

John McPeak

PAI 705

Lecture 6

- In our goal of measuring a concept, the challenge of ensuring content validity is ensuring we have captured the all different dimensions of a concept that are important.
- Multiple observations pertaining to a concept are often required.
- When we do this, we end up with multiple indicators that capture different elements of a given concept.
- Single indicators might not capture enough of the concept. But then we face the challenge of a set of indicators that when combined have content validity with regards to the concept we have in mind.

Do people understand how index insurance works in northern Kenya?

Based on your understanding of the livestock insurance, how often do you have to pay a premium in order to remain insured?

1. once every 6 months
2. once every year
3. once every 2 years
4. you remain insured until compensation has been paid
98. don't know

Based on your understanding of the livestock insurance, when do you expect indemnity payout (compensation)?

1. Anytime animals are lost causing the household to suffer
2. At the end of each RAINY season
3. At the end of each RAINY season when rain fails
4. At the end of each DRY season
5. At the end of each DRY season when forage conditions are WORSE than normal
6. At the end of each DRY season when forage conditions are MUCH WORSE than normal
98. don't know

Based on your understanding of IBLI, what does indemnity payout (compensation) depend on?

1. the number of livestock you have lost
2. predicted number of livestock lost due to poor forage condition close to your home base camp
3. predicted number of livestock lost due to poor forage condition averaged over the normal grazing areas of your community
98. don't know

|

Can create a score of how well it is understood based on these responses. There is a right answer to each.

	Overall Score	How often do you pay to remain insured	What triggers comp.?	What determines the size of the comp.?	If you insure 10 animals and 10% loss predicted?	Premium back if no comp.?	Who pays the comp.?
Friends	-0.63 ***	-0.26 ***	-0.46 ***	-0.76 ***	-0.26	-0.16	-0.33 ***
Chief	-0.33 *	0.13	-0.24	-0.71 ***	0.39	-0.39 **	0.29 **
Meeting	0.24	0.12	0.13	0.29 **	-0.30	-0.08	0.01
Survey	0.40 ***	0.24 ***	0.00	0.27 **	0.02	0.12	0.47 ***
Game	-0.02	0.04	-0.12	0.08	0.43 ***	0.01	0.01
NGO	-0.76 *	-0.49	0.40	-0.88 *		0.01	-0.33
VIP	0.44 ***	0.12	0.32 ***	0.35 ***	-0.07	0.61 ***	0.15 *
Equity	0.72	0.05	0.14	0.57 *	0.72 **	0.06	0.49 *
Constant	0.96 ***	0.03	-0.19	0.05	-1.78 ***	-0.11	-0.86 ***

Can use this to compare different extension sources and methods.

	Code	Assets	Did you buy, receive, sell, or give away [ASSET] between October 2009 and September 2010? 1. Yes 2. No (move to the next asset)	
PRODUCTIVE ASSETS	912	Animal cart		
	913	Water drum		
	914	Plough		
	915	Wheelbarrow		
	916	Sickle		
	917	Pick axe		
	918	Axe		
	919	Hoe (Jembe)		
	920	Spade		
	921	Machete (panga)		
	1003	Qoto (small chisel)		
	1004	Animal bell		
	BASIC HOUSEHOLD GOODS	922	Charcoal/ wood stove (jiko)	
		923	Kerosene stove	
924		Leather/ stick bed		
925		Mattress		
926		Hides/skins/pelts		
927		Metal box/ trunk		
928		Pannier / kiondo		
929		Mosquito net		
930		Gourd		

- Concept of 'change in assets'.
- Scales and indexes are ordinal measures of variables.
 - Both rank order the units of analysis.
 - Both are composites of observations.
- Index is measured by accumulating scores. So my IBLI knowledge score is an index of knowledge about the insurance product.
 - The assets I could also turn into an index by adding up all the yes responses; given them a one and a no a zero, add up the number of yes responses. Yes a cart and yes an animal bell so 2 on the asset index.
 - I might also turn it into an index by, for example, assigning the local cash value of the item in question. An animal cart is worth X shillings and an animal bell is worth Y shillings so the person has a X+Y score on the asset index.
 - I might find there is a scale structure, where knowing yes to one implies yes to a set of other assets. Axe is most basic, Axe and Machete next, ..., Wheelbarrow implies all the other ones are yes.
- There may be a logical ordering that allows you to infer a scale in the intensity revealed by a given pattern.
 - a) Girls and boys learn differently in primary education.
 - b) Girls and boys require different classrooms at the primary level to be educated separately
 - c) Girls and boys require different schools at the primary level to be educated separately.

