PSY792F SEM

Week 8 – Latent Profile Analysis

Mark A. Prince, PhD, MS

Mixture Modeling

- A general class of statistical models focused on finding subtypes of related responses from multivariate data
  - Some examples
    - Latent Class Analysis – all binary indicators
    - Latent Profile Analysis – all continuous or a mix of binary and continuous predictors
    - Latent Transition Analysis – longitudinal extension of LCA
    - Growth Mixture Models – longitudinal extension of LPA

Some points to remember…

- Identifying latent profiles is not grouping people into latent groups (that is called cluster analysis)
- Mixture models define a finite number of profiles and every participant has a probabilistic membership of being in every profile
- It is important to take the probabilistic class membership into consideration when examining predictors of class membership or examining variables that class membership predicts
  - So, **do not** assign people to classes and then run an ANOVA!!!
- Be careful with your language
  - It is ok to use the terms class and profile interchangeably, but specificity is preferred
  - Try not to say things like
    - "those in the abstaining group had better coping" (even though I'm guilty of this)
  - Its better to say things like
    - "the abstaining profile was associated with higher levels of coping"

General Description

- LCA (or LPA – by changing "class" to "profile")
- Used to identify classes of related responses from multivariate categorical data
- Concerned with the structure of cases (i.e., the latent taxonomic structure)
- Person-centered
  - Differs from factor analysis which is concerned with the structure of variables
- Similar to cluster analysis, but based on probabilistic theory

Goals

- Responses from a heterogeneous population assumed to belong to a limited number of homogeneous groups referred to as ”latent classes”
- Latent classes are categories of a latent variable, each one of which contains individuals’ responses who are similar to others’ responses within a class and different from individuals’ responses in other classes
- In other words, mixture models try to maximize within group homogeneity and between group heterogeneity.
- Ultimately, the response patterns to observed categorical (LCA) or continuous (LPA) variables are used to:
    - Discern the number of underlying classes
    - Assign probabilities to each individual for each latent class
    - Determine probabilistic class size

III. *Contributions to the Mathematical Theory of Evolution.*

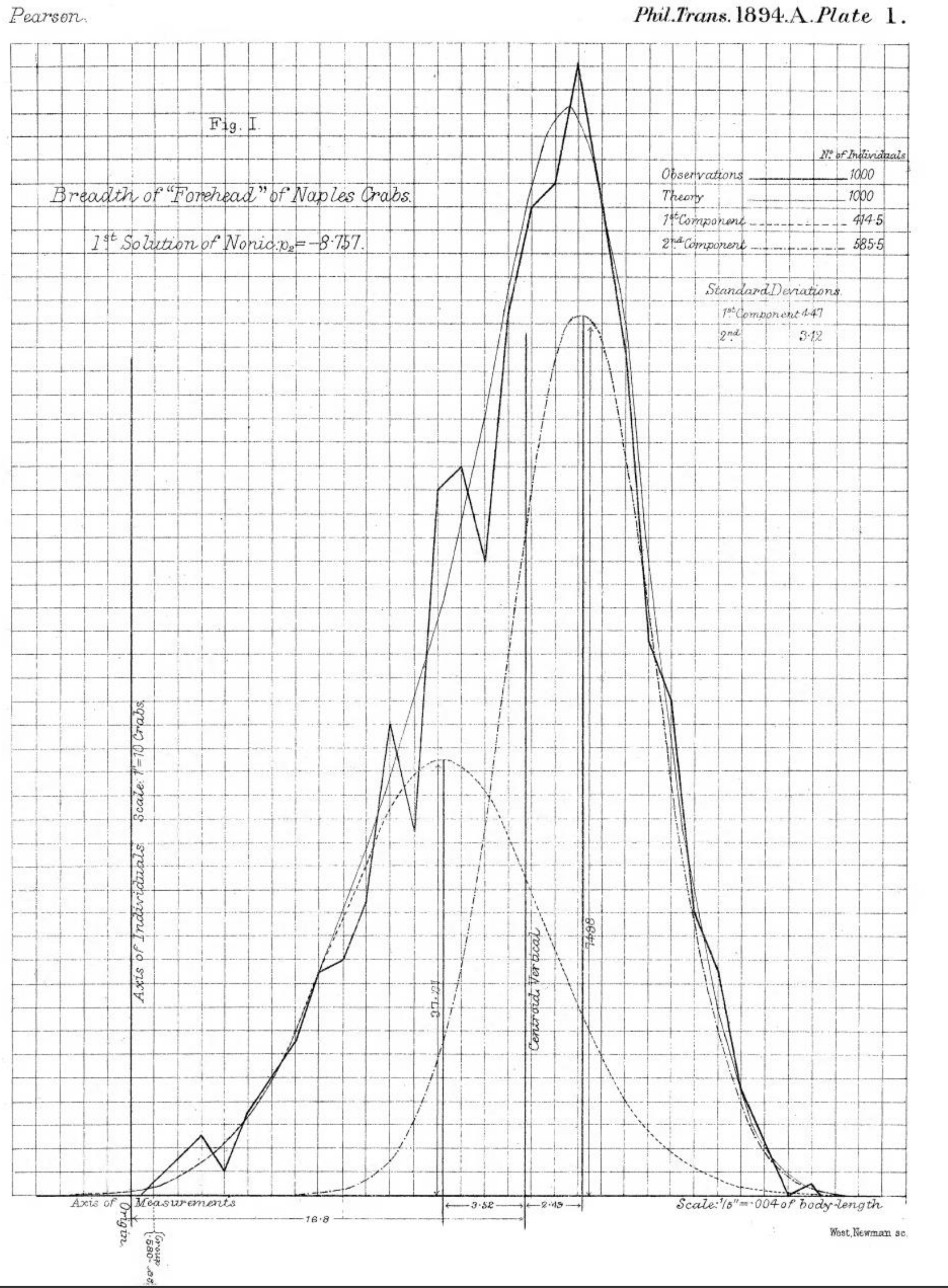
By KARL PEARSON, *University College, London.*

Communicated by Professor HENRICI, *F.R.S.*

Received October 18,—Read November 16, 1893.

1893

Just when you thought,
"This is really cutting-edge methodology"

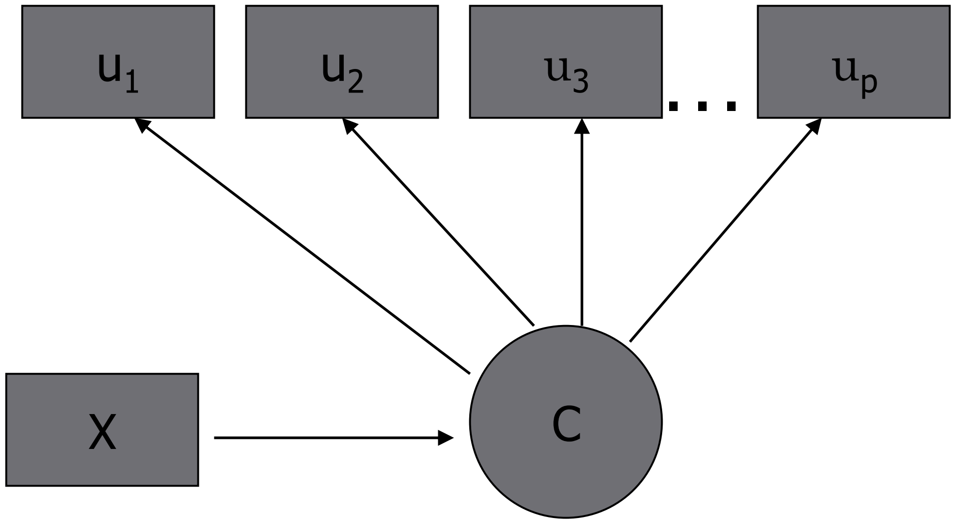


The mechanics of LCA/LPA

- There are 2 primary model parameters
  - The prevalence of each case in a class – latent class probabilities (LCP)
  - The conditional response probability (CRP)
    - The probability that an individual with high probability of being in a particular class will respond ”present” or “yes” to a target item (in LCA) or the typical value for a continuous variable associated with a given profile (in LPA)

LCA Model

- Observed Categorical Items (u’s)
  - Could swap these with continuous variables for LPA
- Categorical Latent Class Variable (c)
- Continuous or Categorical Covariates (x)



Steps for LCA/LPA

- Model comparison and selection
  - How many classes/profiles?
- Examination of class/profile structure
  - Meaningful? Theoretically consistent?
- Interpretation of latent classes/profiles
  - Once you have the best fitting model, what does it tell you??

Model Comparison and Selection

- Find the most parsimonious model
  - Model with the fewest parameters that maximizes the association among the observed variables
- Run a number of models by imposing a different number of classes/profiles on your data
  - Compare comparative fit indices (these are mostly new for us)
    - BIC
      - Closer to 0 is better
    - Entropy (classification quality)
      - Closer to 1 is better (should be greater than .80)
    - Average Latent Class Probabilities (e.g., how distinct are classes/profiles)
      - Closer to 1 is better (should be greater than .90)
    - Probabilistic class sizes
      - Smallest class should represented at least 5% of cases
    - Lo-Mendell-Rubin Test (tech11)
      - Should be significant for the final model and non-significant for a model with k+1 classes
    - Estimated latent class sample statistics should be meaningful (tech7)
      - Substantive interpretation

Interpretation of latent classes/profiles

- Based on conditional response probabilities (CRP)
  - Similar to factor loadings
- LCA
  - CRPs tell us the likelihood of endorsing a item associated with that profile.
- LPA
  - Conditional response means (CRM) define the typical response for each item associated with that item.
- We want discrimination
  - Equal CRPs across classes are not helpful
  - Similar typical values are not helpful

LCA example

- Questions, answer with yes or no
  - Are you happy?
  - Are you sad?
  - Are you angry?

Class	Happy	Sad	Angry
1	.90	.10	.02
2	.05	.80	.70

- Class 1 is highly likely to endorse being happy and highly unlikely to endorse being sad or angry
- Class 2 is highly likely to endorse being sad and angry and highly unlikely to endorse being happy.

