The University of Texas at Austin
Educational Psychology
College of Education

Unidimensional Scaling Methods

Reza Norouzian © 2020-present
The University of Texas at Austin

NOTE: I'll be using the following terms interchangeably:

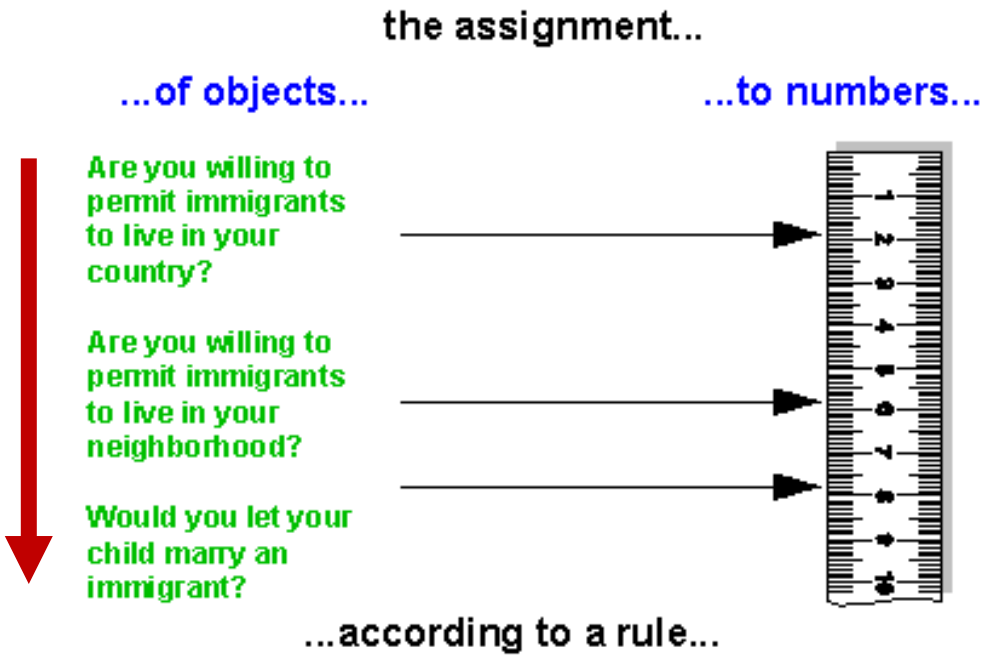
measure = scale

Scaling

- Scaling is the assignment of objects to numbers according to a rule.

EX: We would like to give a single number that describes a person's attitudes towards immigration.

Cumulative Rule
(see Guttman Scaling)

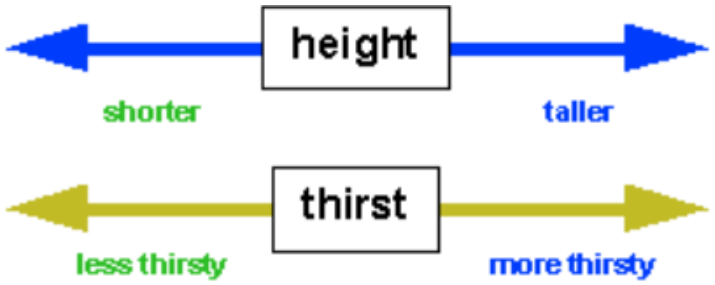


- The objects are statements of attitude or belief.
- Don't confuse scaling with response scale (Likert, dichotomous etc.)!
- Response scale is used to collect the response for an item!
- Via scaling, each item on the scale will have a ***scale value*** determined by experts not respondents.

Dimensionality

- **Uni-dimensional constructs:** Require a single ruler to be measured.

EX: Height, temperature, thirst etc.