Index-Construction Logic

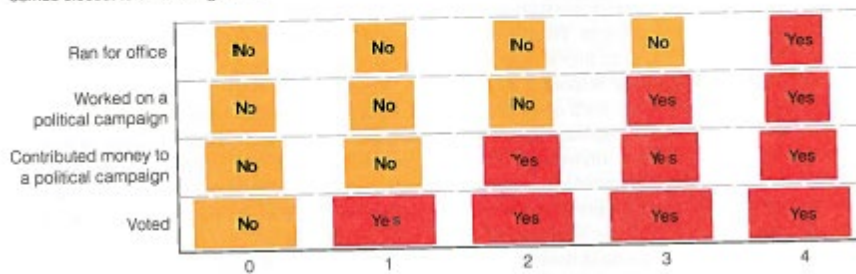
Here are several types of political actions people may have taken. By and large, the different actions represent similar *degrees* of political activism.

To create an *index* of overall political activism, we might give people 1 point for each of the actions they've taken.

**Scale-Construction Logic**

Here are some political actions that represent very different degrees of activism: for example, running for office represents a higher degree of activism than simply voting does. It seems likely, moreover, that anyone who has taken one of the more demanding actions would have taken all the easier ones as well.

To construct a *scale* of political activism, we might score people according to which of the following "ideal" patterns comes closest to describing them.

**FIGURE 6-1**

Indexes versus Scales. Both indexes and scales seek to measure variables such as political activism. Whereas indexes count the number of indicators of the variable, scales take account of the differing intensities of those indicators.

© Cengage Learning®

- In the index construction, the items are different dimensions of political activism.
- When we go to the scale, we think there is a logical ordering of probable patterns.
 - Scales attempt to capture not just different dimensions, but also some form of intensity.

- INDEX CONSTRUCTION

- Selecting items to be in an index.

- Face validity. Is the variable selected plausible as a measure of the concept you are trying to measure.

- I am interviewing in herders and farmers so my asset measures concern farming tools and animal bells.

- Unidimensionality. The measure being used can be thought of as a number line with different observations along the number line.

- Can be zero one; can be height; can be strongly disagree, disagree, neutral, agree, strongly agree.

- General or specific. Am what I am trying to capture is general, household assets in different dimensions of asset ownership, or specific, productive assets used in agriculture?

- Variance. There is going to be differentiation among the responses I get using the item in question.

- Notice we did not ask 'John Deere tractor' or 'harvester'.

- Examining the empirical relationship between the items being used to measure the concept. Does

knowing answers to one question help predict answers to another?

- Bivariate. Relationship between two variables.
 - People who know what triggers the insurance are more likely to understanding the timing of the payout for IBLI.
- Can be positive if both measuring aspects of the same thing, zero if not really measuring aspects of the same thing, negative if measuring it but in opposite directions.
 - Animal ownership and animal bell ownership move together
 - Animal bell and hoe ownership might move in different directions.
- Issues when the items don't seem to co-move
 - Are they not really measuring what you think they are measuring as a coherent concept?
 - Should you refine you concepts?
- Issues when they co- move to a degree that asking both gives you no more information than just asking one.

