Measurement II

Experimental Research

Professor Yamil Velez (10/20/22)

Last class

- Introduced the concept of measurement
- Levels of measurement
  - Nominal, ordinal, interval, ratio
- Measurement scales
  - Items that collectively capture a concept of interest
    - Political engagement
      - Voting, donating, online political involvement
- Criteria for assessing scales
  - Face validity; construct validity

Measurement validity

- We can come up with several items that measure a concept of interest
- But, how do we know that the scale we have created is good or adequate?
  - Face validity — do the items seem to capture the concept?
    - Voting, donating, cat ownership
  - Construct validity — do the items relate to theoretically relevant concepts?
    - If political engagement = interfacing with politics, do all of the items seem to relate to this theoretical mechanism?

Measurement validity

- Other forms of validity
  - Concurrent validity
    - Does the measure relate to other existing measures?
    - Does your new political engagement scale predict performance on other relevant measures?
  - Predictive validity
    - Does your measure predict relevant behaviors/outcomes?
    - Does a measure comprised of self-reports of behavior predict actual voting?

Scale comparisons

- Partisan identity
  - American National Election Study (ANES)
  - 7-point scale

Question Text:

"Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?"
(IF REPUBLICAN OR DEMOCRAT) "Would you call yourself a strong (REPUBLICAN/DEMOCRAT) or a not very strong (REPUBLICAN/DEMOCRAT)?"
(IF INDEPENDENT, OTHER [1966 and later: OR NO PREFERENCE]:) "Do you think of yourself as closer to the Republican or Democratic party?"

Scale comparisons

- Multi-item scale
- Likert items

TABLE 2. Partisan Identity by Study

	Blog Study, excludes leaners	NY State Poll, includes leaners	YouGov Study, includes leaners	Student Study, excludes leaners
"How important is being a [Democrat/ Republican] to you?"				
Extremely important (%)	29.1	11.6	23.4	6.9
Very important (%)	39.8	37.7	35.3	47.5
Not very important (%)	26.9	36.6	32.7	43.8
Not important at all (%)	4.1	14.1	8.6	1.8
"How well does the term [Democrat/ Republican] describe you?"				
Extremely well (%)	23.0	11.3	18.6	6.5
Very well (%)	55.0	47.6	38.9	56.0
Not very well (%)	21.2	31.9	32.6	36.1
Not at all (%)	0.9	9.3	9.9	1.4
"When talking about [Democrats/ Republicans], how often do you use "we instead of "they"?"				
All of the time (%)	17.1	9.5	15.1	6.5
Most of the time (%)	36.6	13.8	22.3	21.7
Some of the time (%)	27.6	26.7	24.5	33.6
Rarely (%)	14.5	22.4	18.7	23.5
Never (%)	4.2	27.5	19.5	14.8
"To what extent do you think of yourself as being a [Democrat/ Republican]?"				
A great deal (%)	49.9	26.6	33.4	22.1
Somewhat (%)	39.8	44.8	35.1	53.9
Very little (%)	9.6	20.6	20.0	22.1
Not at all (%)	0.7	8.1	11.6	1.8

Question

- On what basis do you think Huddy et al. tried to make the case for their new measure of partisanship?

TABLE 3. Determinants of Past Electoral Activity

	1 Blog Study, Partisan Strength	2 Blog Study, Partisan Strength & Identity	3 NY State Poll, Partisan Strength	4 NY State Poll, Partisan Strength & Identity
Partisan strength	.26 (.07)***	.13 (.08)	.34 (.12)***	.08 (.14)
Partisan identity	–	.57 (.16)***	–	.94 (.25)***
Ideological issue intensity	.12 (.13)	.12 (.13)	.47 (.16)***	.44 (.16)***
Knowledge	.90 (.36)**	.91 (.36)**	.80 (.18)***	.81 (.18)***
Education	.91 (.12)***	.95 (.12)***	1.31 (.35)***	1.49 (.36)***
Income	–	–	.55 (.20)***	.53 (.20)**
Gender (male)	− .25 (.06)***	− .23 (.06)***	.13 (.10)	.10 (.10)
Age (decades)	.21 (.02)***	.20 (.02)***	.20 (.04)***	.18 (.04)***
White	–	–	.11 (.26)	.10 (.27)
Black	–	–	.41 (.29)	.29 (.30)
Missing race/ethnicity	–	–	.24 (.34)	.16 (.34)
Children in household	–	–	− .11 (.13)	− .11 (.13)
/Cut 1	*1.84 (.38)*	*2.13 (.39)*	*3.40 (.43)*	*3.72 (.47)*
/Cut 2	*3.11 (.38)*	*3.40 (.39)*	*4.46 (.44)*	*4.79 (.48)*
Pseudo R^2	*.06*	*.06*		
N	*1828*	*1828*	*632*	*632*