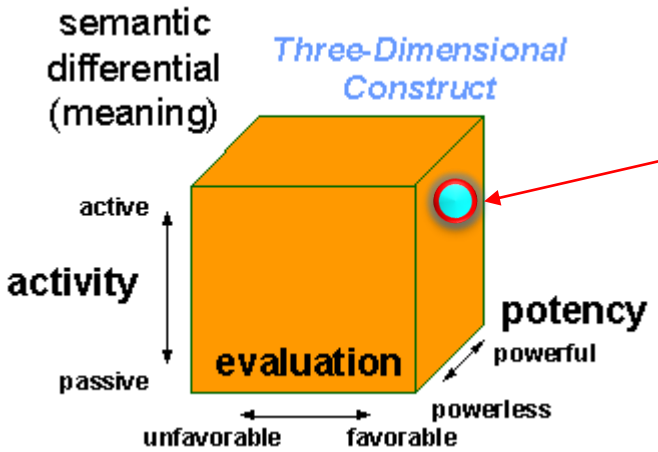


- **Multi-dimensional Constructs:** Require multiple rulers to be measured.

EX: **Semantic Differential** theory holds that phrases can be distinguished or differentiated from each other along three dimensions:

  - *activity, evaluation, and potency*

3 rulers = 3 dimensions



Meaning of "Ballet" (high on 3 dimensions)

Thurstone Scaling

- All the methods covered here are unidimensional scaling methods.

- Method of Equal-Appearing Intervals:

1. **Develop the focus:**

   EX: Describe attitudes that people might have towards persons with AIDS

2. **Generate a pool of similarly worded (e.g., in grammar, structure) statements:**

Each describing the focus (60-80 statements is desirable)

Take a look at these sample statements

3. **Rating the scale items:**

Ask potential judges ("experts") to rate each statement on a **1** to **11** scale on how much each statement indicates a ***favorable*** attitude towards people with AIDS.

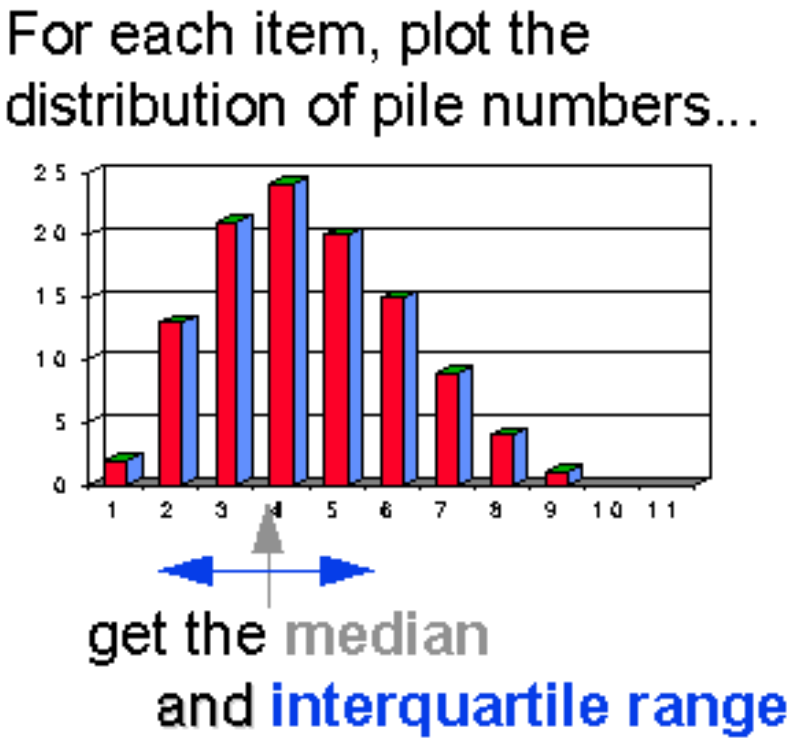
1 = "extremely unfavorable attitude towards people with AIDS" …
11 = "extremely favorable attitude towards people with AIDS".

Caution!

We are not asking judges to try our measure as if they are respondents (i.e., we are NOT measuring judges on the construct)!

We're asking judges to use their expertise to rate how strongly each item measures the construct as defined at the "focus" stage to create a ruler!

Thurstone Scaling cont,d



- **Computing scale values for each item:**

- For each statement, compute the **Median (**`scale value`**)** and the **Interquartile Range (**`IQR`**)** of the ratings from the judges.