- The example from the text investigates the consequences of the 'scientific perspective' of doctors on the quality of care provided to patients.
- Bivariate relationships presented in figure 6-2.
 - How does a group identified by a response to one question break out when we group them by responses to another question?

a. Greatest Teaching Contribution

		Physician	Researcher
Ultimate Medical Interest	Total patient management	49%	13%
	Basic mechanisms	51%	87%
		100% (268)	100% (159)

b. Reading Preferences

		Effectiveness	Rationale
Ultimate Medical Interest	Total patient management	68%	30%
	Basic mechanisms	32%	70%
		100% (78)	100% (349)

c. Reading Preferences

		Effectiveness	Rationale
Greatest Teaching Contribution	Physician	85%	64%
	Researcher	15%	36%
		100% (78)	100% (349)

FIGURE 6-2

Bivariate Relationships among Scientific Orientation Items. If several indicators are measures of the same variable, then they should be empirically correlated with one another, as you can observe in this case. Those who choose the scientific orientation on one item are more likely to choose the scientific orientation on other items.

Percent Interested in Basic Mechanisms

		Greatest Teaching Contribution	
		Physician	Researcher
Reading Preferences	Effectiveness of treatments	27% (66)	58% (12)
	Rationale behind treatments	58% (219)	89% (130)

FIGURE 6-3

Trivariate Relationships among Scientific Orientation Items. Indicators of the same variable should be correlated in a multivariate analysis as well as in bivariate analyses. Those who choose the scientific responses on greatest teaching contribution and reading preferences are the most likely to choose the scientific response on the third item.

© Cengage Learning®

Percent Interested in Basic Mechanisms

		Greatest Teaching Contribution	
		Physician	Researcher
Reading Preferences	Effectiveness of treatments	51% (66)	87% (12)
	Rationale behind treatments	51% (219)	87% (130)

FIGURE 6-4

Hypothetical Trivariate Relationship among Scientific Orientation Items. This hypothetical relationship suggests that not all three indicators would contribute effectively to a composite index.

© Cengage Learning®

- Consider the multivariate analysis in figure 6-3.
 - Of a group identified by responses to questions 1 and 2, and sorted into

categories, how do they respond to a third question?

- Multivariate is often handled by calculating crosstabs.

INDEX SCORING

- What is a desirable range of values for the scores?
 - You want as much gradation as possible
 - But also don't want to refine it so much / extend the scale so much there is hardly anyone in each category.
- Should items in the index be given equal weight or different weights?
 - Animal bell equals 1, animal cart equals one, gourd equals one?
 - Or by cash value one the local market to weight?
- Handling missing data.
 - Can exclude the missing cases if there are few.
 - Consider first if there is something systematic about the missing observations that you might lose if you do this.
 - I am missing lots of information about livestock owning households as they move out of the area during the dry season.
 - Is it the case that a missing answer can be interpreted as a 'no' and you can fill in the missing zero?
 - If yes, continue; if no, skip to question 61.
 - Of those answering yes....
 - Is there a pattern in the rest of the data that make it so that all of the people who answered the other

questions the way this person did also answered in a given way on the missing value so I can fill it in?

- Could I fill in the neutral point, sample mean, other approach and not cause more of a problem than I get if I eliminate the case missing the value?
 - Missing GPS data for one village but have them for all the other neighboring villages, take the average for the neighboring ones.
 - Generally, try it a few different ways and see if different approaches matter and make a decision.

Index validation.

- Item analysis.
 - Investigate the extent to which the overall index score is related to the scores on individual items.
 - Line them up to see what percent has what score when cross cut with a particular question.
 - See how the different components contribute to the index.
- External validation
 - Is the rank order of the index related to logically related results for other measures / questions that should get at the underlying concept from a different perspective? Table 6-1.

TABLE 6-1**Validation of Scientific Orientation Index**

	<i>Index of Scientific Orientation</i>			
	<i>Low</i> 0	1	2	<i>High</i> 3
Percent interested in attending scientific lectures at the medical school	34	42	46	65
Percent who say faculty members should have experience as medical researchers	43	60	65	89
Percent who would prefer faculty duties involving research activities only	0	8	32	66
Percent who engaged in research during the preceding academic year	61	76	94	99

Some indices of interest to IR:

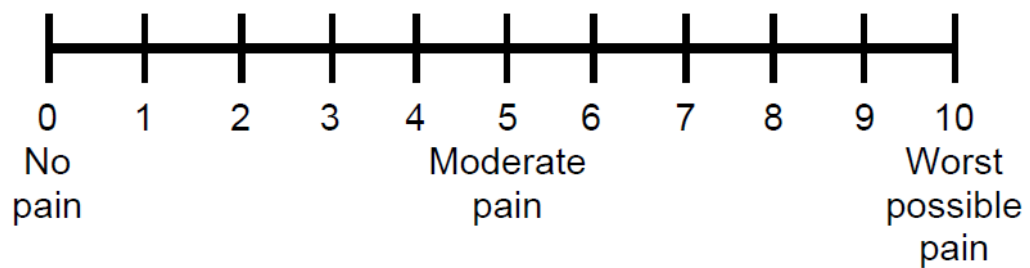
- Polity IV <http://www.systemicpeace.org/polityproject.html>
- TI-CPI <http://www.transparency.org/research/cpi/>
- Bertelsmann Transformation Index <https://www.bti-project.org/en/home/>
- Doing Business Index <http://www.doingbusiness.org/rankings>
- HDI <http://hdr.undp.org/en/content/human-development-index-hdi>
- PISA <http://www.oecd.org/pisa/pisaproducts/>
- Fragile States Index <https://fragilestatesindex.org/>
- Worldwide Governance Indicators: <http://info.worldbank.org/governance/wgi/Home/Reports>
- MCC <https://www.mcc.gov/who-we-fund/indicators>
 - MCC scorecards <https://www.mcc.gov/who-we-fund/scorecard/fy-2020/AF>
- Center for Global Development Commitment to Development Index. <https://www.cgdev.org/commitment-development-index-2018>
- Foreign Policy rankings. <http://foreignpolicy.com/2015/02/03/top-twenty-five-schools-international-relations/>

SCALES

Wong-Baker FACES Pain Rating Scale



0–10 Numeric Pain Rating Scale



Reprinted from Pain: Clinical Manual, McCaffery M, et al, P. 16, Copyright 1999, with permission from Elsevier.

- Tap into the intensity structures among the indicators.

- Bogardus social distance scale. Measurement technique to determine the willingness of people to participate in social relations of differing degrees of closeness with other kinds of people.
 1. Are you willing to allow refugee children into the United States?
 2. Are you willing to allow refugee children into New York State?
 3. Are you willing to allow refugee children into Central New York?
 4. Are you willing to allow refugee children into your neighborhood?
 5. Are you willing to allow refugee children into your spare bedroom?
- Range from 'easy items' to 'hard items'.
- Example of reverse social distance scale.
- For a recent immigrant, ask the following set of questions. Consider typical Caucasian Americans you have met since your arrival. Do you think they:
 5. Would be welcoming if you become a citizen?
 4. Would be welcoming if you move into their neighborhood?
 3. Would be welcoming with you living next door?
 2. Would be welcoming with you becoming a close friend?
 1. Would be welcoming of a member of your family becoming kin by marriage?

- Knowing how many are agreed to tells you which are agreed to. If only one 'would be welcoming' is present, it should be in response to 5. If you tell me there are two 'would be welcoming' it should be to 4 and 5.

Thurstone scales. A panel of judges evaluates a set of indicators that relate to the concept under investigation.

- Each judge gives a numeric score to evaluate how strong they think the indicator is as a measure of the concept in question.
 - First, filter by ones where there is not consensus (maybe look at the variance in the scores given by the judges and toss those that are higher than a threshold you select).
 - Then from those for which there is general agreement, pick one or more from each number of the scale; mean 1, mean 2, mean 3,....
 - Then you could succinctly sum up the degree of the concept by the strongest item scored; this person agreed up to the indicator ranked 6 so I can assume they also agreed to 1-5.

Likert scaling.

- [Puppy – monkey – baby](#) ads should be banned from TV.
 - Strongly agree (2)
 - Agree (1)
 - Neutral (0)
 - Disagree (-1)
 - Strongly disagree (-2)

Semantic Differential.

- Pick words at either ends of the extreme of a way of describing something. Array them on a scale from one end to the other.
 - Sort of like Likert, but instead of just Agree or Disagree, now different dimensions one could use to evaluate.
 - How do you evaluate The Weeknd's performance at the Super Bowl Halftime show?

