Assignment 3: Scale Construction

DESCRIPTION

In this assignment, you will practice the steps involved in constructing a subjective measurement scale. This process will start with identifying a concept that you are interested in measuring. In identifying this concept, you can (but do not have to) draw on the problems you have been working on in your team project. For instance, you might be interested on measuring how much someone trusts a conversation partner that exhibits deviant behavior, how socially connected people feel when working together on a collaborative task, or how comfortable people might feel with sharing their environments with a robot. You will also identify categories along which you can later compare this concept (measured categories such as participant gender or manipulated categories such as robot type). You will not use these categories in this assignment, but will need it to complete the next assignment, which will involve statistical analysis of the data collected as part of this assignment.

In the next step, you will identify a set of survey questions using rating scales and collect data. Once you have your data, you can start exploratory factor analysis. The analysis will start by extracting factors using a Principal Component Analysis and identifying how many of these factors should be retained for the analysis. This will be followed by applying a factor rotation to your factor matrix. You will then identify candidate items to construct scales for the concept you are interested in measuring. The final step will be testing scale reliabilities.

EXPECTED OUTCOME

Because you will be collecting data, I suggest that you start early. Expected outcome is a one-page write-up that resembles an extended version of the "Measurements" section of a research paper. This section should reflect all the decisions you made and follow the style and format that an empirical research article would follow (see template). Links to sample reports of scale construction are provided on Moodle. Upload your submission following the naming convention "Assignment3-StudentLastName.zip" to the Assignment 3 submission on Moodle. Further directions are provided on the last page of this handout.

RESOURCES

See the Moodle page for resources including the Scale Construction lecture notes and the assignment submission template.

STEP 1: IDENTIFY CONCEPTS FOR EXPLORATION

The first step in constructing scales is identifying the concepts you are interested in measuring with these scales. In the context of an experiment, these concepts would act as dependent or response variables. Therefore, you are encouraged to contextualize your concepts for this assignment in the context of the experiment you designed for the previous assignment as long as measuring the concept you described seems feasible. Below are examples of concepts one might be interested in measuring:

- Trust, e.g., How much one trusts a conversational or task partner
- Social connectedness, e.g., How connected one feels to his/her social environment
- Cooperativeness, e.g., How cooperative one perceives his/her conversation or task partner

As a part of this assignment and in preparation for the next assignment, you will also find independent variables that you think might affect these concepts (categorical variables such as gender or whether someone has a pet). You will be collecting the data on these variables but will not be using this data in this assignment.

STEP 1 INSTRUCTIONS

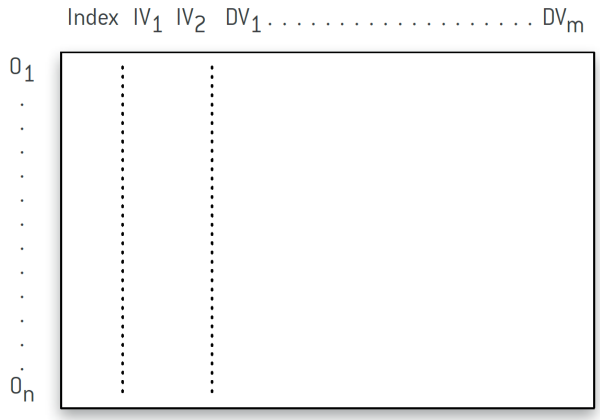
Identify three concepts (dependent variables) that you are interested in measuring and two categorical variables (independent variables) that you think might affect these dependent variables (e.g., "social connectedness" as a dependent variable and "whether one owns pets" as an independent variable). Provide in your write-up a description of these concept and variables.

STEP 2. PREPARE QUESTIONNAIRE AND COLLECT DATA

Next you will prepare a set of questions with rating scales that you think can measure the concepts you are interested in. As we discussed in the lectures on subjective measures, there is a set of principles one can follow to design good survey questions. These are also several forms of rating scales such as Likert-like scales, two-dimensional rating grids, and graphical representations of different rating levels. The lecture slides will provide you with a long list of such principles and examples of different forms of rating scales. Also, an example questionnaire is provided to illustrate the variety of questions and rating scales that can be used in a research study.