Note. Entries are ordered probit coefficients with robust standard errors in parentheses. All variables range between 0 and 1 except for age, which is measured in decades. "Other race" is the omitted category for race/ethnicity in the NY State poll. All tests of significance are two-tailed. Models in the NY State poll are based on multiple imputed values for income, which makes it difficult to provide a value for the Pseudo R^2.

* $p <$. 1. ** $p < .05$, *** $p < .01$.

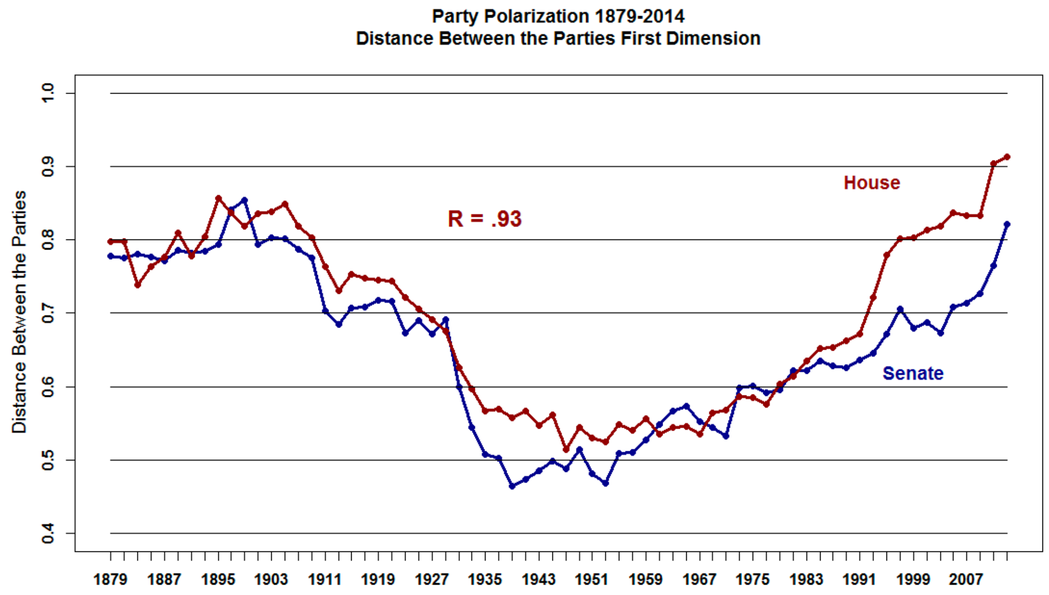
	Blog Study		Student Study		YouGov Study		
	1	2	3	4	5	6	7
	Issue Intensity	Issue Intensity & Partisan Identity	Issue Intensity	Issue Intensity & Partisan Identity	Issue Intensity	Issue Intensity & Ideological Identity	Issue Intensity, Partisan & Ideological Identity
Partisan strength	.13 (.02)***	.03 (.02)	.13 (.04)***	.03 (.04)	.21 (.04)***	.15 (.04)***	.04 (.05)
Partisan identity	–	.38 (.04)***	–	.47 (.10)***	–	–	.22 (.06)***
Ideological issue intensity[a]	.06 (.04)**	.07 (.04)**	−.05 (.08)	−.06 (.07)	.13 (.05) ***	.06 (.05)	.06 (.05)
Ideological identity	–	–	–	–	–	.19 (.04)***	.14 (.04)***
Knowledge	.25 (.09)***	.25 (.09)***	−.09 (.09)	−.08 (.08)	.30 (.05)***	.25 (.05)***	.26 (.05)***
Education	.14 (.04)***	.17 (.04)***	–	–	.23 (.04)***	.23 (.04)***	.24 (.04)***
Income	–	–	–	–	.004 (.003)	.004 (.003)	.004 (.003)
Gender (male)	−.07 (.02)***	−.06 (.02)***	−.02 (.03)	.01 (.03)	.07 (.02)***	.06 (.02)***	.07 (.02)***
Age (decades)	.02 (.01)***	.02 (.00)***	–	–	.02 (.01)**	.02(.01)***	.01 (.01)**
White	–	–	−.06 (.05)	−.09 (.05)*	−.04 (.04)	−.05 (.05)	−.05 (.05)
Black	–	–	−.11 (.07)	−.16 (.07)**	.08 (.06)	.09 (.06)	.07 (.06)
Latino	–	–	–	–	.06 (.05)	.04 (.05)	.03 (.05)
Asian	–	–	−.01 (.05)	−.03 (.05)	–	–	–
Constant	−.13 (.10)	−.33 (.10)***	.20 (.10)**	−.02 (.10)	−.43 (.07)***	−.42 (.07)***	−.43 (.07)***
Adj. R^2	*.06*	*.10*	*.07*	*.17*	*.18*	*.21*	*.22*
N	*1825*	*1825*	*212*	*212*	*933*	*933*	*933*