	Very Much	Somewhat	Neither	Somewhat	Very Much	
Enjoyable	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Unenjoyable
Simple	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Complex
Discordant	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Harmonic
Traditional	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Modern

FIGURE 6-5

Semantic Differential: Feelings about Musical Selections. The semantic differential asks respondents to describe something or someone in terms of opposing adjectives.

© Cengage Learning®

Guttman scaling.

- Some items under consideration may prove to be more extreme indicators than other items.
 - 2012 General Social survey, support for a woman's right to an abortion:
 - Woman's health is seriously endangered (87%)
 - Pregnant as a result of rape (76%)
 - Woman is not married (41%)

TABLE 6-3
Index and Scale Scores

	Response Pattern	Number of Cases	Index Scores	Scale Scores	Total Scale Errors
Scale types	+++	763	3	3	0
	++-	633	2	2	0
	+--	201	1	1	0
	---	191	0	0	0
Mixed types	-+-	43	1	2	43
	+ - +	7	2	3	7
	--+	4	1	0	4
	-++	4	2	3	4

Total scale errors = 58

$$\begin{aligned}
 \text{Coefficient of reproducibility} &= 1 - \frac{\text{number of errors}}{\text{number of guesses}} \\
 &= 1 - \frac{58}{1,846 \times 3} = 1 - \frac{58}{5,538} \\
 &= 0.9895 = 99\%
 \end{aligned}$$

Note: This table presents one common method for scoring mixed types, but you should be advised that other methods can also be used.

© 2016 Cengage Learning®

- The scale types and the mixed types. Working off the basis of the scale types.
- The mixed types present a challenge to assigning a score based on the answers. A 1 for example is the + - - pattern, and we are assigning the value to a - + - or - - + observation.
- Introduces the idea of a coefficient of reproducibility. If you take the scale scores, how often do you assign the right pattern to the actual responses? How often do the index and scale scores align?
 - Coefficient of reproducibility = $1 - \{\text{number of errors} / \text{total number of responses to questions}\}$.
 - In the case in the book, 1846 people answered the survey questions concerning the index, there were three questions used to construct the index. When we take the index and move to a scale, how often is the assumed pattern of the index right?
 - 58 responses misalign scale and index. $1 - (58/5538) = .9895$.
 - Often use a 90% or 95% threshold on this to accept a scale interpretation. If not passing the threshold, use as an index.
- This is context specific. What works in one setting as a scale might not in another.
 - A set of questionnaire items in and of themselves is not a scale, but the observations generated from

these questionnaire responses might let you use a scale interpretation of the responses.

- The same questions might be acceptable to use the responses as a scale in one setting but not in another.
- Scale is making a higher claim than an index
- Claims there is an ordering of intensity
- Claims that if you tell me the scale number value, I can predict which ones they said yes to.
- How confident are you to go from a 'how many' index claim to a 'how much' scale claim in interpreting your results?
 - How much is a stronger inferential claim.

TYPOLOGIES

Nominal classifications of observations in terms of their attributes in two or more variables.

Weberian 'ideal types'.

“Bread and circuses”

Share of the population in the urban area by type:

Stable Democracies Urban Concentration= 23%	Stable Dictatorships Urban Concentration = 30%
Unstable Democracies Urban Concentration = 35%	Unstable Dictatorships Urban Concentration = 37%

Ades and Glaeser argue that to stay in power, throw money (bread and circuses) at the urban population to keep them from revolt. However, this will draw further population inflows.

Often useful for putting structure on analysis.

Herd size when first interviewed

Cash income when first interviewed.

Above median and below median.