Selecting the Final Scale Items in R

```
1.  th <- read.csv('https://raw.githubusercontent.com/rnorouzian/d/master/ths.csv')
2.  if(!require(tidyverse)) install.packages("tidyverse") ; library(tidyverse)
3.  stats <- data.frame(t(apply(th, 2,function(x) c(median(x), IQR(x)))))
4.  colnames(stats) <- c("Median", "IQR")
5.  arrange(th, Median, desc(IQR))# Median in ascending and IQR in descending order
```

- If you want to see *whole scores* for the median column (scale values) pick an odd-number of judges! Here **13** judges rated 60 items.
- Depending on how well items covered the construct, you may see a range of median sets.
- Pick one item from each unique median set (1s, 2s etc.); the one with the smallest IQR.
- You may see several similar ones, select the statement that makes the most sense.
- If, say, you find ALL items with the median of 4 to be confusing statements, skip to the next median set (e.g., 5s)

	Median	IQR
item12	1	10
item23	1	10
item46	1	9
item59	2	8
item47	3	10
item24	4	10
item48	4	10
item33	4	7
item56	4	5
item57	4	5
item60	5	10

Administering the Scale

- Randomly scatter your final items on the scale before using it.
- A respondent's responses are as follows.
- How to score this?

```
yard_stick= c(6,5,1,10,11,3,5,9)
respondent= c(6,5,  10,11,  9)
```

- 3 empty spaces = 3 disagreed
- The scoring method:

```
(score= mean(respondent))# 8.2
```

- An inappropriate scoring scheme:

```
(score=
length(respondent)/length(yard_stick)
*100) # Percent favorable toward AIDS
```

Agree	Disagree	People with AIDS are like my parents. **6** ← median scale value from judges
Agree	Disagree	Because AIDS is preventable, we should focus our resources on prevention instead of curing. **5**
Agree	Disagree	People with AIDS deserve what they got. **1**
Agree	Disagree	Aids affects us all. **10**
Agree	Disagree	People with AIDS should be treated just like everybody else. **11**
Agree	Disagree	AIDS will never happen to me. **3**
Agree	Disagree	It's easy to get AIDS. **5**
Agree	Disagree	AIDS doesn't have a preference, anyone can get it. **9**

Simulate Thurstone Scaled items (QM students)

- We could simulate a couple of Thurstone scaled items in R in a few lines of code (just a start; you need functional programming):

```
1. N <- 13  # participants or judges
2. latent = rnorm(N) # true latent construct
3. item1 = latent + rnorm(N, mean=0, sd=0.2)
4. item2 = latent + rnorm(N, mean=0, sd=0.3)

# cut items into Thurstone categories:
1. item1 = findInterval(item1, vec=c(-Inf,0,Inf))-1
2. item2 = findInterval(item2, vec=c(-Inf,0,Inf))-1
3. items <- cbind(item1, item2)
```

- As long as we're willing to assume (we often are) that the latent construct has a normal distribution, the above method can be used to simulate categories for items.

Likert Scaling

1. **Develop the focus:** As in all scaling problems e.g., see Thurstone scaling.

2. **Generate a pool of similarly worded (e.g., in grammar, structure) statements:**

   Each describing the focus (60-80 statements is desirable)

3. **Rating the scale items:**

Ask potential judges ("experts") to rate each statement on a **1** to **5** scale (very common) on how ***favorable*** each statement is as regards the construct of interest.

1 = Strongly unfavorable to the construct, . . .
5 = Strongly favorable to the construct

Selecting the Final Scale Items in R

- **Selecting the Items:**

- Throw out any items that have a low correlation with the total (summed) score across all items. This is easily done with any software e.g., R, SPSS etc.

- A rule of thumb: Eliminate items with an item-total correlation < **+0.6.**

- Let's try this with a 10-item (*space limitation*) Likert measure rated by 12 judges:

```
lk <- read.csv("https://raw.githubusercontent.com/rnorouzian/d/master/lk.csv")

if(!require(psych)) install.packages("psych") ; library(psych)

psych::alpha(lk)$item.stats # look at the column titled r.drop:
```

	n	raw.r	std.r	r.cor	r.drop
item1	12	0.93	0.92	0.93	0.89
item2	12	0.78	0.72	0.72	0.68
item3	12	0.96	0.96	0.98	0.94
item4	12	0.32	0.39	0.37	0.26
item5	12	0.83	0.84	0.83	0.80
item6	12	0.64	0.66	0.66	0.60
item7	12	0.61	0.66	0.59	0.55
item8	12	0.83	0.83	0.82	0.77
item9	12	0.73	0.73	0.73	0.65
item10	12	0.66	0.63	0.58	0.56

Check the items (wording, meaning) before elimination!

