CASoS ENGINEERING

http://www.sandia.gov/CasosEngineering

Opinion-Driven Behavioral Dynamics Model of Tobacco Product Use: Results and Validation

Patrick D. Finley, Thomas W. Moore,
Nancy S. Brodsky, Gregory J. Lambert

Sandia National Laboratories

2014 Joint Statistical Meetings
Boston, MA

August 4, 2014

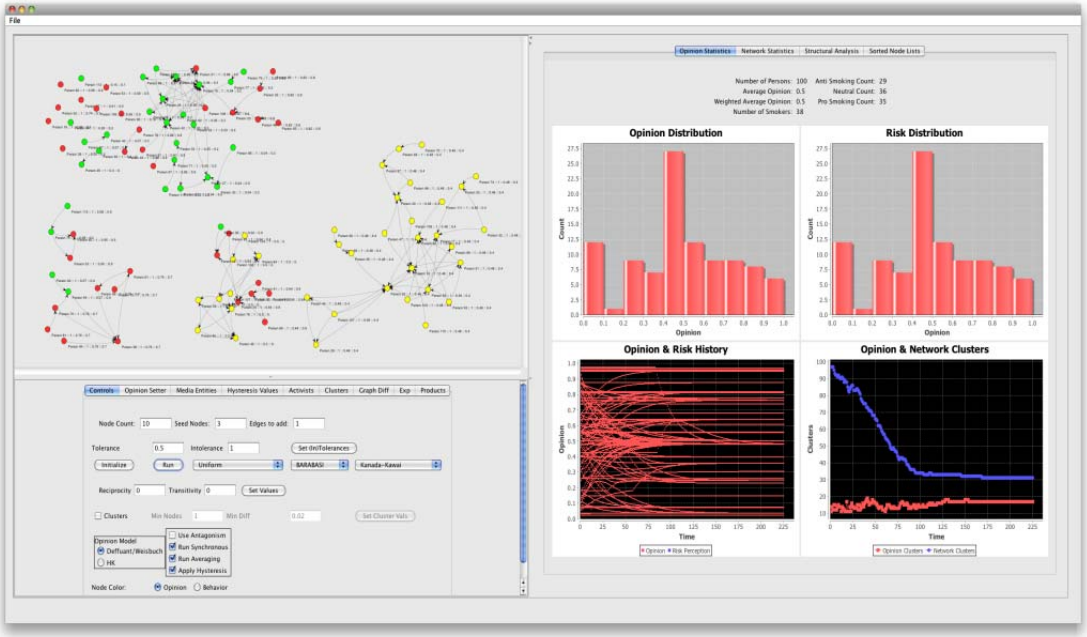
This work was funded by the U.S. Food and Drug Administration (FDA) through a contract with the U.S. Department of Energy/Sandia National Laboratories. The information included in this presentation reflects the views of the authors and should not be construed to represent FDA's positions or policies.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND2014-16379 PE

Sandia National Laboratories

SnapDragon Models Tobacco Use

- Agent-based model of tobacco use within a social network
  - Information about tobacco product use flows among individuals in network
  - This information affects individuals' decisions to start or stop using tobacco
- Media influences
  - Public Health messages
  - Industry advertisements
- Incorporates Risk
  - Product Risk Perception
  - Individual Risk Affinity



- Model used to:
  Explore effects of market and regulatory changes on tobacco use