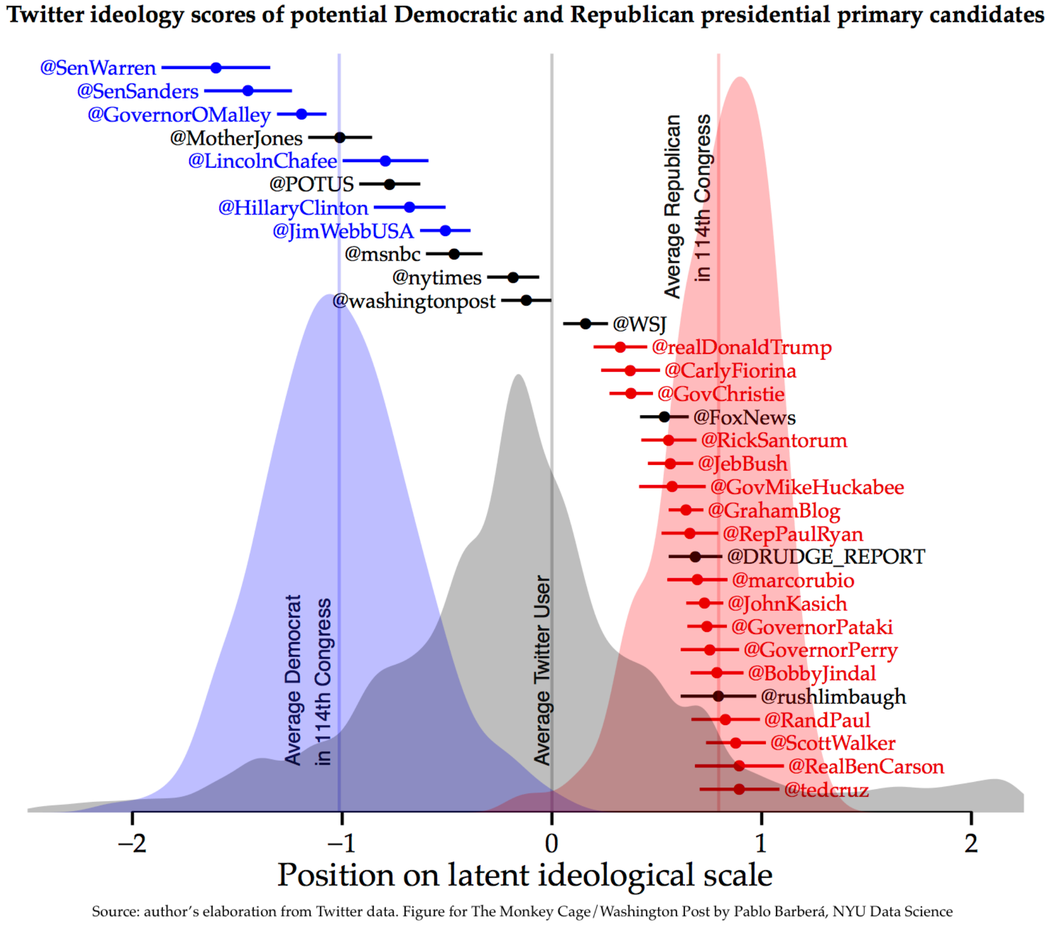
Note. Entries are OLS regression coefficients with robust standard errors in parentheses. All variables range between 0 and 1 except for age, which is measured in decades. All tests of significance are two-tailed. * $p < .1$, ** $p < .05$, ***$p <$. 01.
a. Ideological issue intensity is measured in the blog study as a folded measure of ideological consistency across three issues; in the student study it is measured as a scale created by combining three issues weighted by their subjective importance and then folded at the ideological midpoint; in the YouGov study, it is measured by combining five issues weighted by their subjective importance and then folded at the neutral midpoint. See the text for details.

