Why so Many Items? Complex Constructs and Item Selection

An example with a Subjective Well-Being Survey

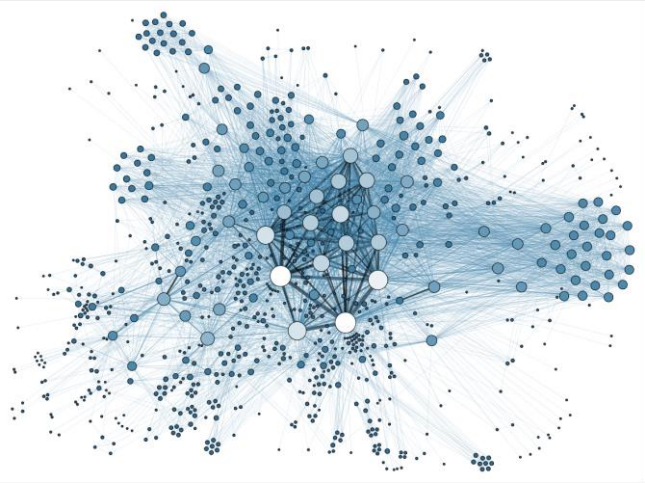
Daniel Katz

w/

Dr. Karen Nylund-Gibson and Dr. Michael Furlong

UC SANTA BARBARA

Complex Constructs to the Fore



- Popular: Measuring “non-academic” constructs in schools via self report
- For instance, CORE districts in California, attempting to measure:
  - Social Emotional Skills
  - Growth Mindset
  - Self-Efficacy
  - School Climate
- All complex: construct-wise and self-report-wise
- How does this complexity enter item selection?

Subjective Well-Being as Archetype

- With these complex constructs: does it make sense to sum survey responses, treat them as continuous?
  - Theory might expect non-ordered solutions (different clusters of students/non-invariance)
- How do theoretical considerations (vs statistical methodological considerations) enter item selection?
- A classic example of a complex construct: Subjective Well-Being (SWB)
  - Typical validation: items selected based on fit to factor model
  - This fit does not indicate adherence to theory

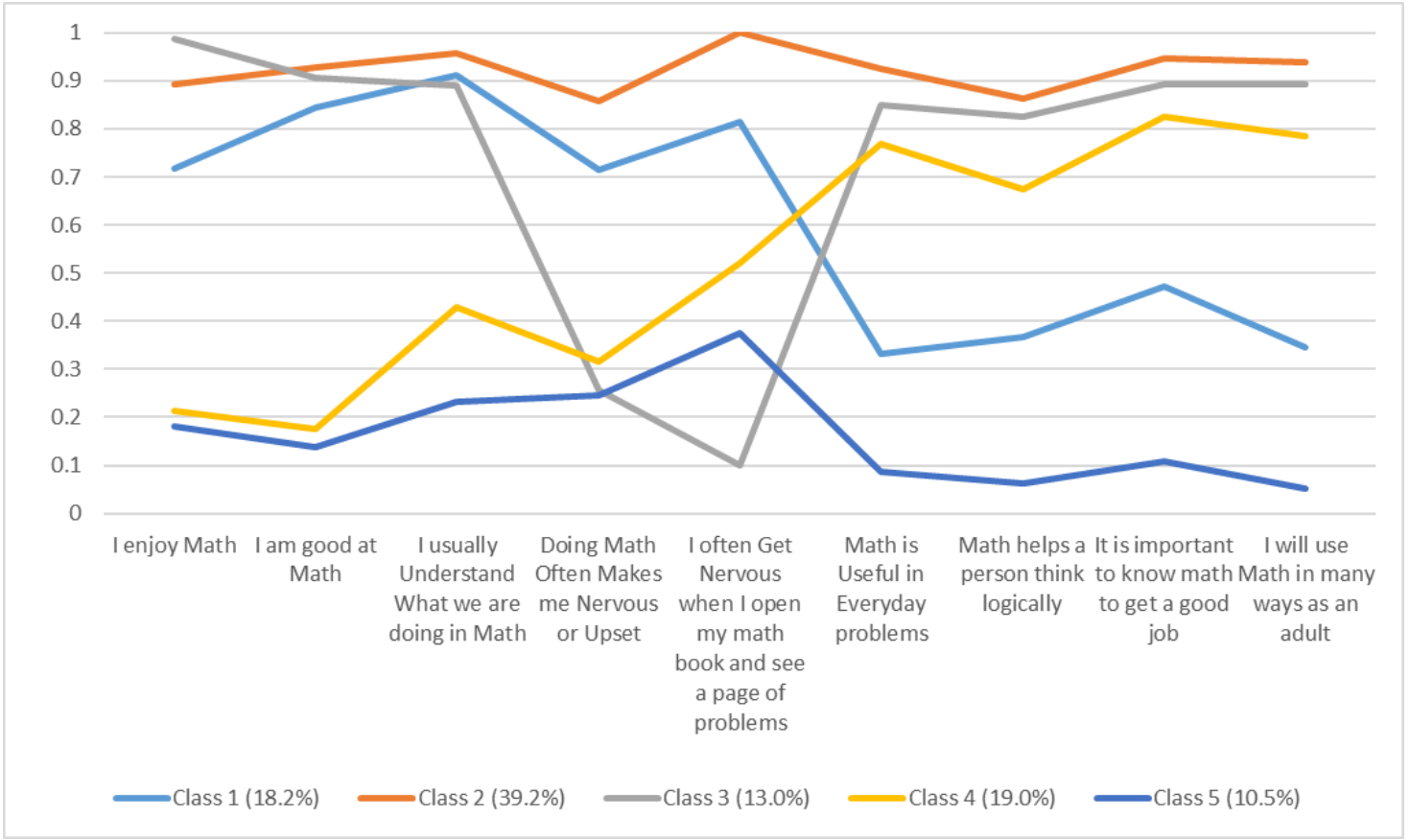
Typical Treatment of SWB Surveys

- Break down SWB into individual parts and sum the item scores
- However, a lot of assumptions about continuousness, equal weighting of items, etc
- In this case, factor analysis would not be adequate for validation
  - Not allowing for this is “theory avoidant” (Alexandrova and Haybron, 2016)
- So what would not be theory avoidant?
  - Allow for models that represent what we think!

One Solution: Mixture Modelling with Distal Outcomes

- Latent Class Analysis (LCA), a form of mixture modelling, allows for disorder
  - Caveat: completely exploratory
- Assumes no ordering, so clusters of students, or classes, can emerge in a non-ordered, heterogeneous way
- If there is disorder: theoretically could see a group of students that have similar overall SWB but very different response profiles across a set of items
- However, many SWB surveys were validated via only-factor analysis

Disorder


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
I enjoy Math
I am good at Math
I usually Understand What we are doing in Math
Doing Math Often Makes me Nervous or Upset
I often Get Nervous when I open my math book and see a page of problems
Math is Useful in Everyday problems
Math helps a person think logically
It is important to know math to get a good job
I will use Math in many ways as an adult
Class 1 (18.2%)
Class 2 (39.2%)
Class 3 (13.0%)
Class 4 (19.0%)
Class 5 (10.5%)

Research Questions:

So, what happens with items used to measure SWB? Does order emerge?

RQ1: Is there a mixture-model solution that shows disorder among classes/students as one may expect?

Example: No set of clusters that indicates classes of students that progress from low to high on each item

RQ2: How do these classes relate to an overall one-item measure of life satisfaction?

Can classes of respondents with different sumscores have equivalent reported overall life satisfaction?

Sample 1

- *N* = 1,908 high school students
- 48.1% Male, 50.5% Female, 28 total marked as missing or

<u>Grade Breakdown:</u>

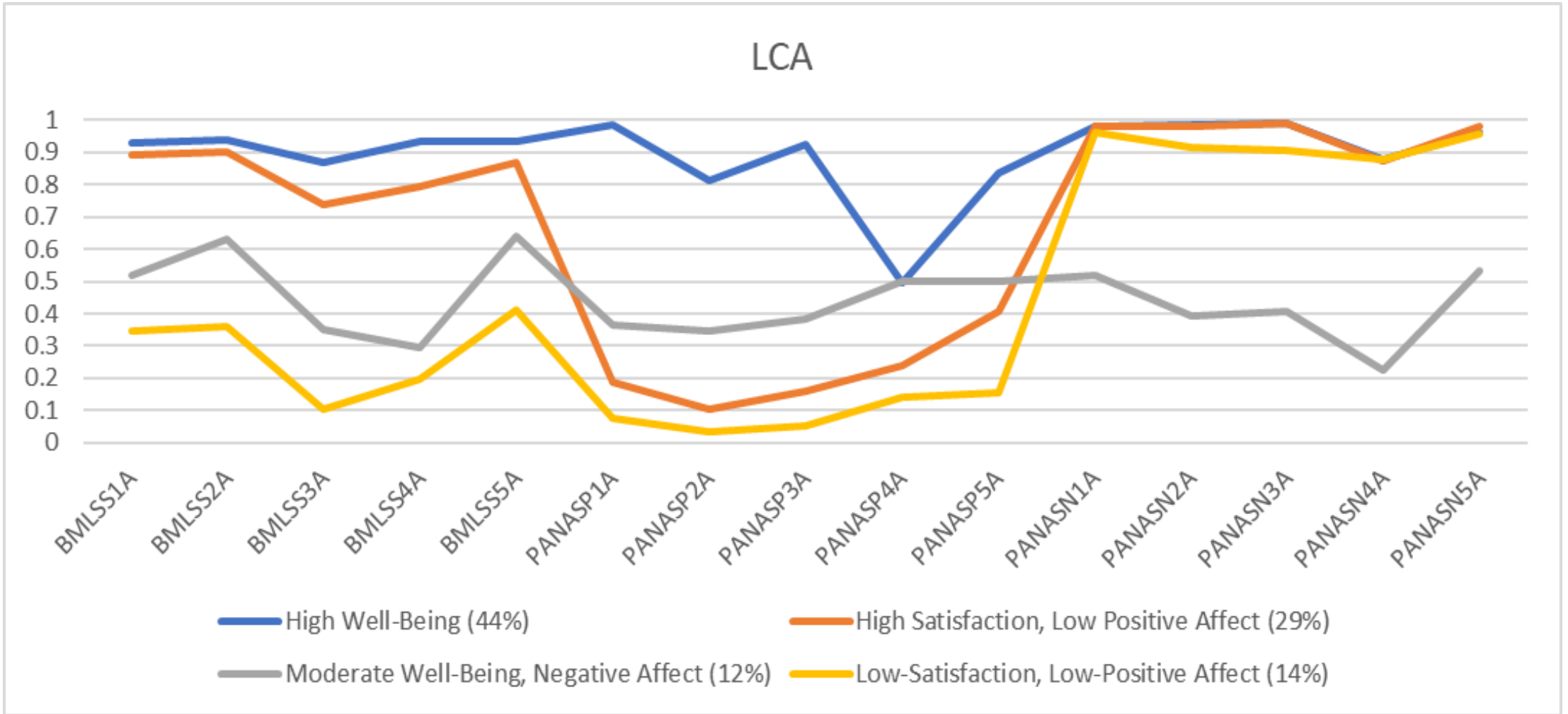
- 9th Grade: 26.4%
- 10th Grade: 26.0%
- 11th Grade: 25.0%;
- 12th Grade: 22.6%
- 48.6% Latino/a or Hispanic; 38.2% White; 7.4% “two or more groups”; 3.1% Asian; 1.3% Black or African American.

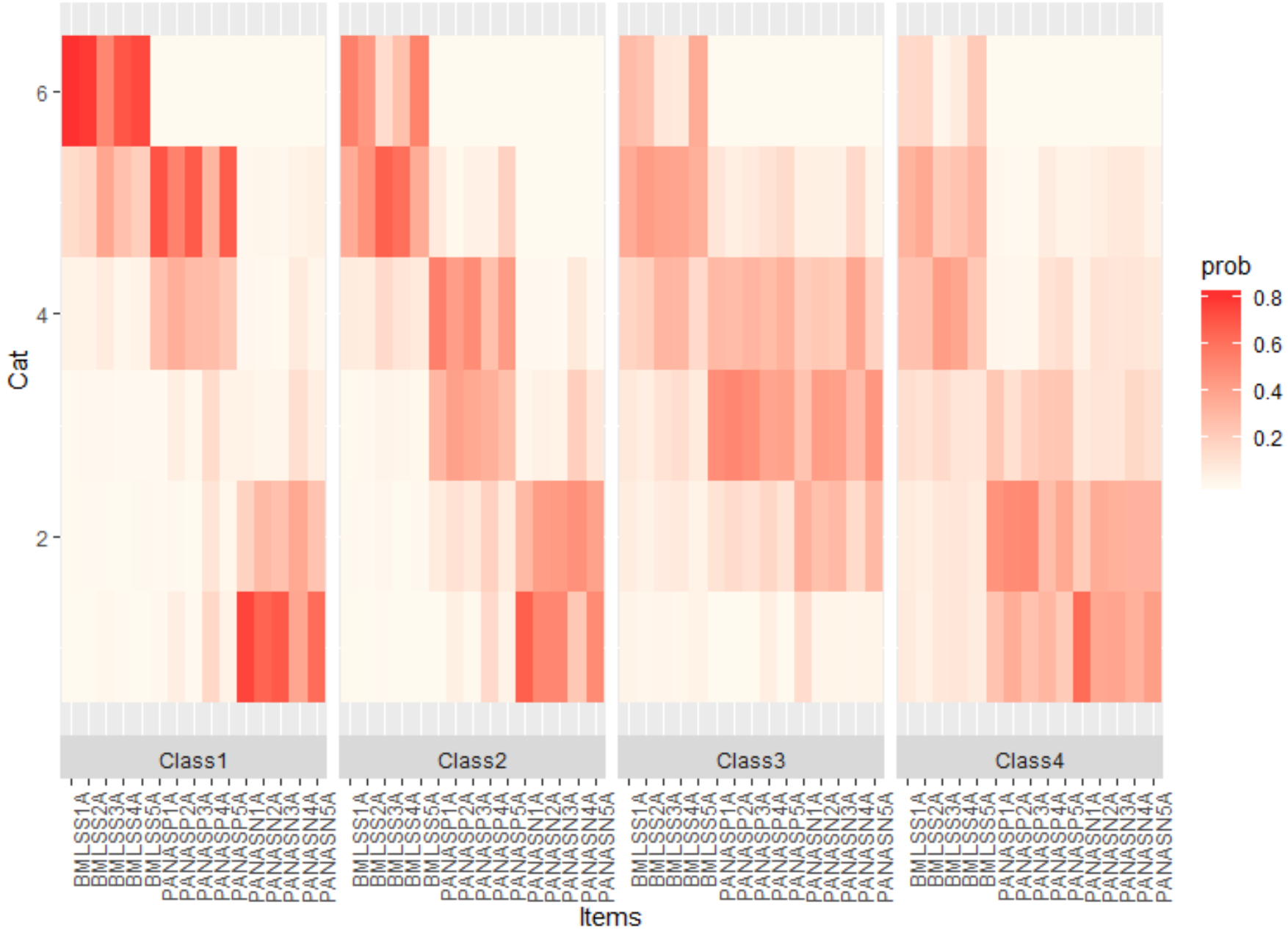
Surveys/Measures

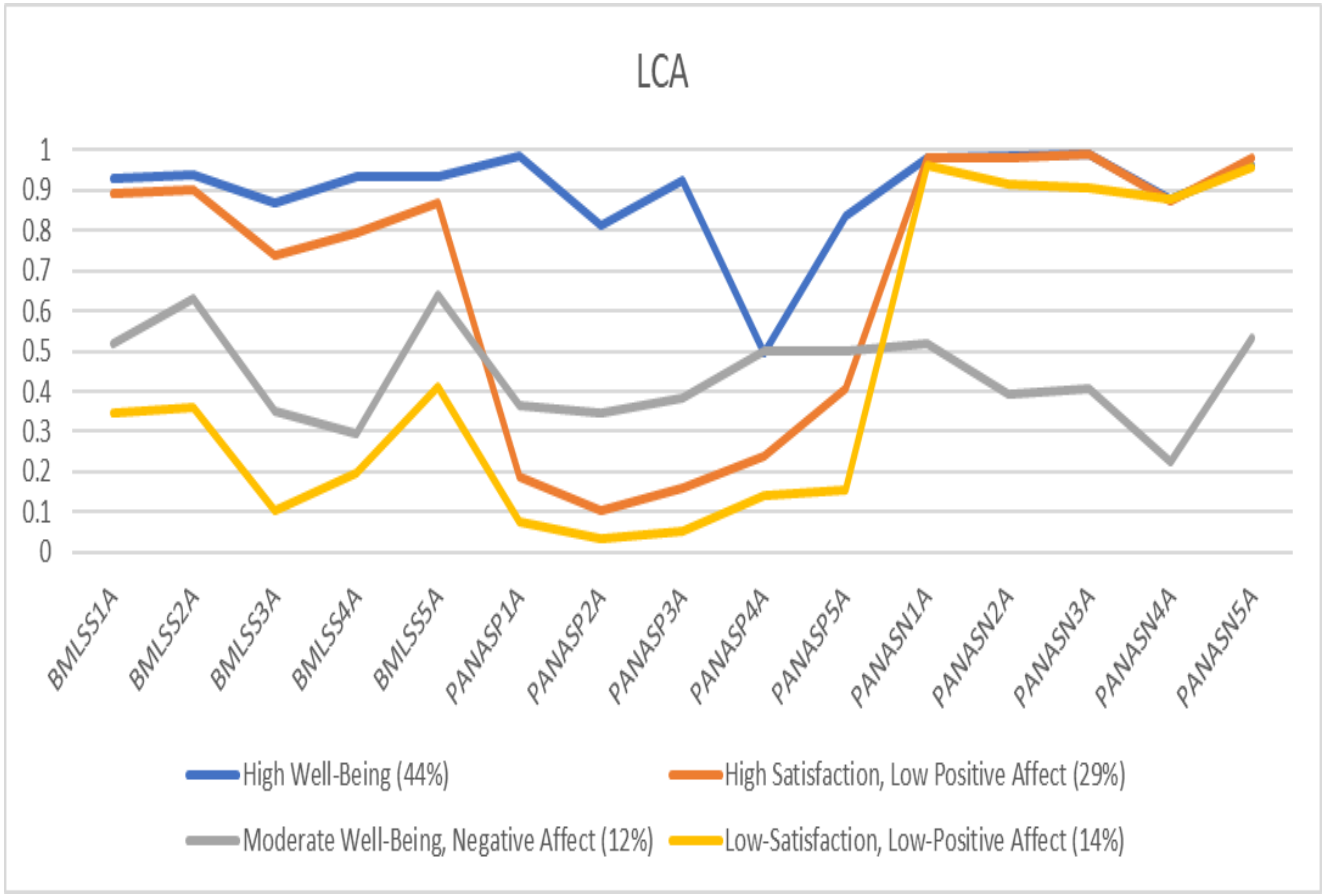
- Used some traditionally combined measures:
  - Measure 1: 5-ItemBrief Multidimensional Measure of Satisfaction with Life Scale (BMLSS)
  - Measure 2: Positive and Negative Affect Schedule (PANASP and PANASN)
  - Both often taken as a sum: this is problematic if the construct of interest is not thought to be perfectly continuous
  - Often combined based on Diener's dual-factor mental health model (see: Suldo, S. M., Thalji-Raitano, A., Kiefer, S. M., & Ferron, J. M., 2016).
  - For presentation, items were dichotomized, but ordinal-lca provided the same solution

Note: Matter of Rhetoric

Since item selection was not based on anything but factor analysis, would not expect a non-ordered solution


LCA
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
BMLSS1A
BMLSS2A
BMLSS3A
BMLSS4A
BMLSS5A
PANASP1A
PANASP2A
PANASP3A
PANASP4A
PANASP5A
PANASN1A
PANASN2A
PANASN3A
PANASN4A
PANASN5A
High Well-Being (44%)
High Satisfaction, Low Positive Affect (29%)
Moderate Well-Being, Negative Affect (12%)
Low-Satisfaction, Low-Positive Affect (14%)


Cat
6
4
2
Class1
Class2
Class3
Class4
BMLSS1A
BMLSS2A
BMLSS3A
BMLSS4A
BMLSS5A
PANASP1A
PANASP2A
PANASP3A
PANASP4A
PANASP5A
PANASN1A
PANASN2A
PANASN3A
PANASN4A
PANASN5A
Items
prob
0.8
0.6
0.4
0.2



One Item Score for Each Class: Average

High Well-Being Class: 90.4

Next Highest Class: 81.1

Moderate Well-Being Class: 58. 9

Low Well-Being: 55.9

Overall: 79.9

Interpretations

- Going from a lower class to higher class was associated with increasing reported life-satisfaction: If you want to use one item, it looks like you can.
- Typical statistical validation methods aren't always helping surveys and measurement.
  - Think about how many ways a person might think about their well-being
  - The desire to have general questions is problematic
- Leads to questions about why use multiple items in surveys and ways in which we select items (redundancy!)
- Subjectivity in item appraisals in these sorts of surveys is complex
  - Between person measurement even possible?
- What are we measuring? Item appraisals, affect, semantics?
- Brings about the challenges of between person measurement

Thank you

Please get in touch: I’m continuing to think about implications and limitations, and would love to hear other thoughts.

dkatz@ucsb.edu

- Alexandrova, A., & Haybron, D. M. (2016). *Is Construct Validation Valid?* Retrieved from http://www.journals.uchicago.edu/t-and-c

- Suldo, S. M., Thalji-Raitano, A., Kiefer, S. M., & Ferron, J. M. (2016). Conceptualizing high school students' mental health through a dual factor model. *School Psychology Review,45*(4), 434–457. https://doi.org/10.17105/spr45-4.434-457.CrossRefGoogle Scholar

Measure1: 5-ItemBrief Multidimensional Measure of Satisfaction with Life Scale (BMLSS)

- 5 items each beginning with the stem: I would describe my satisfaction with ___
  - Family
  - Friendship
  - School Experience
  - Myself
  - Where I live

- 6 point Likert
  - 1—Very Dissatisfied
  - 2 -- Somewhat Dissatisfied
  - 3- A little Dissatisfied
  - 4– A little Satisfied
  - 5 – Somewhat Satisfied
  - 6 – Very Satisfied

Measure 2: Positive and Negative Affect Schedule (PANASP and PANASN)-

10 items (it actually comprises 15 items)

Thinking about yourself and how you normally feel, how much do you generally feel:

- Joyful
- Scared
- Upset
- Delighted
- Sad
- Cheerful
- Nervous
- Gloomy
- Alert
- Determined

5-Category Likert Scale:

1 = Not At All
2 = A Little
3 = Moderately
4 = Quite a bit
5 = Extremely

LPA (default MPLUS-Only: Diagonal, class invariant—Variances the same, no covariance)

Number of classes	Log likelihood	BIC	ABIC	p-value of BLRT	p-value of LMRT	Entropy	BF	cmP
1	-42975.12	8617.728	86081.418	_	_	_	#NUM!	1
2	-40529.76	81406.80	81260.659	<.0001	0.0283	0.872	0	0
3	-39149.451	78766.98	78570.003	<.0001	<.0001	0.889	3E-284	0
4	-38436.154	77461.18	77213.371	<.0001	<.0001	**0.907**	**0.00000**	0
5	-38048.001	76284.67	76507.028	<.0001	**0.0002**	0.865	1E+28	0
6	**-37791.597**	**76413.65**	**76064.181**	<.0001	0.175	0.867	0	0