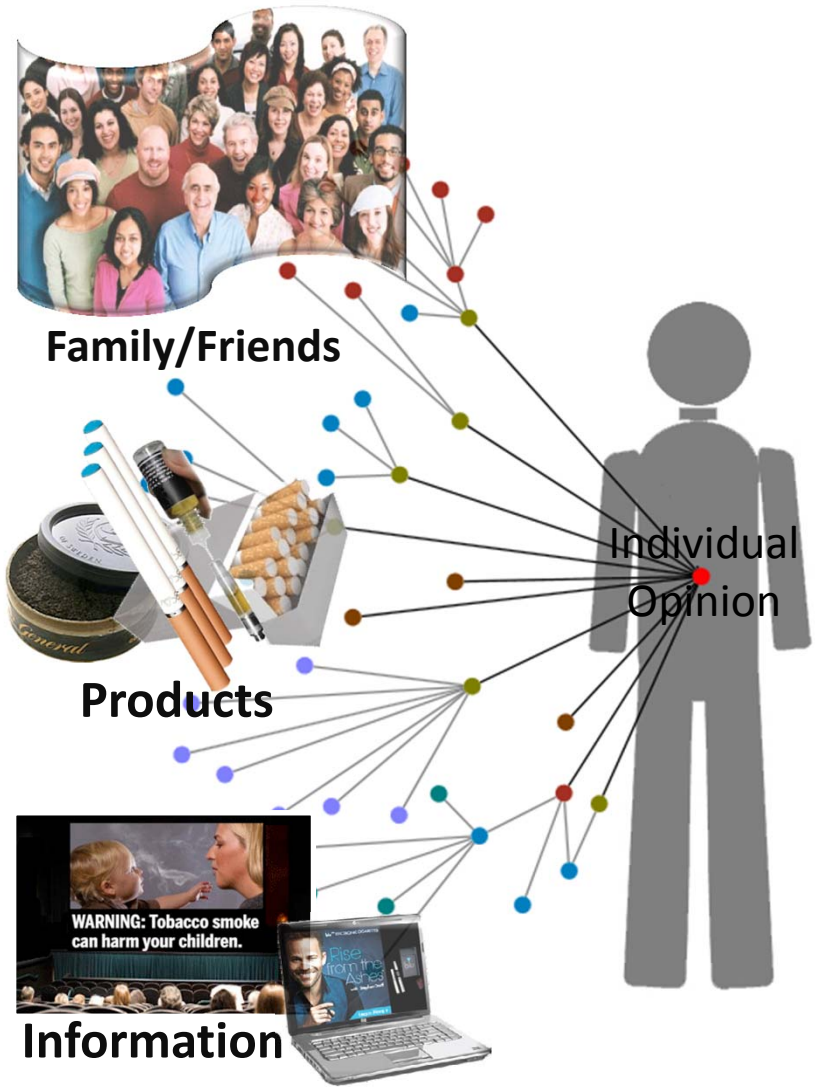
Sandia National Laboratories

Social networks: one way to understand tobacco use

"The evidence is sufficient to conclude that there is a causal relationship between peer group social influences and the initiation and maintenance of smoking behaviors during adolescence."

