Using Latent Variable Mixture Modeling to Explore Sample Diversity

Dr. Rick Sawatzky, PhD, RN
Nursing Department - Trinity Western University

Dr. Chris G Richardson, PhD
Research Scientist, CHÉOS
Assistant Professor, Dept of Health Care and Epidemiology, UBC
Emerging tobacco dependence

• Tobacco use typically begins in adolescence

• Emerging tobacco dependence is multidimensional

• Diversity influences emergence and maintenance of dependence: Gender
  – Girls experience higher levels of nicotine dependence
  – Girls have a shorter time to withdrawal symptoms
  – Nicotine influences the subjective and reinforcing effects of tobacco use to a lesser degree in woman
  – Subjective and reinforcing effects of non-nicotinic stimuli associated with smoking may be greater in women
Theoretical model for the DTDS

Nicotine
- Strong cravings
- Feel right
- Concentrate

Smoking
- Emotional
  - Stressed
  - Angry
  - Nervous
- Sensory
  - Handling cigarettes
  - Taste
  - Blowing out smoke
- Social
  - Fit in at school
  - Look cool
  - Feel popular
Diversity

• Diversity is often meant to refer to known differences that are relevant to a particular analysis.
  – Demographic differences
  – Other relevant differences

• The assumption is that one knows what the relevant differences are *a priori*.

• We examined the plausibility of sources of diversity that are *not* known *a priori* by evaluating sample heterogeneity with respect to:
  a. The measurement structures of the DTDS subscales
  b. The profiles of the DTDS summary scores
Data and methodology

B.C. Youth Survey on Smoking and Health 2

- Cross-sectional survey conducted in 2004
- 8125 students
- 1,425 smoked at least once in previous month
- 55% female, average age of 16 years
- Average of 4.1 cigs/day

Methods

- Factor Mixture Analysis
- Latent Profile Analysis
Factor mixture analysis

An exploratory technique used to identify subgroups of respondents that differ with respect to the measurement of each of the DTDS dimensions.

Research questions:

1. Are there subgroups of adolescents that respond to the DTDS questions in different ways?

2. Are these subgroups explained by demographic variables and other variables associated with tobacco dependence (e.g., life time # of cigarettes smoked)?
Factor mixture model diagram

Latent classes

Ethnicity
Gender
Grade
Depression
Lifetime smoking

Social

deps6
deps2
deps4
depw3
deps7
deps5

ε ε ε ε ε ε
## FMAs of DTDS subscales

<table>
<thead>
<tr>
<th>Subscale</th>
<th>$K$</th>
<th>BIC</th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicotine</td>
<td>1</td>
<td>13621.91</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>13517.94</td>
<td>0.60</td>
<td>0.40</td>
</tr>
<tr>
<td>Emotion</td>
<td>1</td>
<td>43585.25</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>43162.00</td>
<td>0.60</td>
<td>0.40</td>
</tr>
<tr>
<td>Social</td>
<td>1</td>
<td>13629.07</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>13332.22</td>
<td>0.64</td>
<td>0.36</td>
</tr>
<tr>
<td>Sensory</td>
<td>1</td>
<td>15880.41</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>15559.85</td>
<td>0.65</td>
<td>0.35</td>
</tr>
</tbody>
</table>

*K* = number of latent classes. *P* = number of model parameters. BIC = Bayesian Information Criterion.
Social dimension of the DTDS

Smoking helps me feel good at school

Class 1 (n = 909)
Class 2 (n = 519)
Social dimension of the DTDS

This item does not provide much information about adolescents in class 1.
Explaining latent classes

Logistic regression analysis results

- Assumption: the latent classes are discrete.
- The following variables did explain latent class membership ($p < .05$) in some models but only to a small extent ($R^2 < 3\%$):
  - Gender, grade, ethnicity
  - Lifetime # of cigarettes smoked

Inclusion of explanatory variables in the FMA

- Assumption: all relevant explanatory variables are included.
- Grade and lifetime # of cigarettes smoked significantly explained latent class membership for social and sensation subscales.
- Grade, lifetime # of cigarettes smoked, depression and ethnicity explained latent class membership for the emotion subscale.
- Nicotine dependence subscale did not converge.
Explaining latent classes

Inclusion of explanatory variables in the FMA

- Assumption: the latent classes are discrete.