Question

- Which of the following measures would you say best approximates elite-level "ideology" (e.g., whether a politician is more conservative/liberal than another)?

Question





Source: author's elaboration from Twitter data. Figure for The Monkey Cage/Washington Post by Pablo Barberá, NYU Data Science

Measurement validity

- Forms of validity such as face and content validity are inherently subjective, and conditional on people's beliefs
- How do we really know that items tap into a measure we're creating?
- How do we know we're not measuring multiple things as opposed to a single underlying concept?

Example

- 6-item scale
  - In the past week or so, have you experienced the following emotions every day, several days, one or two days, or never?
    - Irritable
    - Hot-headed
    - Cranky
    - Frightened
    - Worried
    - Nervous

Question

- Is there an underlying concept tying these items together?
- Are we measuring one or two things?
  - Negative emotions
    - All six emotions have a negative valence in terms of people experience them
  - Anger vs. anxiety
    - Anger: Irritable, cranky, hot-headed
    - Anxiety: Frightened, worried, nervous
  - Intensity of emotion
    - High intensity: hot-headed, frightened
    - Lower intensity: cranky, nervous

Moving from items to scales

- How do we adjudicate between these different representations?
- Theory
  - If you're interested in positive vs. negative emotions, you might prefer the first representation
    - Example: Do negative attack ads make people feel more positive or negatively about the candidate who is on the "offense"?
  - If you're interested in distinguishing between anger and anxiety, you might prefer #2
    - Example: How economic downturns affect specific emotional states?
  - If you're interested in intensity, #3 might be preferable