Table 1: Livelihood groups in the study region sample

Herd size	Lower Cash Income	Higher Cash Income
Lower Herd Size	Left Out 29%	Moving From 21%
Higher Herd Size	Staying With 21%	Combining 29%

2: Average herd size and income measures by sub-group

Group	Herd Size TLU (per capita)	Total Income per capita per day	Cash Income as % of Total Income	Total Income variability (cv)
Left out	7.3 (0.9)	\$0.20	29%	1.32
Moving From	7.2 (0.9)	\$0.27	46%	0.90
Staying With	23.7 (3.5)	\$0.34	21%	0.82
Combining	26.0 (4.4)	\$0.46	35%	0.63

Table 3: Average resilience Indices, by sub-group (1 is most resilient, 0 is least resilient)

Group	Income Recover Resilience Index	Income Threshold Resilience Index	Asset Recover Resilience Index	Asset Threshold Resilience Index
Left out	0.827	0.156	0.464	0.035
Moving From	0.621	0.305	0.486	0.004
Staying With	0.608	0.237	0.444	0.302
Combining	0.533	0.531	0.425	0.446

Figure 3: Herd size over time, percentage change on March 2000 herd size.

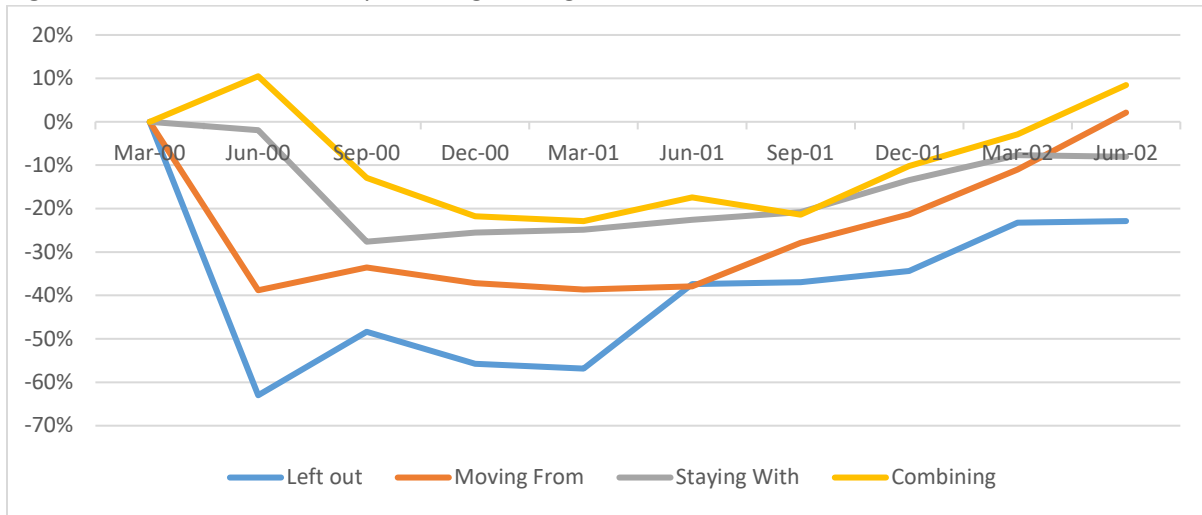


Table 8: Human and Animal Disease Incidence

	Seasons with Human Health Events in Drought	Seasons with Human Health Events in Recovery	Share of all livestock mortality due to disease shock.
1) Left Behind	27.0%	11.0%	31%
2) Combining	38.3%	15.8%	40%
3) Staying With	15.6%	11.5%	21%
4) Combining	23.8%	13.0%	27%
	t12 **, t13 ***, t14 , t23 ***, t24 ***, t34 **,	t12 , t13 , t14 , t23 , t24 , t34 ,	t12 , t13 * , t14 , t23 ***, t24 ** , t34 ,

*** significant difference at 1%, **significant difference at 5%, * significant difference at 10%

Table 10: Price Per TLU in Drought and Recovery.

	Price per TLU in Drought	Price per TLU in Recovery
1) Left Out	\$86.23	\$78.02
2) Moving From	\$94.89	\$125.78
3) Staying With	\$85.64	\$98.20
4) Combining	\$81.19	\$118.33
Significant difference in means by groups, t-statistics	t12 , t13 , t14 , t23 , t24 , t34 ,	t12 ***, t13 ** , t14 ***, t23 * , t24 , t34 * ,

*** significant difference at 1%, **significant difference at 5%, * significant difference at 10%

These are useful for looking across 'types'.

Good as independent variables, framework for analysis.

Hard to use as dependent variables.

Explaining 'why' a household is in each bin.

That may need a different research approach.