Once you create a set of questions to measure the concepts and categorical variables, you should recruit participants to provide you with data. Because you are collecting this assignment for training purposes, you should not be concerned about the representativeness of your sample or IRB approval. However, because your statistical analysis will require a certain patterned variability to produce results, do not make up data. You can collect data from other students in class, your friends, or through online markets such as Amazon's Mechanical Turk. I recommend using Google Docs to create your questionnaire as an online form. Once you create your form, you can post it on an online forum or market or share it with your friends to collect data.

After collecting your data, you should create a data matrix as shown in the lecture on scale construction. The columns of the matrix should include your variables starting with an index column followed by the categorical variables and questionnaire items. The rows should include your observations. Below is an illustration of how your data matrix should look.



STEP 2 INSTRUCTIONS

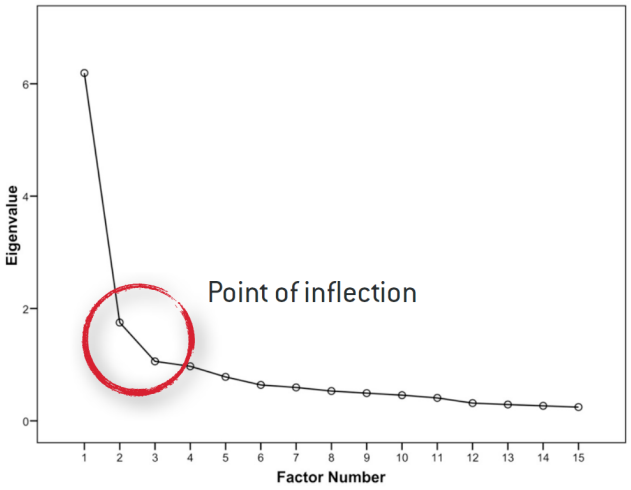
Create a minimum of six questions for each of the concepts and one question for each of the categorical variables that you are trying to measure. This should create a minimum of 21 columns in your data matrix including the index column. Attach these questions and your data matrix in your write-up.

STEP 3. EXTRACT FACTORS USING PCA AND DETERMINE THE NUMBER OF FACTORS TO RETAIN

The first step in factor analysis is to extract factors from the data using a Principal Component Analysis. The PCA will extract all the latent factors that account for the variance in the data. The amount of variance that each factor accounts for in the data is expressed as an eigenvalue (the percent of variance accounted by the individual factor/number of factors identified in data). The following R command will produce extract the factors and calculate the eigenvalues:

```
ev <- eigen(cor(data[1]))[2]
```

The PCA will most likely extract more factors than are significant to our analysis. You will have to identify a subset of these factors that might be significant for the scales you are interested in building. Which subset of the factors that you should retain in your analysis is an arbitrary decision. However, guidelines such as the Kaiser Criterion (Kaiser, 1960) and the Scree test (Cattell, 1966) are commonly used. The Kaiser Criterion suggests that only factors with eigenvalues that are larger than 1 should be retained in the analysis.



Alternatively, the Scree test provides a graphical method in which eigenvalues are plotted and visually inspected the find a "point of inflection" where the smooth decrease of eigenvalues

[1] Replace **data** with your data table (only the columns for questionnaire items) that you loaded into R.

[2] You will need to install the following R packages in order to use some of the functions you will need to complete this assignment in R:

```
install.packages("psych")
install.packages("nFactor")
install.packages("FactoMineR")
```

appears to level off to the right of the plot. The number of eigenvalues plotted to the left of this point of inflection is the number that should be retained in the analysis. It is common that both methods would suggest the same number of factors and you can use either method for this assignment. The following R commands will plot the eigenvalues in your dataset and provide you with additional visual tools to determine the point of inflection.

```
                    library(nFactors)
ap <- parallel(subject=nrow(data), var=ncol(data), rep=100,
                        cent=.05)
          nS <- nScree(ev$values, ap$eigen$qevpea)
                    plotnScree(nS)
```

References

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-51.

Cattell, R. B. (1966). The Scree test for the number of factors. Multivariate Behavioral Research, 1, 245-76.

STEP 3 INSTRUCTIONS

Extract the factors from your data and identify the number of factors using either the Kaiser Criterion or the Scree test. Provide in your write-up the eigenvalues for the extracted factors, the Scree plot, and how many factors you decided to retain in your analysis.

STEP 4. ROTATE FACTORS AND CALCULATE LOADINGS

The next step in your analysis is to calculate the loadings for each variable on each factor and apply a factor rotation that will spread the variability more evenly among variables. As reviewed in the lecture on scale construction, the loading Λ_{ji} of variable X_j on factor F_i is the correlation between that factor and the variable.

$$\texttt{corr}(F_i, X_j) = \Lambda_{ji}$$

You can use the standard R package function `factanal()` to calculate the loadings for the whole factor model and apply the factor rotation with the number of factors you identified in the previous step. Various methods of factor rotation such as varimax and promax exist. For this assignment, we will use the varimax method, which maximizes the squared loading variance across variables (as opposed to across factors in the promax method). The following R commands will calculate the loadings for the given number of factors, rotate the factors using the varimax method, and print out the factor matrix for inspection.

```
fit <- factanal(data, numberOfFactors[3], rotation="varimax")
       print(fit, digits=2, cutoff=.3, sort=TRUE)
```

The factor matrix that the the analysis should produce will look something like the following table.

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
Cheerful	0.61				0.47		
Friendly	0.91						
Warm	0.64		0.34	0.32		0.57	
Happy	0.66						0.42
Loyal		0.92					
Honest		0.60					
Cooperative		0.76					-0.55
Knowledgeable			0.67				
Intelligent			0.94				
Sensible			0.64				
Attractive		0.33		0.71			
Cute				0.70			
You_like_the_robot				0.86			
Behaves_humanlike					0.93		
Optimistic						0.70	
Responsible		0.38		0.31		0.48	0.53
Attentive							0.69
Looks_humanlike			0.32		0.42		
Robot_likes_you	0.33						

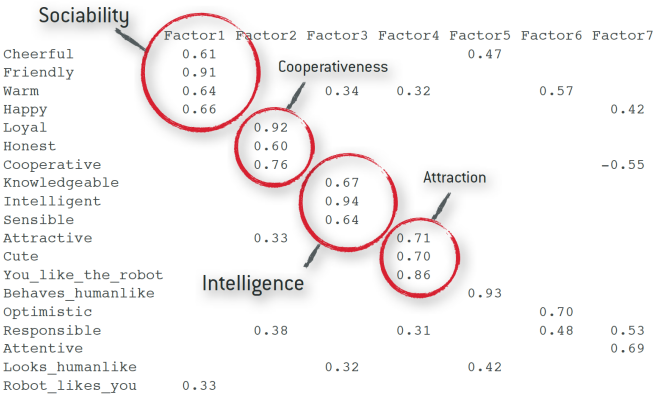
STEP 4 INSTRUCTIONS

Calculate loadings of your variables on factors and apply a factor rotation using the number of factors you identified in your previous step. Include in your write-up the factor matrix that the analysis produces.

[3] Replace **numberOfFactors** with the number of factors you decided to retain in your analysis in the previous step.

STEP 5. IDENTIFY ITEMS FOR SCALE CONSTRUCTION

The next step in your analysis is studying the loadings of each item on factors to identify which set of items might be combined to construct a scale that you think could be a measure of the concept that you identified in the first step. This is mostly a qualitative process. Even though the loadings will help you identify which items will be good predictors of a factor that your analysis revealed, you have to qualitatively judge whether these items will reasonably represent the concept you are interested in measuring. The table below shows examples of scales (provisionally) constructed from items that heavily load on the same factor:



	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
Cheerful	0.61				0.47		
Friendly	0.91						
Warm	0.64		0.34	0.32		0.57	
Happy	0.66						0.42
Loyal		0.92					
Honest		0.60					
Cooperative		0.76					-0.55
Knowledgeable			0.67				
Intelligent			0.94				
Sensible			0.64				
Attractive		0.33		0.71			
Cute				0.70			
You_like_the_robot				0.86			
Behaves_humanlike					0.93		
Optimistic						0.70	
Responsible		0.38		0.31		0.48	0.53
Attentive							0.69
Looks_humanlike			0.32		0.42		
Robot_likes_you	0.33						

What loadings are shown in the factor matrix is determined by the cutoff correlation that you used in printing your table. In the example provided above $r > 0.3$ is used as the cutoff point, which in general means good correlation. However, include in your provisional scales only the items that have a significantly high loading on the factors. $r > 0.5$ is a good cutoff point for this purpose. An alternative approach is to include all items with correlation $r > 0.3$ and eliminate them if the scale reliability test shows that these items are not reliable contributors to the scale.

STEP 5 INSTRUCTIONS

From the factor matrix, identify three provisional scales and the items that will make up these scales. Either use the $r > 0.5$ cutoff point for this or include all items above $r > 0.3$ and eliminate one by one based on the results of the scale reliability test. Provide with your write-up a list of your provisional scales and the items that make them up.

STEPS 6. TEST SCALE RELIABILITY — CRONBACH'S α

In the final step of your scale construction, you will conduct scale reliability tests for the provisions scales you constructed in the previous step. A common statistic used for how reliable a scale is a numerical coefficient called Cronbach's α (Cronbach, 1951). The α takes values between negative infinity and 1. As a rule of thumb, a reliability of $\alpha > 0.70$ is considered acceptable for scales. The following R commands will calculate the α and print a detailed report of the analysis. Here scale is a data matrix for the provisional scale. The columns of this matrix include the items that construct the scale and the rows are the observations.

```
alpha(scale, na.rm = TRUE)
```

Below is part of the report that the R command will print out. The alpha value is highlighted in bold. The report will also provide alpha values for sub-scales that would be created by removing individual items from the scale. If removing an item seems to make a significant difference and you don't think that this item is not crucial in how well your scale measures the concept of interest, you can consider removing it from your scale. Your final goal should be to arrive at a scale with an α of 0.70 or larger with individual items that contribute to the measurement in a conceptually reasonable way.

```
raw_alpha std.alpha G6(smc) average_r mean  sd
    0.81      0.81    0.79      0.52  3.5 1.0

Reliability if an item is dropped:
         raw_alpha std.alpha G6(smc) average_r
Cheerful      0.83      0.83    0.77      0.62
Friendly      0.69      0.69    0.62      0.43
Warm          0.76      0.75    0.71      0.51

Happy         0.78      0.77    0.72      0.53
```

References

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297-334

STEP 6 INSTRUCTIONS

Test your three provisional scales for reliability iteratively by including and removing items. Provide with your write-up the results of the test for each scale.

WRITE-UP

Write a one-page report of your process. The report should be a description of the decisions you made at each step of this process. Support your decisions with results from your tests in your description. This report should look like an extended version of the "Measurements" section of a research paper. Sample reports of the scale construction process are provided under "Resources" section of the assignment page on Moodle.

Your final submission should be an archive file following the naming convention "Assignment3-StudentLastName.zip" and uploaded to the Assignment 3 submission link on Moodle. Include in your submission the following materials:

- A copy of the questionnaire that you send your participants and descriptions of which question is trying to measure which concept in .docx or .pdf format[‡]
- Your full data matrix[‡]
- Your eigenvalues and Scree plot[†]
- The factor matrix[†]
- Final list of scales and items for each scale along with the α values for each scale[†]

[†] Include at the end of your report as an appendix in the same .docx or .pdf file.
[‡] Include in the .zip file.

GRADING (100 POINTS TOTAL)

- Question design, initial scale construction, and data collection (40/100)
  - Identifying concepts for measurement (13/40)
  - Designing questions and rating scales to measure concepts of interest (13/40)
  - Data collection and formatting (14/60)
- Factor analysis and scale construction (40/100)
  - Extracting factors, identifying the subset of factors to retain, calculating loadings, and factor rotation (13/40)
  - Constructing provisional scales (13/40)
  - Conducting reliability tests (14/40)
- Quality of the written report (20/100)