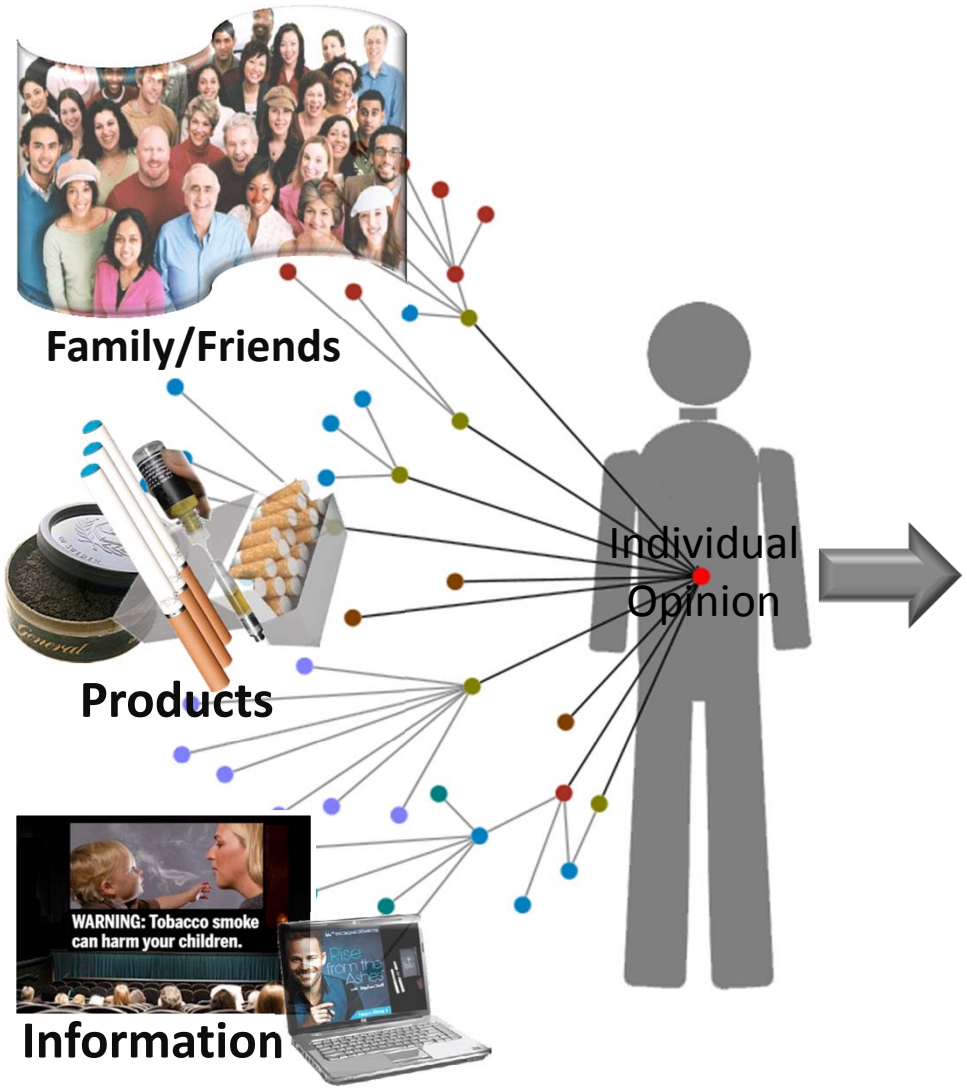
- *Surgeon General's Report, 2012*

Network Influences Individual Opinions



Sandia National Laboratories

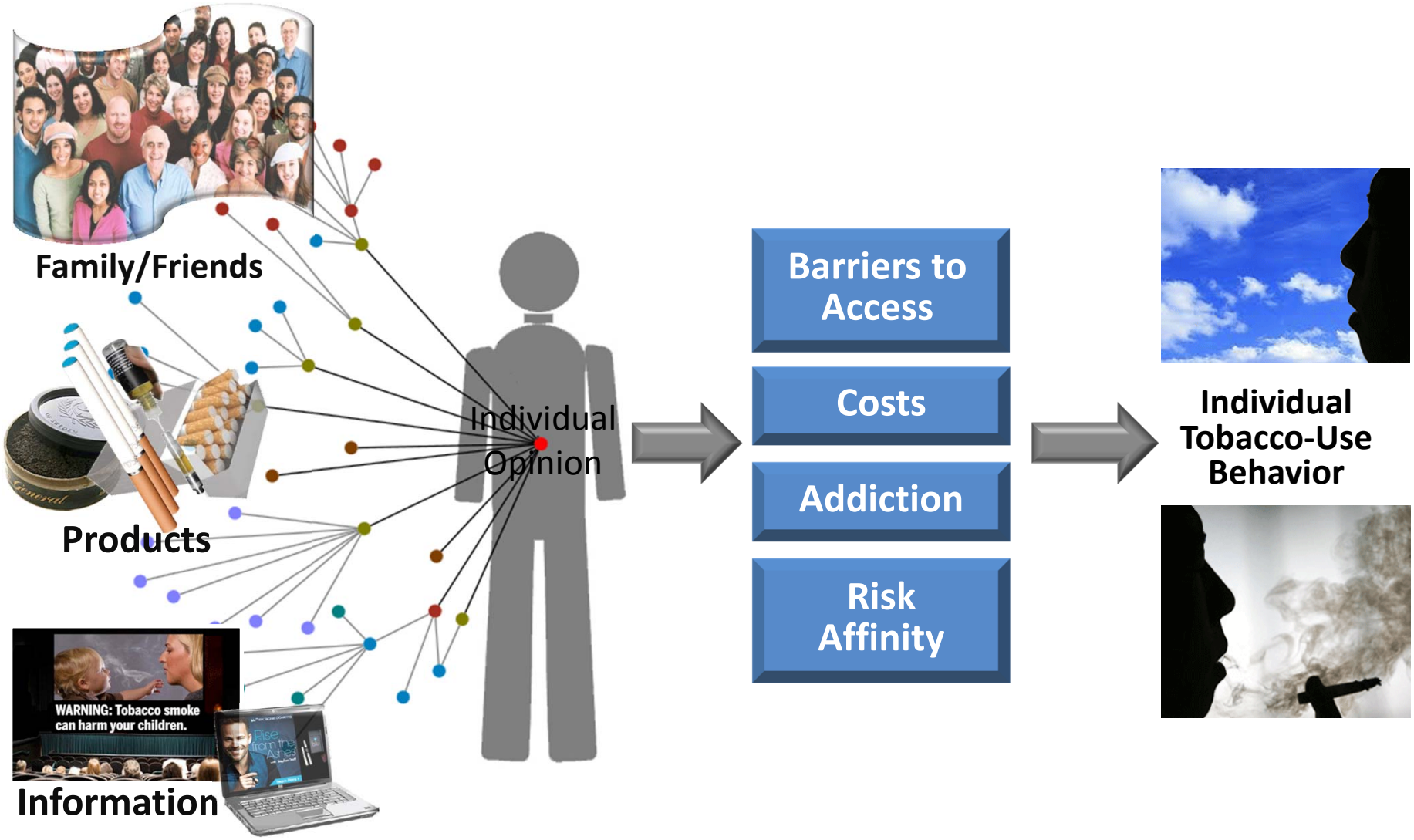
Opinion Drives Behavior Transitions



Individual Tobacco-Use Behavior

Sandia National Laboratories

Many Factors Contribute to Behavior Change



Sandia National Laboratories

Tobacco Opinion Moves on Social Networks