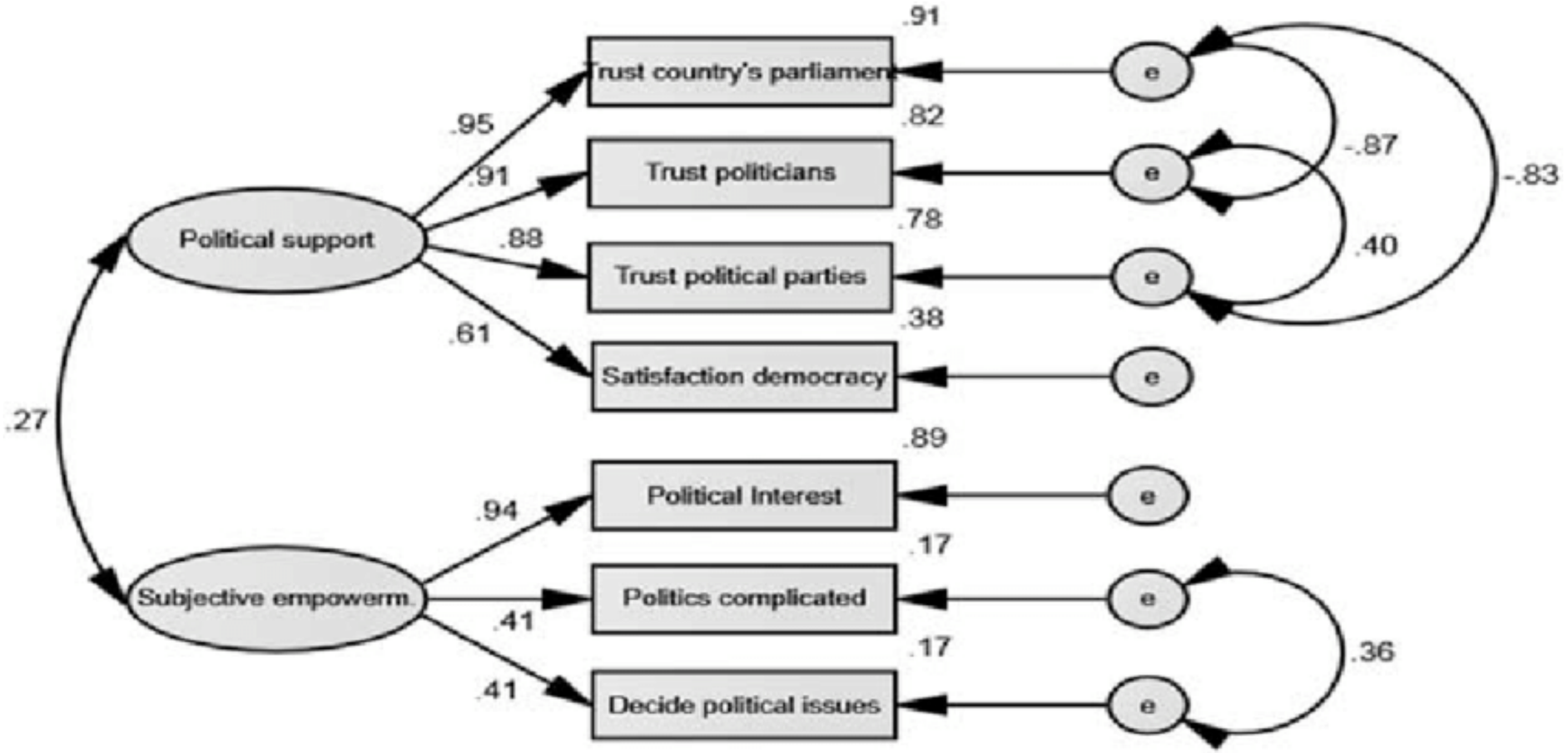
Moving from items to scales

- Data-driven strategies
  - If different measures are tapping into the same underlying concept, how should we expect them to relate to each other?
  - If voting and donating **both** tap into political engagement, should we expect no relationship, a weak relationship, or a strong relationship?
    - Should we expect high engagement people to both vote and donate and low engagement to opt out of doing both?

Moving from items to scales

- Factor analysis
  - A method that takes many measures and condenses them into a smaller set of factors (or “dimensions”)
  - In the emotions example, the dimensions could be (1) anger vs. anxiety, (2) low vs. high negativity, and (3) low vs. high intensity
  - Often referred to as a “dimension reduction” technique because it simplifies the underlying data
    - Instead of six scores to represent each survey question, a smaller set of scores to represent the underlying dimensions

Moving from items to scales



Moving from items to scales

- Factor analysis
  - Factors and items
    - Factors are the latent variables we are interested in (e.g., engagement, democracy, anger)
    - Items are measures of those latent variables
  - Estimating models to recover how items relate to these factors
    - Step 1: Define number of factors
    - Step 2: Estimate factor loadings (estimates of how well items are correlated with the factor)
    - Step 3: Estimate factor scores

