**PC & CFA Practice**

**#1 Dataset 🡪 talent.sav**

**Use 🡪 SES, locus of control, self concept, motivation, reading, writing, math, science, civics**

**Terms to know:**

Total vs. Common variance

Uni-polar vs. bi-polar

Univocal, multivocal & nonvocal variables

Types of factors: General, group, unique, and undefined

Also good to know the differences between types of rotation

**Anticipating the Factor Structure:**

1. How many factors do you expect to find when you factor the set of “personality” variables and "performance" variables together? Tell what variables do you expect to load on each below.

**Principal Component Analysis: Use Direct Oblimin rotation**

1. How many factors pass the λ > 1.00 criterion? \_\_\_\_\_\_
2. How much variance is accounted for by these factors? \_\_\_\_\_\_
3. What kind of variance is this? \_\_\_\_\_\_\_\_\_\_\_\_\_
4. Is there any interest in considering a 1-more-factor solution? \_\_\_\_\_\_ Why or why not?
5. Is there any interest in considering a 1-fewer-factor solution? \_\_\_\_\_\_ Why or why not?
6. How much variance is accounted for by these factors:

1-more-factor solution? \_\_\_\_\_\_\_ 1-fewer-factor solution? \_\_\_\_\_\_\_

1. What kind of variance is this? \_\_\_\_\_\_\_\_\_\_\_\_\_
2. Pick the # factor solution and use the rotated structure matrix with a cutoff of .4 to interpret each of the factors (be sure to identify the type of factor, tell the variables that “load” on each factor, if that factor is uni- or bi-polar, and give it a "name"):
3. How would the interpretation of the factors change if you used a cutoff of .3?
4. Which cutoff do you prefer and why?
5. Obtain the 1-more-factor solution and use the rotated structure matrix with a cutoff of .4 to interpret each of the factors (be sure to identify the type of factor, tell the variables that “load” on each factor, if that factor is uni- or bi-polar, and give it a "name"):
6. Do you prefer the # factor solution (from #8) or the 1-more-factor solution and why?
7. Why might someone prefer whichever you don’t like?

**Common Factor Analysis:**

1. How many factors were “kept”? \_\_\_\_\_\_
2. How much variance is accounted for by these factors? \_\_\_\_\_\_\_
3. What kind of variance is this? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
4. Using the rotated structure matrix with a cutoff of .4 to interpret each of the factors (be sure to identify the type of factor, tell the variables that “load” on each factor, if that factor is uni- or bi-polar, and give it a "name"):
5. How would the interpretation of the factors change if you used a cutoff of .3?
6. Which cutoff do you prefer and why?
7. What are the similarities and differences between the principal component analysis solution and the common factor analysis solution?
8. What is the “conceptual” difference between the results/interpretation/use of these two analyses?

**#2 Dataset 🡪 teenage.sav**

**Use 🡪 from items “i1” to “i40”. These are the 40 items of the Teenage Inventory of Social Skills**

**Sometimes factor analysis is used to understand a set of scale/questionnaire items, to help decide how many and what should be the subscales. Generally this is not a great idea – much better to test a priori hypotheses about scale structures using confirmatory factoring. But, I wanted you to get a taste of the exploratory principal components used in this way…**

**Principal Component Analysis:**

1. Run this analysis with Promax rotation of the items.
2. How many factors pass the λ > 1.00 criterion? \_\_\_\_\_\_
3. How much variance is accounted for by these factors? \_\_\_\_\_\_

Well, that got ugly quickly, didn't it?!?

Factoring individual items often leads to "messy" solutions such as this -- the whole reason we build scales is that individual items are pretty unreliable, and individually do a poor job of "tapping" the construct of interest -- so we shouldn't be too surprised with they give us unattractive factor solutions such as this. One common solution is to look at the first ?-many factors, to see if we get at least a hint of the structure we were hoping for, even though there many be additional factors in the solution.

Remember, the TISS was designed to do was to provide separate measurements of "positive" and "negative" social behaviors. So, let's look to see what the first 2 PCs look like!