$$x_i(t+1) = x_i(t) + \frac{1}{|N_i|}\sum_{N_i} \mu_{ik}[x_k(t) - x_i(t)]$$

Update Rule: *Adjust individual's opinion by mean scaled opinion differences of opinion and neighbors'*

$$N_i \in S_i : |x_k(t) - x_i(t)| \leq \varepsilon_i$$

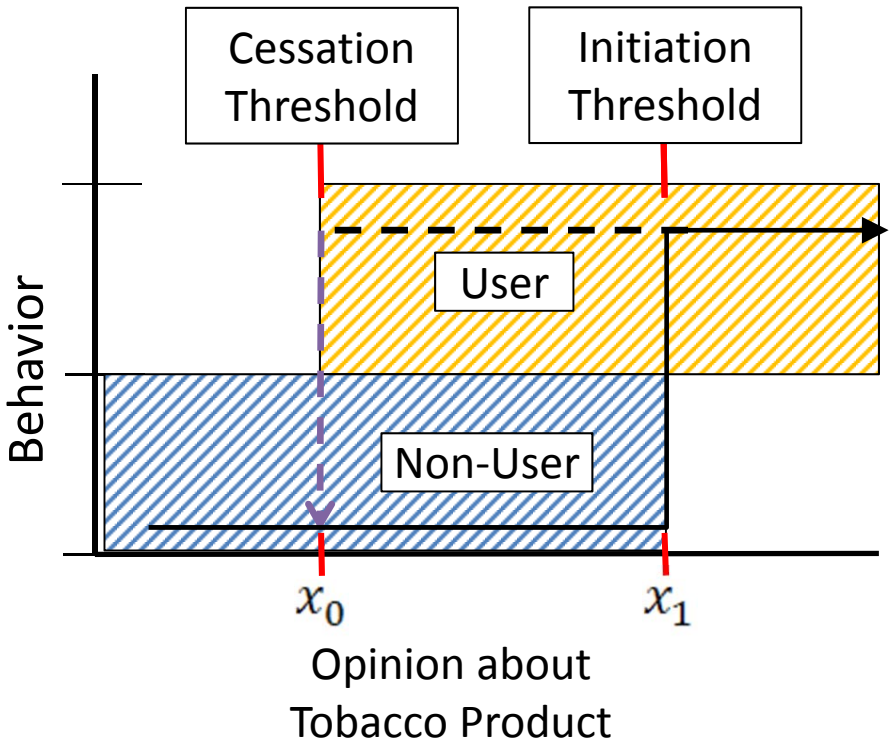
Effect of tolerance*: Ignore neighbors whose opinion is outside individual's tolerance.*