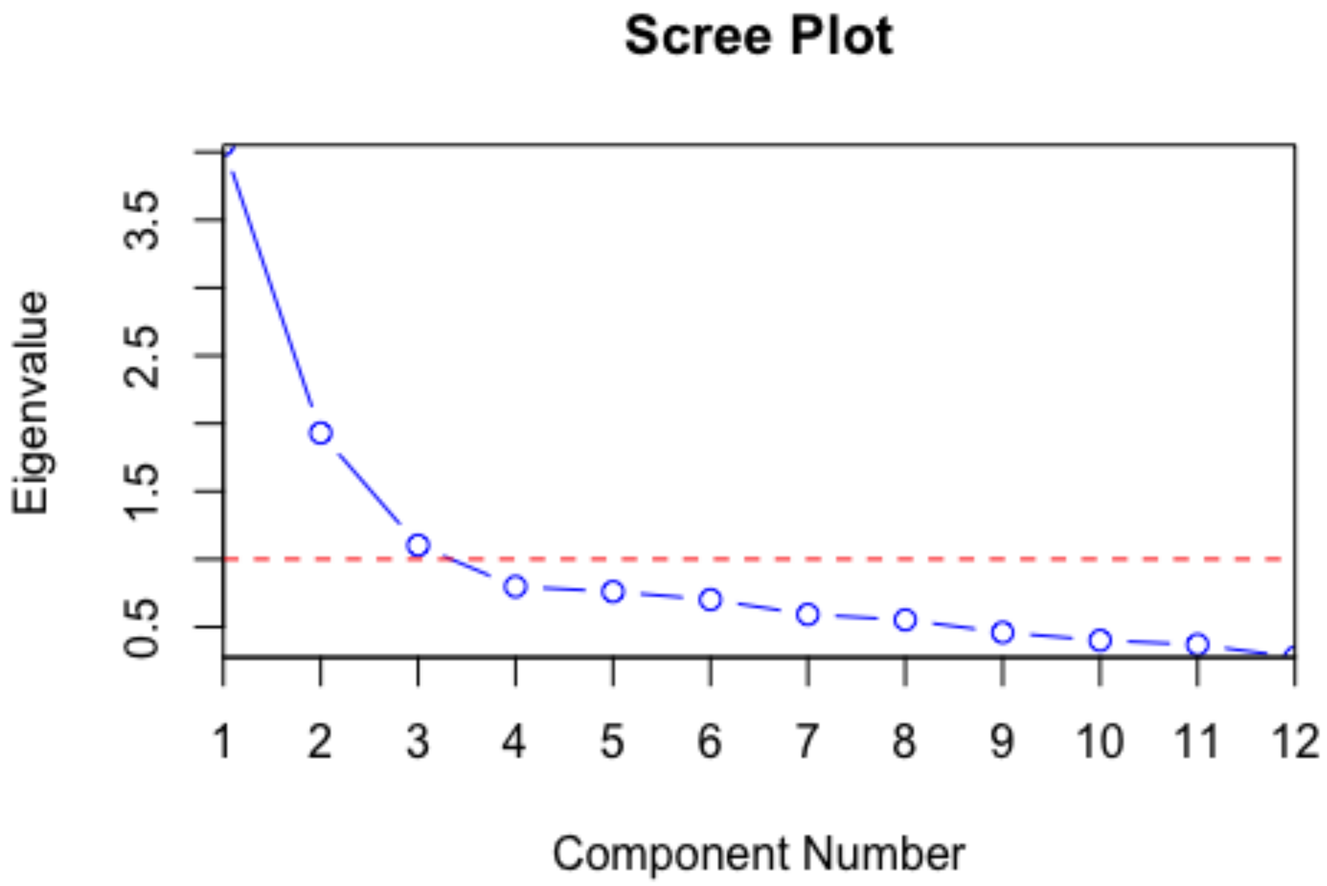
Moving from items to scales

- Step 1
  - Define number of factors
    - Theory
      - “Based on my theory of political engagement, I expect political engagement to be comprised on a single dimension: people who are engaged and those who are not.”
      - Factor = 1
    - Inductive
      - Estimate a model with different numbers of factors
      - Find out which model best fits the data

Moving from items to scales

- How do we figure out the number of factors?
  - Screeplot
    - Presents eigenvalues for each factor
    - In plain language: provides a sense of the amount of variation in items that is accounted for by each factor
    - General rule of thumb is to stop when you see diminishing marginal returns from additional factors (“elbow” rule)

Moving from items to scales



Moving from items to scales

- Step 2
  - Estimate factor loadings
    - Recall factor loadings capture how item relate to a factor
    - Scaled to run from 0-1, where 1 suggests the item is perfectly correlated with the overall factor score and 0 suggests there is no correlation
    - Higher factor loadings suggest an item is capturing the construct

Moving from items to scales

Table 1
Structure of initial themes and subthemes.

Theme	Subtheme	Example items
Latent engagement		I visited Facebook pages or profiles of politicians or public figures
		I read posts linked to a political/social cause on Facebook
Expressive forms of engagement	Public expressive forms of engagement	I posted or shared something (status, photo, meme, link,…) linked to a political/social cause on Facebook in a way it was publicly visible
	Group expressive forms of engagement	I posted or shared something (status, photo, meme, link,…) linked to a political/social cause on Facebook in a closed group
	Private expressive forms of engagement	I sent something (status, photo, meme, link,…) linked to a political/social cause to someone using Facebook Messenger
	Platform driven forms of engagement	I changed my profile picture on Facebook to support a certain political/social cause
System forms of engagement	Within system engagement	I contacted a politician, political party or public figure through a private message on Facebook
	Counter system engagement	I have spread someone's personal information (e.g. phone number, address) without their permission on Facebook in the context of a certain political/social cause

Moving from items to scales

Table 3
Exploratory factor analysis of the Social Media Political Participation Scale (n = 595).