1. Re-factor the items requesting specifically 2 principal components and a Promax rotation (remember the idea had been to construct 2 scales – one for positive behaviors and one for negative behaviors).
2. How much variance is accounted for by these PCs? \_\_\_\_\_\_
3. Interpret the PCs using a cutoff of .4 applied to the Promax rotated Structure matrix.

* Tell if each rotated PC is a general, group or unique factor
* Tell if each rotated PC is unipolar or bibolar
* Identify the variables that load on each rotated PC
* Try to give a name to each rotated PC

Promax rotated PC1 –

Promax rotated PC2 –

1. Identify any multivocal variables:
2. Does this factoring give support to this particular 2-factor structure for the TISS (positive vs. negative behaviors)? \_\_\_\_\_ Explain your answer:

**#3 Dataset 🡪 clinical.sav**

**Use 🡪 sepfam, findep, ruls, soss, frss, fass, stanx, tranx, dep, stress**

1. The purpose of this analysis is to help determine how many different kinds of information are obtained from these variables. Which type of “factor analysis” would you use for this purpose and why?

**Obtain a default solution with an oblimin rotation and answer each of the following questions.**

1. How many factors pass the λ > 1.00 criterion? \_\_\_\_\_\_
2. How much variance is accounted for by these factors? \_\_\_\_\_\_
3. What kind of variance is this? \_\_\_\_\_\_\_\_\_\_\_\_\_
4. Is there any interest in considering a 1-more-factor solution? \_\_\_\_\_\_ Why or why not?
5. Is there any interest in considering a 1-fewer-factor solution? \_\_\_\_\_\_ Why or why not?
6. Pick a factors solution. Decide if you prefer a .3 or a .4 cutoff for interpreting the factor solution, explain your preference, and interpret the factors:

**Common Factor Analysis:**

1. How many factors were “kept”? \_\_\_\_\_\_
2. How much variance is accounted for by these factors? \_\_\_\_\_\_\_
3. What kind of variance is this? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
4. Using the rotated structure matrix with a cutoff of .4 to interpret each of the factors (be sure to identify the type of factor, tell the variables that “load” on each factor, if that factor is uni- or bi-polar, and give it a "name"):
5. How would the interpretation of the factors change if you used a cutoff of .3?
6. Which cutoff do you prefer and why?
7. What are the similarities and differences between the principal component analysis solution and the common factor analysis solution?
8. What is the “conceptual” difference between the results/interpretation/use of these two analyses?

**#4 Dataset 🡪 success.sav**

**Use 🡪 from items “scs” to “debil”**

1. The purpose of this analysis is to help determine how many different kinds of information are obtained from these variables. Which type of “factor analysis” would you use for this purpose and why?

**Obtain a default solution with an oblimin rotation and answer each of the following questions.**

1. How many factors pass the λ > 1.00 criterion? \_\_\_\_\_\_
2. How much variance is accounted for by these factors? \_\_\_\_\_\_
3. What kind of variance is this? \_\_\_\_\_\_\_\_\_\_\_\_\_
4. Is there any interest in considering a 1-more-factor solution? \_\_\_\_\_\_ Why or why not?
5. Is there any interest in considering a 1-fewer-factor solution? \_\_\_\_\_\_ Why or why not?
6. Pick a factors solution. Decide if you prefer a .3 or a .4 cutoff for interpreting the factor solution, explain your preference, and interpret the factors:

**#5 Dataset 🡪 selfdesc.sav**

**Use 🡪 from items “index of self-esteem” to “authority behavior”.**

1. The purpose of this analysis is to identify the underlying latent traits for these 10 variables. Which type of “factor analysis” would you use for this purpose and why?

**Obtain a default solution with an oblimin rotation and answer each of the following questions.**

1. How many factors pass the λ > 1.00 criterion? \_\_\_\_\_\_
2. How much variance is accounted for by these factors? \_\_\_\_\_\_
3. What kind of variance is this? \_\_\_\_\_\_\_\_\_\_\_\_\_
4. Is there any interest in considering a 1-more-factor solution? \_\_\_\_\_\_ Why or why not?
5. Is there any interest in considering a 1-fewer-factor solution? \_\_\_\_\_\_ Why or why not?
6. Pick a factors solution. Decide if you prefer a .3 or a .4 cutoff for interpreting the factor solution, explain your preference, and interpret the factors: