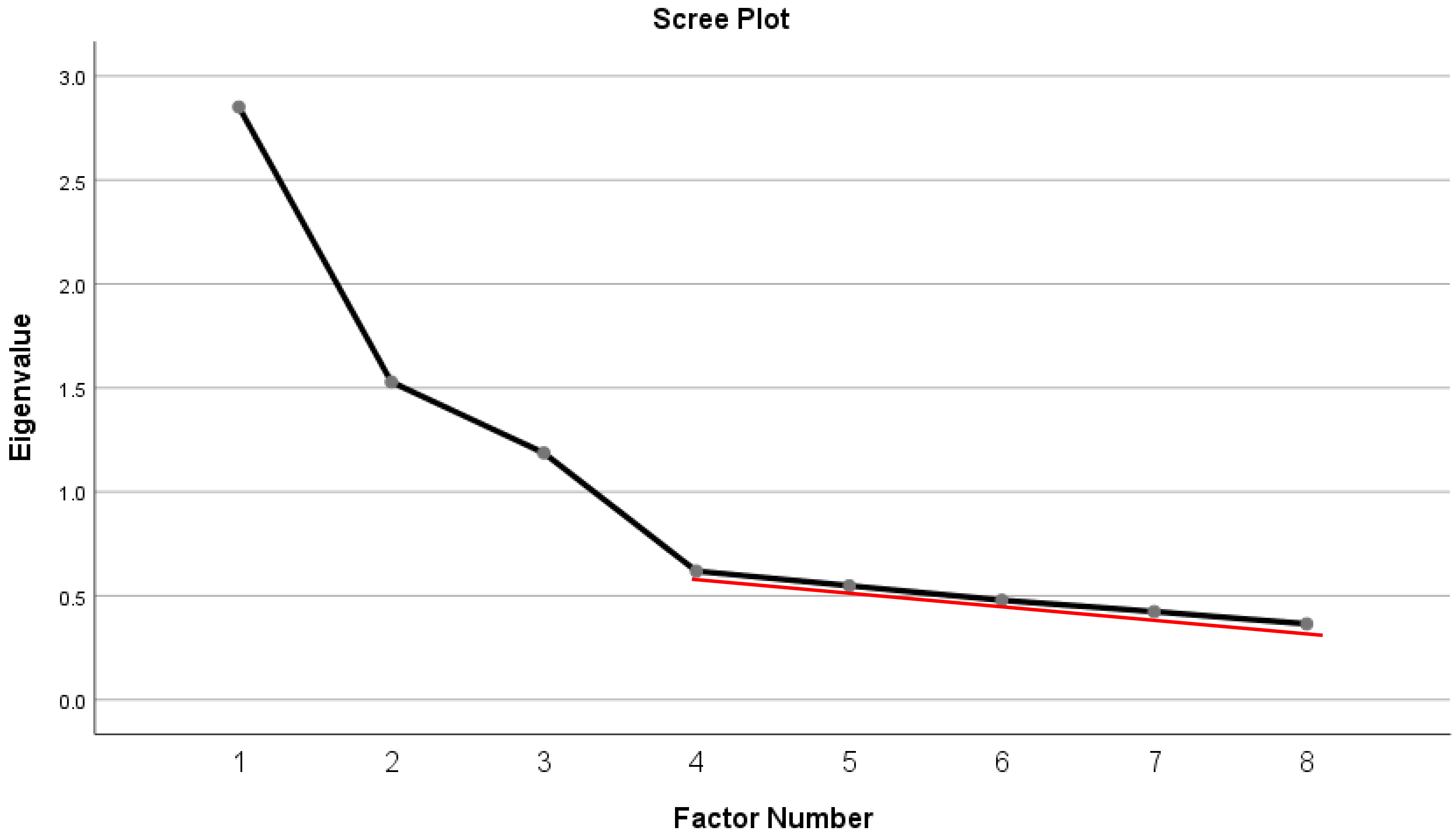


$r_{x1,x2} = a*b$ $r_{x4,x5} = d*e$

$r_{x2,x3} = b*c$ $r_{x5,x6} = e*f$

$r_{x1,x3} = a*c$ $r_{x4,x6} = d*f$

What about $r_{x1,x4}$?



Here is the scree plot for the music data.
Three factors would be extracted because the eigenvalues after the third one fall on a straight line.

- The output below is from an oblique rotation of the same variables:

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	2.851	35.641	35.641	2.401	30.009	30.009	2.187
2	1.528	19.105	54.746	1.018	12.726	42.735	1.151
3	1.187	14.838	69.583	.761	9.516	52.250	1.530
4	.618	7.731	77.314				
5	.548	6.852	84.166				
6	.479	5.981	90.147	Sum is 4.18			Sum is 4.87
7	.424	5.296	95.442				
8	.365	4.558	100.000				

Extraction Method: Principal Axis Factoring.

a. When factors are correlated, sums of squared loadings cannot be added to obtain a total variance.

- Note that percentages of variance are not given for the rotated solution.
- Also note that the eigenvalues and %s of variance no longer add to the original totals.
- This is because the variances for the factors overlap and can’t be separated (see the note under the table).

Reasons for cross-loadings

- These include:
  - The item is not a pure measure of the intended factor (note that this is not necessarily bad)
  - The two (or more) factors on which items cross-load are not distinguishable
    - In this case, the factors could be collapsed.
    - This results in a broader factor than may have been intended, but is not necessarily a bad thing.
    - Overly narrow factors may be just as problematic as overly broad factors.
    - The extent to which narrow rather than broad factors are preferred depends on the theory underlying the construct(s) being measured.

M*plus* Syntax for two model identifications

- Setting factor variances to 1.0:

```
DATA: FILE IS xxx.dat;
VARIABLE: NAMES ARE xxx ;
MODEL:
        F1 by r1* r2-r10;
        F1@1;
```

- The asterisk (*) after r1 indicates that its loading should NOT be set to 1.
- The syntax "F1@1" sets the factor variance to 1.

- The values under "estimates" are the loadings. Here I show the standardized loadings.
- The values under "S.E." are the standard errors and "Est/S.E." is the *z*-test for significance of the parameter estimate.

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 F1       BY
    X1                 0.778      0.088      8.878      0.000
    X2                 0.884      0.084     10.560      0.000
    X3                 0.839      0.085      9.813      0.000

 F2       BY
    X4                 0.892      0.085     10.440      0.000
    X5                 0.718      0.091      7.913      0.000
    X6                 0.832      0.087      9.515      0.000

 F2       WITH
    F1                 0.212      0.109      1.954      0.051

 Variances
    F1                 1.000      0.000    999.000    999.000
    F2                 1.000      0.000    999.000    999.000

 Residual Variances
    X1                 0.385      0.070      5.496      0.000
    X2                 0.208      0.064      3.261      0.001
    X3                 0.286      0.065      4.398      0.000
    X4                 0.194      0.072      2.695      0.007
    X5                 0.474      0.080      5.928      0.000
    X6                 0.297      0.072      4.155      0.000
```

Loading coefficients

Factor correlation

Factor variances

Residual variances

Reasons for Poor Fit of CFA Models

- Below I list several of the most common reasons that CFA models fail to fit well:
  - Failure to meet proportionality constraints
  - Item meaning/wording similarity
  - Positive/negative keying
  - Distribution similarities
  - Order effects
- Evidence of all of these can be sometimes be obtained through sources of evidence we discussed previously (chi-square test and other fit indices, *z*-tests of parameter estimates).
- However, ***model residuals*** and so-called ***modification indices*** are most useful in identifying these problems.

Residuals for items with similar keying

- Remember that items 3, 5, 8, 9 and 10 were negatively keyed.

Standardized Residuals (z-scores) for Covariances

	R1	R2	R3	R4	R5
R1	0.620				
R2	14.517	0.264			
R3	-2.524	0.548	0.190		
R4	6.690	11.012	0.508	0.431	
R5	-0.349	1.335	8.447	0.094	0.422
R6	-0.576	3.181	-4.161	2.649	-3.814
R7	-2.323	-0.495	-2.866	-0.028	-5.218
R8	-2.333	-6.985	-1.363	-4.636	0.468
R9	-7.270	-10.767	-1.947	-7.952	-0.304
R10	-6.077	-10.574	2.946	-8.348	-1.243

Residuals for items with similar distributions

- We've already seen from the EFA examples of items with similar distributions that this can cause extra factors to form.
- On the next slide I show some of the standardized residuals obtained by fitting a two-factor CFA model to these data.
- Recall that the odd and even numbered items had different distributions.
- Even though the correct model was fit, there are many large residuals
- There are large positive residuals for items with the same distribution.
  - This is because the items are more correlated than can be explained by the factor.
- There are large negative residuals for items with different distributions.
  - This is because the items are less correlated than can be explained by the model.

Residuals for items with similar distributions

- This problem can be remedied, to some extent, by treating the variables as categorical (which they are) in M*plus* and using the WLSMV estimator.

```
title: cfa demonstration using difficulty data;
data: file is difficulty data all.dat;
variable: names are x7-x30;
        categorical are x7-x30;
analysis: estimator = WLSMV;
model:
        F1 by x7* x8 - x18;
        F2 by x19* x20-x30;
        F1@1.0; F2@1.0;
output: sampstat residual modindices;
plot: type is plot2;
```