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Emotion (R² = .01)</th>
<th>Social (R² = .01)</th>
<th>Sensation (R² = .03)</th>
<th>Nicotine (R² = .03)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Green</td>
<td>Red</td>
<td>Green</td>
<td>Red</td>
</tr>
<tr>
<td>Grade</td>
<td>Red</td>
<td>Green</td>
<td>Red</td>
<td>Green</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Red</td>
<td>Green</td>
<td>Red</td>
<td>Green</td>
</tr>
<tr>
<td>Depression</td>
<td>Red</td>
<td>Red</td>
<td>Red</td>
<td>Green</td>
</tr>
<tr>
<td># cigarettes smoked</td>
<td>Red</td>
<td>Red</td>
<td>Red</td>
<td>Green</td>
</tr>
</tbody>
</table>

- Green = Not significant
- Red = Significant (p < .05)
Explaining latent classes

### Inclusion of explanatory variables in the FMA

- Assumption: all relevant explanatory variables are included.

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Emotion</th>
<th>Social</th>
<th>Sensation</th>
<th>Nicotine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td>nc</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
<td>nc</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td>nc</td>
</tr>
<tr>
<td>Depression</td>
<td></td>
<td></td>
<td></td>
<td>nc</td>
</tr>
<tr>
<td># cigarettes smoked</td>
<td></td>
<td></td>
<td></td>
<td>nc</td>
</tr>
</tbody>
</table>

- = Not significant
- = Significant ($p < .01$)
Factor mixture analysis results

Implications

• Adolescents may not interpret and respond to some DTDS items in a consistent manner.

• Some inconsistencies are explained by differences in gender, grade, ethnicity, depression and # of cigarettes smoked in a lifetime.

• Most of the variance in latent class membership (sample heterogeneity) remains unexplained.
Traditionally, look at differences on total scores each dimension of the DTDS

But “diversity” leads to multiple pathways to addiction

LPA identifies subgroups of respondents with similar DTDS profiles
DTDS profiles for the 3 classes
Correlations for the 3 classes

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Class 1 (n=255)</th>
<th>Class 2 (n=391)</th>
<th>Class 3 (n=476)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social with</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicotine</td>
<td>.22</td>
<td>.17</td>
<td>.46</td>
</tr>
<tr>
<td>Emotion</td>
<td>.03</td>
<td>.15 (NS)</td>
<td>.37</td>
</tr>
<tr>
<td>Sensation</td>
<td>.30</td>
<td>.55</td>
<td>.61</td>
</tr>
<tr>
<td>Nicotine with</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotion</td>
<td>.46</td>
<td>.07 (NS)</td>
<td>.72</td>
</tr>
<tr>
<td>Sensation</td>
<td>.17</td>
<td>.17</td>
<td>.48</td>
</tr>
<tr>
<td>Emotion with</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensation</td>
<td>.16</td>
<td>-.01 (NS)</td>
<td>.40</td>
</tr>
</tbody>
</table>
Explaining latent profiles

Logistic regression analysis results

- The following variables did significantly explain latent class membership:
  - Lifetime # of cigarettes smoked, depression & grade

Inclusion of explanatory variables in the LPA

- Lifetime # of cigarettes smoked, depression & grade significantly explained latent class membership in all models.
Explaining latent classes

Can indentify latent class membership BUT...

• Class structure should be informed by covariates
• ALL relevant covariates should be included
• Examine stability and fit of models as covariates included
• More research needed on this methodology