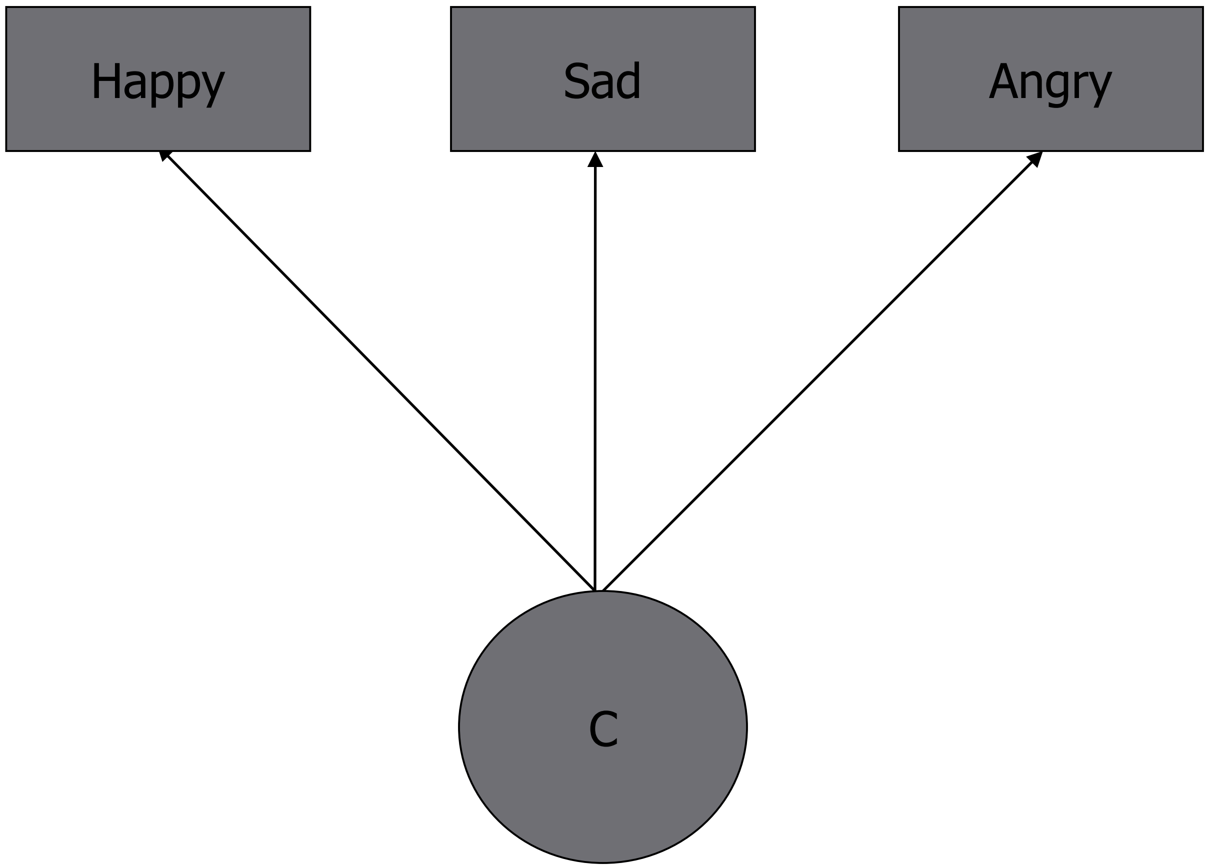
LPA example

- Questions, answer on a scale from 1-10.
  - Are you happy?
  - Are you sad?
  - Are you angry?

Class	Happy	Sad	Angry
1	8.5	2.3	3.1
2	0.5	8.3	7.7

- Class 1 is associated with high levels of happiness and low levels of sadness and anger
- Class 2 is associated with low levels of happiness and high levels of sadness and anger

In either case the model being tested is…

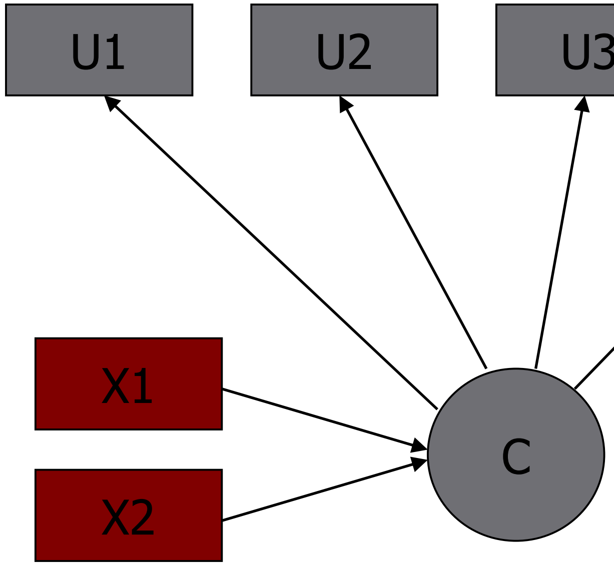


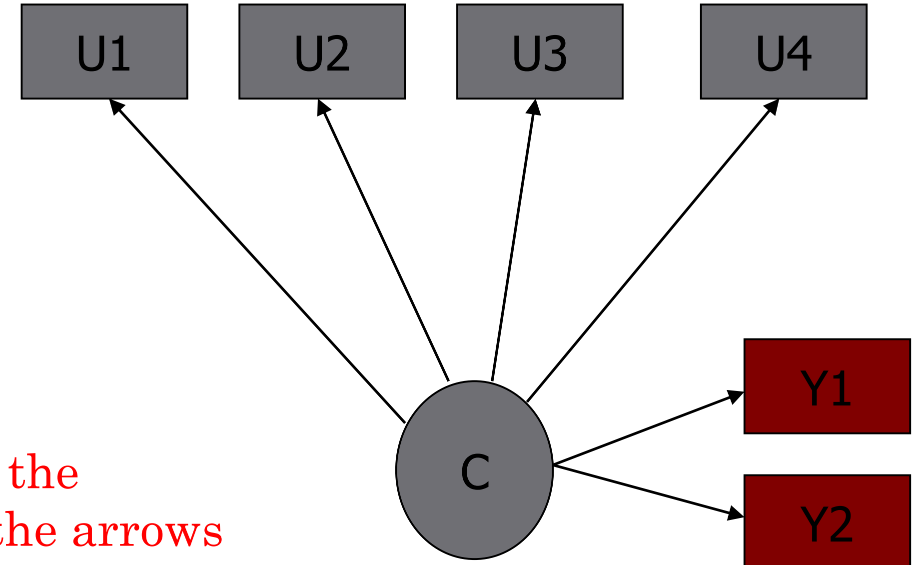
Comparing Model Fit Example

- Real data
  - Project Drinking
    - Protective Behavioral Strategy Use – use of strategies to reduce alcohol use or reduce the risk of alcohol related consequences
      - Three subscales: stopping/limiting drinking, manner of drinking, serious harm reduction
    - Perceived Effectiveness – the perceived effectiveness of PBS used
    - Alcohol use – typical quantity = typical number of drinks in a drinking day
    - Alcohol related consequences – number of alcohol related consequences in the past month
- See excel sheet

Some extensions of LCA/LPA

- Incorporate antecedent variables to predict class membership
- Incorporate outcome variables predicted by the categorical class variable
  - These two extensions will change the class structure

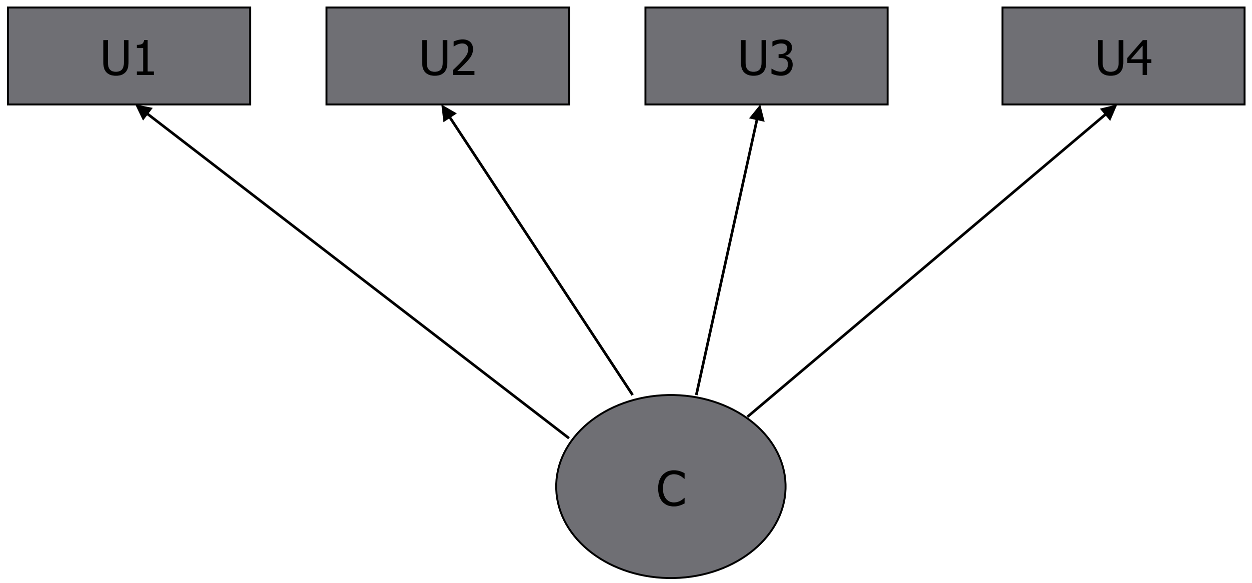




Note the direction of the arrows. On the left, the arrows point to C so you need an on statement.
On the right, the arrows point to the outcomes so they become indicators

Some more extensions of LCA/LPA

- Test for class differences using auxiliary statement
  - Lots of options (BCH) seems to be the current best
  - These will not change your model
- You use a Wald Test, that you interpret like ANOVA
  - Omnibus test, and pairwise tests (when more than 2 classes)



Effect size to compare means

- Can compare the difference in means across classes using Cramer's V.
- V = SQRT(χ^2/nm)
  - χ^2 is the chi-squared value from the wald test
  - n = sample size
  - m = the smaller of (rows – 1) or (columns – 1)
- Used to test independence – has recommended cutoffs based on degrees of freedom for a chi-squared test
- DF=1 (0.10 = small effect) (0.30 = medium effect) (0.50=large effect)
- DF=2 (0.07 = small effect) (0.21= medium effect) (0.35 = large effect)
- DF=3 (0.06 = small effect) (0.17 = medium effect) (0.29 = large effect)

How to write the code…

- New classes option
- If no predictors then can just have a usevariables list
- If you have predictors need a model statement
- Can have different predictors within each class/profile – need to use a class specific model section
- If you have specification problems you can fix the variance to 0 for your indicators within each class

A note on starting values

- Default in Mplus is 25 5
  - This means take 25 starting values and see if at least 5 converge
- What is a starting value?
  - Imagine a multivariate mountain range
  - You can't see in multidimensional space, and your goal is to find the highest peak in the range
  - You have 25 mountain climbers and you drop them off somewhere and say "climb up, and tell me when you get to the top"
  - Your rule is that at least 5 of them reach the same "highest peak" you will treat it as the highest peak in the range

How to use starting values when a model doesn't converge

- When a model won't converge you should dramatically increase your starting values
  - 100 10
  - 500 25
  - *keeping in mind that its easier to get 500 5 to converge than 500 25 – but you want to feel good about your solution
- Once you get starting values, you can save them for use in subsequent models to be able to replicate your solution

How to write up the results

- Describe variables (continuous or categorical)
- Describe model building and selection process including fit indices and criteria you used
  - Make a table of comparative fit indices
- Only describe the final model in your results section
- Name the classes/profiles in your final model
- In discussion link to theory or clinical utility

Example Tables *(from Prince et al 2016, not linked to write-up)*

Table 2

Comparative Latent Profile Analysis Overall Model Fit Statistics for Client Reports on the Therapeutic Alliance

	Overall model fit			
	1-class	2-class	3-class	4-class
saBIC	1048.07	795.55	404.27	665.85
Entropy	1.00	.936	.971	.967
ALC-Prob.	—	.958–.987	.98–1.0	.92–1.0
% sample/class				
1	65 (100%)	22 (35%)	21 (34%)	9 (15%)
2		40 (65%)	27 (44%)	13 (21%)
3			14 (23%)	26 (42%)
4				14 (23%)
LMR (*P* value)	—	196.29 (.00)	102.12 (.00)	50.23 (.54)

Note. saBIC = sample size adjusted Bayesian information criterion; ALC-Prob. = average latent class probability for most likely latent class membership; LRM = Lo-Mendell-Rubin adjusted likelihood ratio test. Wald tests of equality were used to compare profiles in models with more than one class on follow-up percent days abstinent. Thus, in the 1-class model the total *N* is 65, as there were no classes to compare; the total *N* is 62 for the two- through four-class models due to the loss of three participants who failed to provide follow-up data.

Table 4
Descriptive Information for 3-Profile Model of Client Ratings of Therapeutic Alliance

	Low-CA, n = 21 (34%) M (SD)	Medium-CA, n = 27 (44%) M (SD)	High-CA, n = 14 (22%) M (SD)
Variables defining latent profiles			
Client alliance mean	4.98 (.65)	6.18 (.24)	6.77 (.09)
Client alliance min	4.05 (.60)	5.63 (.40)	6.13 (.21)
Client alliance max	5.58 (.81)	6.34 (.29)	6.98 (.03)
Client alliance difference (range)	1.53 (1.04)	1.01 (.57)	.85 (.21)
Client session number difference (min-max latency)	3.27 (4.73)	2.90 (5.03)	2.63 (2.98)
Control variables			
PDA Pre-Tx	.30 (.25)	.38 (.27)	.43 (.24)
PDA Tx	.66 (.10)	.80 (.24)	.89 (.16)
Auxiliary outcome variable	M (SE)	M (SE)	M (SE)
PDA 4m FU	.59 (.09)	.77 (.05)	.91 (.05)
Session attendance	8.23 (.87)	9.11 (.70)	11.14 (.56)

Note. low-CA = weak client alliance profile; medium-CA = moderate client alliance profile; high-CA = strong client alliance profile; PDA = percent days abstinent; Pre-Tx = pretreatment; Tx = treatment; 4m FU = 4-month follow-up.