➢ What if we got a *negative* item-total correlation for an item?!

- **Could mean two things:**

1) The item runs opposite in meaning and its scoring must be reversed (*next slide*),
2) A really terrible item!

Administering the Scale

- The final score for each respondent on the scale is the ***sum*** of their Likert ratings for all of the items (for our data in R: `rowSums(lk)`).

- On some scales, you might have items that are reversed in meaning from the overall direction of the scale (**reversal items**).

- You will need to reverse the response value for each of the reversal items before summing for the total.

- For example, if the respondent gave a 1, you make it a 5; if they gave a 2 you make it a 4; 3 ⇒ 3; 4 ⇒ 2; and, 5 ⇒ 1.

- Finally, odd-number of response scales (0-4) allow for a "Neutral-type" response option (e.g., undecided, I don't know) in the middle, but if needed, you could use even-number of response scales (0-5) to make respondents pick the options with no "Neutral-type" response.

Simulate Likert Scaled items (QM students)

- We can also create a latent construct and then create items that deviate from the latent construct (just a start; you need functional programming):

```
1. N <- 5  # participants or judges
2. latent = rnorm(N) # true latent construct
3. item1 = latent + rnorm(N, mean=0, sd=0.2)
4. item2 = latent + rnorm(N, mean=0, sd=0.3)

# cut items into 5-scale Likert categories:
1. item1 = findInterval(item1, vec=c(-Inf,-2.5,-1, 1,2.5,Inf))
2. item2 = findInterval(item2, vec=c(-Inf,-2.5,-1, 1,2.5,Inf))
3. items <- cbind(item1, item2)
```

- As long as we're willing to assume (we often are) that the latent construct has a normal distribution, the above method can be used to simulate categories for items.

Guttman Scaling

- Also sometimes known as ***cumulative*** scaling or ***scalogram*** analysis.
- We aim to predict ***item responses*** perfectly, knowing only the ***total score*** for the respondent.

- Imagine a 10-item cumulative scale. If the respondent scores a 4, it should mean that he/she agreed with the *first four items*.

- But rarely does such a pattern hold in practice as regards psycho-educational constructs! But, let's follow a methodical approach to develop such a scale:

1. **Define the focus:** *U.S. citizens' attitudes towards immigration*

- Be clear about all aspects of the focus,
- EX: Here you mean any type of immigration (legal and illegal)? from anywhere (Europe, Asia, Latin and South America, Africa) etc.
- This way judges can exactly rate your items in terms of depth of attitude.

Guttman Scaling cont,d

2. Develop the Items (60-80 statements desirable**):**

1. I believe that this country should allow more immigrants in.
2. I would permit a child of mine to marry an immigrant.
3. I would be comfortable if a new immigrant moved next door to me.
4. It would be fine with me if new immigrants moved onto my block.
5. I would be comfortable with new immigrants moving into my community.
6. I would be comfortable if my child dated a new immigrant.

3. Rating the scale items:

- Ask judges ("experts") to dichotomously rate each item indicating whether the item ***is*** favorable toward immigration or it ***is not.***

- Because items are of Yes/No type, you can likely get more experts to rate your statements.

Selecting the Final Scale Items in R

4. Selecting the Items:

- Recall, the goal is to determine whether agreeing with one item can '***imply***' agreeing with other item(s) coming before it. But, how can we select and order our items on our measure to build a cumulative scale.

- For our 6-item immigration scale, agreeing with which items would imply agreeing which other items?

- Required R packages:

```
install.packages("BiocManager");library(BiocManager)
BiocManager::install("Rgraphviz"); library(Rgraphviz)
install.packages("DAKS") ; library(DAKS)
```

Selecting the Final Scale Items in R, contd.

- **30 experts have rated the 6 items (0=Yes, 1=No):**

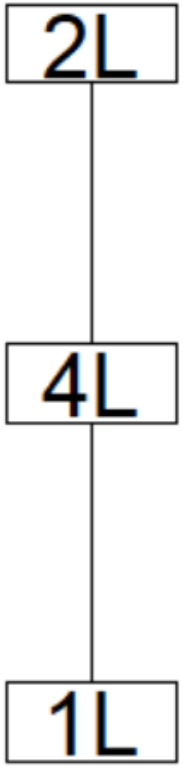
```
gut <- read.csv('https://raw.githubusercontent.com/rnorouzian/d/master/gut.csv')
```

- **3 Inductive item tree analysis (iita) algorithms:**

```
mini <- iita(gut, v = 1)
corr <- iita(gut, v = 2)
orig <- iita(gut, v = 3)
```

- **The tree diagram for all algorithms:**

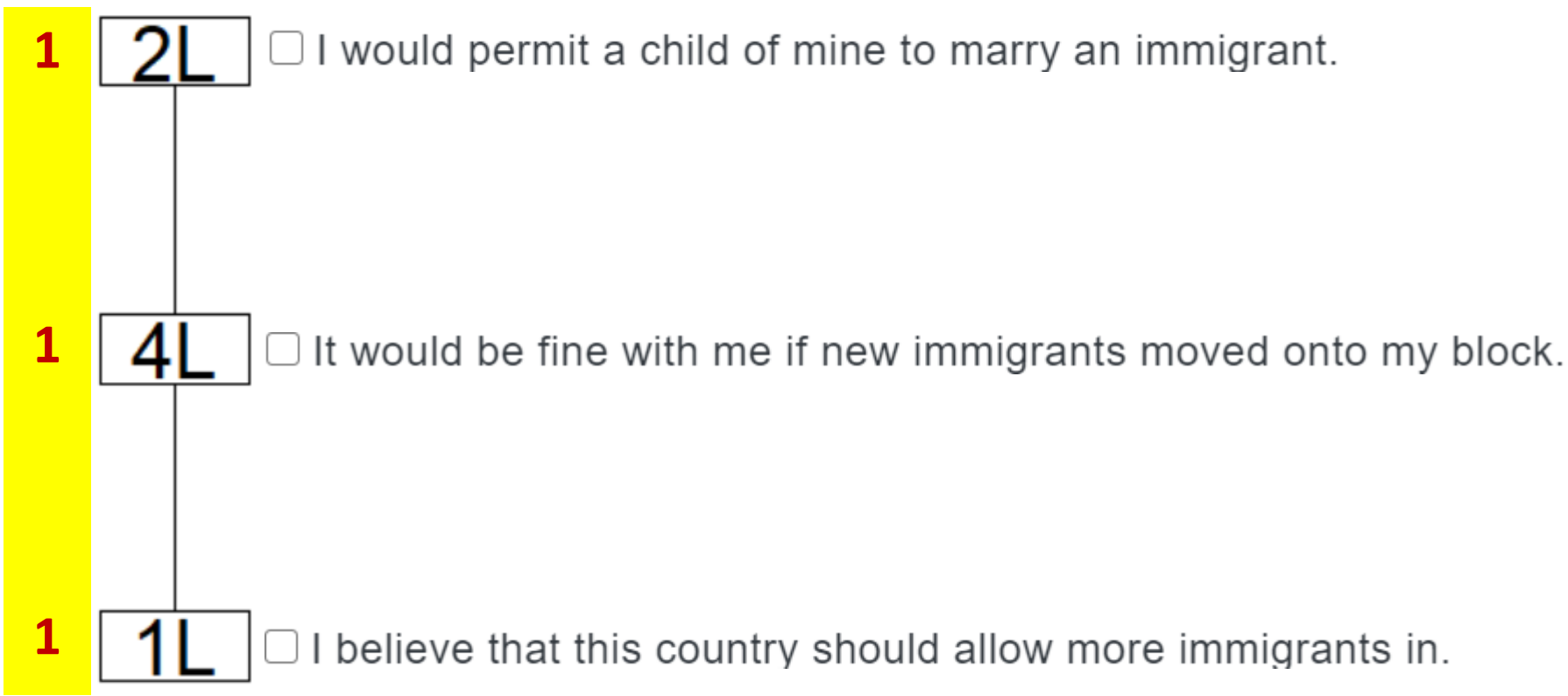
```
hasse(mini$implications, 6)
hasse(corr$implications, 6)
hasse(orig$implications, 6)
```



- Agreeing with item 2 implies agreeing with items 4 and 1.
- Items 3, 5, and 6 not shown (*see later slides*).
- You now know both the items and their order for your cumulative scale!

Administering the Scale

- **Create a hierarchical scoring:** Each respondent's score is sum of the scale values of the items they agree with. If our cumulative scale works as expected we can see a score of 2 would mean someone has agreed with items 1 AND 2 but not item 3!



Behind the Scene

- The previous plot omitted items 3, 5, 6. Let's see why?

- In R: `mini$implications`

```
{(1L, 2L), (1L, 3L), (1L, 4L), (1L, 5L), (1L, 6L), (3L, 1L), (3L, 2L),
 (3L, 4L), (3L, 5L), (3L, 6L), (4L, 2L), (5L, 1L), (5L, 2L), (5L, 3L),
 (5L, 4L), (5L, 6L), (6L, 1L), (6L, 2L), (6L, 3L), (6L, 4L), (6L, 5L)}
```

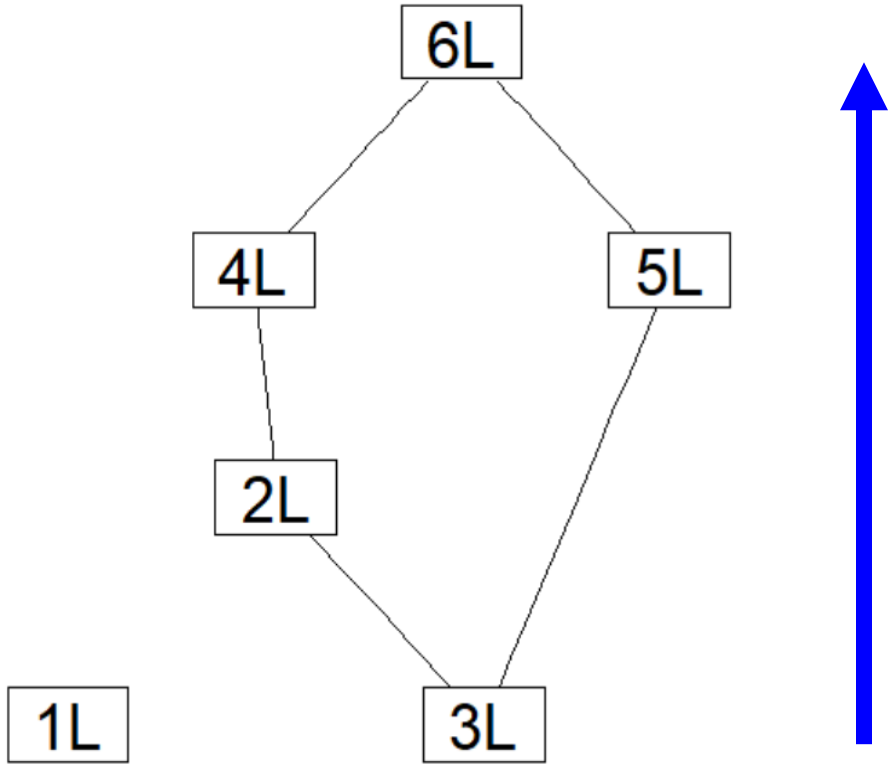
- Agreeing with ***any*** other item than 1 implies agreeing with 1.
- Agreeing with ***any*** other item than 3 implies agreeing with 3.
- Agreeing with ***any*** other item than 5 implies agreeing with 5.
- Agreeing with ***any*** other item than 6 implies agreeing with 6.

- So 3, 5, 6 are equally hierarchically structured as item 1 is and thus repetitive!
- No item implies 2 but all items are implied by 2, so item 2 is the head of the hierarchy.
- Item 4 is only implied by item 2, so must be included below 2.

```
1 implied by 2,3,4,5,6
3 implied by 1,2,4,5,6
4 implied by 2
5 implied by 1,2,3,4,6
6 implied by 1,2,3,4,5
```

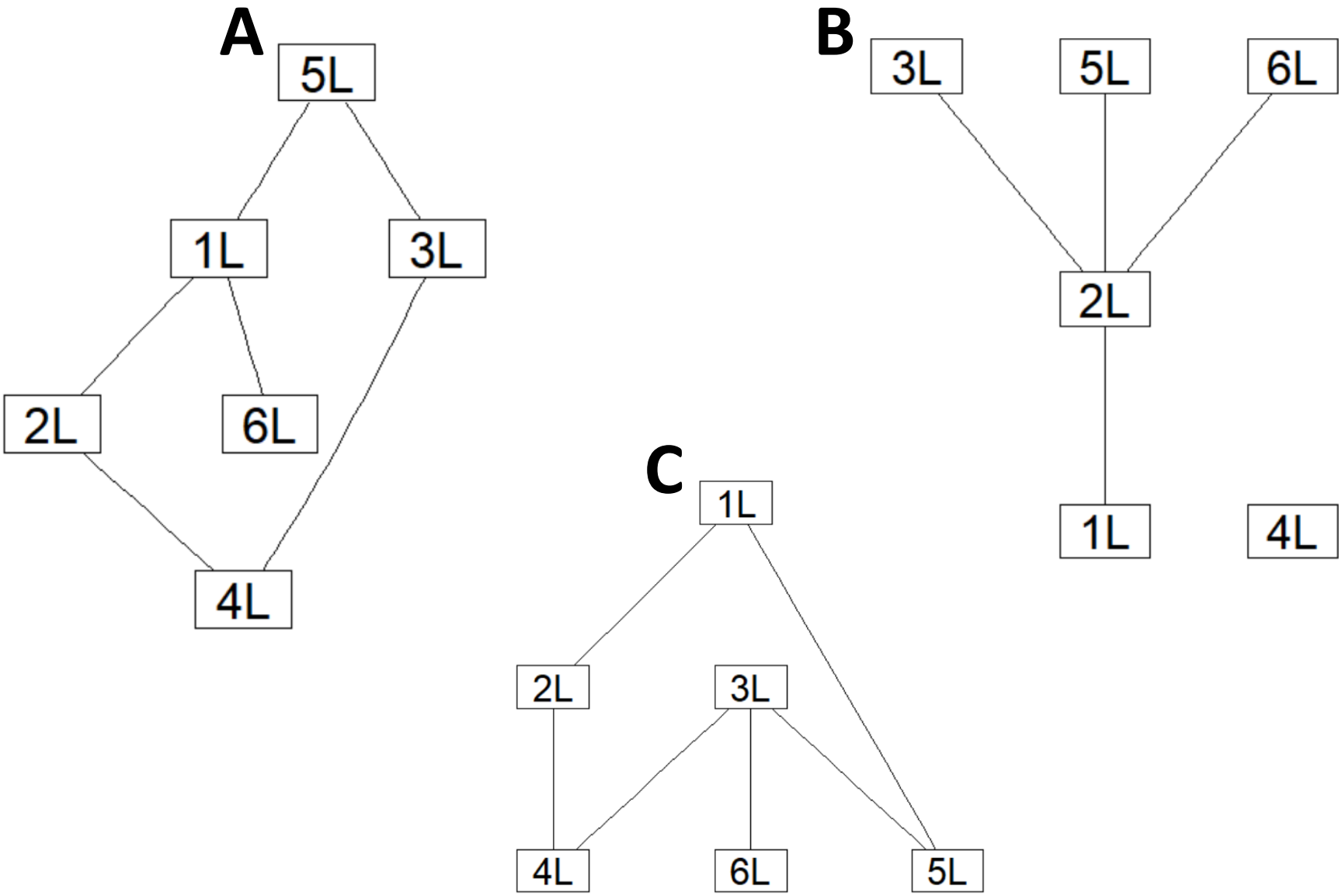
Question: Would the following make a good cumulative scale?

- Which items would you pick for your cumulative scale?



{(2L, 4L), (2L, 6L), (3L, 2L), (3L, 4L), (3L, 5L), (3L, 6L), (4L, 6L), (5L, 6L)}

A Few Examples of Other Possible Trees



Guttman's Scale

- Can you think of one practical application of Guttman's cumulative scale in your area of research/ field of study?

- Can you think of one unique benefit of Guttman's scale in general? (hint: it happened in our own demonstration)

Simulate Guttman Scaled items (QM students)

- Read this paper.