Item	Factor 1	Factor 2	Factor 3	Factor 4
LE01	**0.876**	−0.019	0.145	0.090
LE02	**0.868**	0.002	0.080	0.068
LE03	**0.828**	0.016	0.077	0.137
LE04	**0.638**	−0.061	0.213	0.193
LE05	**0.616**	0.059	0.264	0.279
CE01	−0.014	**0.859**	0.167	0.086
CE02	0.001	**0.852**	0.078	0.022
CE03	−0.036	**0.832**	0.122	0.101
CE04	−0.033	**0.763**	0.170	0.141
FE01	0.087	0.086	**0.755**	0.167
FE02	0.296	−0.017	**0.649**	0.102
FE03	0.159	0.106	**0.600**	0.162
FE04	0.031	0.192	**0.597**	0.238
FE05	0.133	0.205	**0.586**	0.164
EE01	0.141	0.064	0.191	**0.804**
EE02	0.138	0.037	0.197	**0.784**
EE03	0.200	0.116	0.289	**0.638**
EE04	0.334	0.000	0.352	**0.469**
EE05	0.073	0.352	0.007	**0.395**
EE06	0.118	0.272	0.319	**0.375**
EE07	0.118	0.245	0.332	**0.359**

Note: *LE* Latent engagement, *CE* Counter engagement, *FE* Follower engagement, EE Expressive Engagement.

Table 4
Means and standard deviations for the items of the Social Media Political Participation Scale.

Latent Engagement (LE)		**Mean**	**SD**
LE01	I read posts linked to the climate debate	3.04	1.33
LE02	I read news articles or other kinds of information (e.g. an opinion piece) linked to the climate debate	3.10	1.30
LE03	I read comments linked to the climate debate	2.96	1.42
LE04	I watched videos linked to the climate debate	2.59	1.14
LE05	I visited pages or profiles of politicians or public figures in the context of the climate debate	2.08	1.06
Counter engagement (CE)		**Mean**	**SD**
CE01	I broke into someone's account to get information about the climate debate	1.07	0.44
CE02	I broke into someone's account to post something in their name about the climate debate	1.06	0.38
CE03	I spread someone's personal information (e.g. phone number, address) without their permission in the context of the climate debate	1.07	0.37
CE04	I did something with someone's personal information that was spread on Facebook in the context of the climate debate (e.g. sent a text)	1.08	0.38
Follower engagement (FE)		**Mean**	**SD**
FE01	I shared, spread a Facebook event or invited people for it in the context of the climate debate	1.19	0.58
FE02	I indicated that I would attend or was interested in a Facebook event in the context of the climate debate	1.54	0.91
FE03	I signed a petition on the climate debate after I saw it on Facebook	1.35	0.79
FE04	I shared a petition on the climate debate which was organised by someone else	1.19	0.60
FE05	I became a member of a Facebook group concerning the climate debate	1.20	0.61
Expressive engagement (EE)		**Mean**	**SD**
EE01	I posted or shared something (status, meme, link,…) concerning the climate debate in a closed Facebook group	1.53	0.97
EE02	I commented on something concerning the climate debate in a closed Facebook group	1.52	0.93
EE03	I liked something or reacted with an emotion on something concerning the climate debate in a closed Facebook group	1.78	1.14
EE04	I sent something relating to the climate debate to someone in a private message on Facebook/Messenger	1.94	1.12
EE05	I trolled in the context of the climate debate	1.29	0.78
EE06	I commented on something concerning the climate debate in a way it was publicly visible	1.29	0.71
EE07	I posted or shared something (status, meme, link,…) concerning the climate debate in a way it was publicly visible	1.41	0.87

Note: *LE* Latent engagement, *CE* Counter engagement, *FE* Follower engagement, EE Expressive Engagement.

Hands-on example