S_i: Set of out-degree neighbors
N_i Out-degree neighbors within tolerance
ε: Tolerance
μ: Plasticity
x: Opinion

Sandia National Laboratories

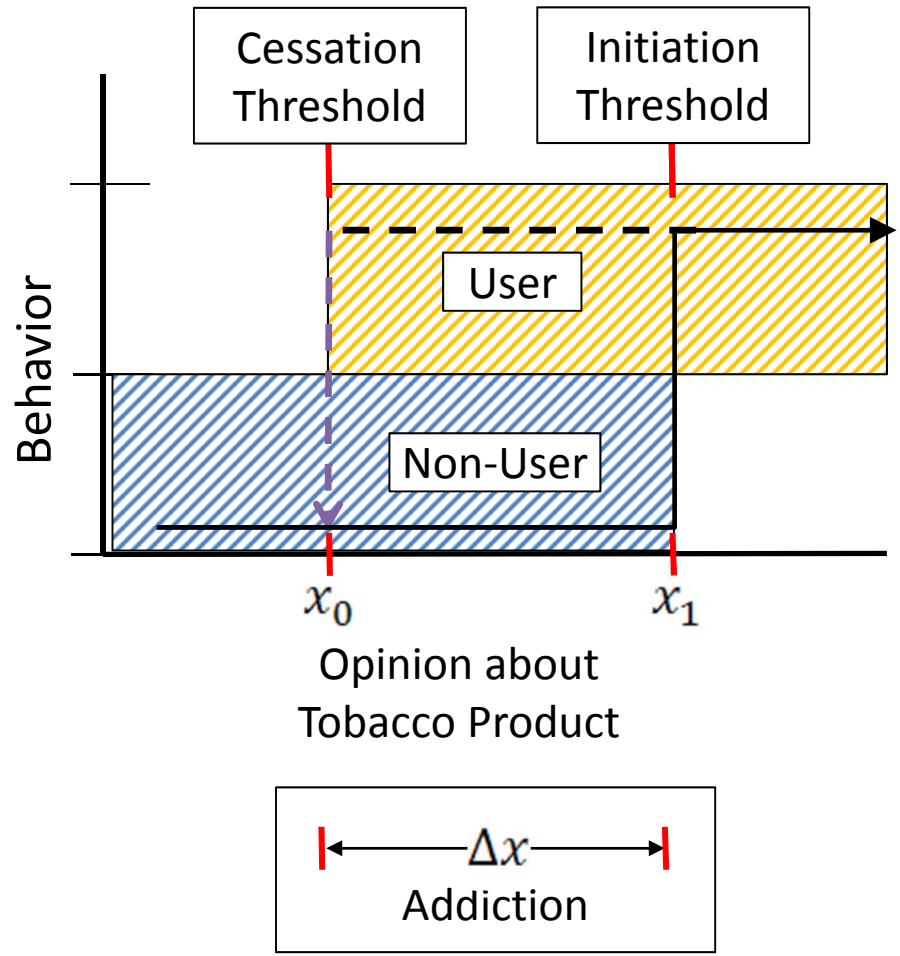
Opinion to Behavior Mapping with Addiction

- Individual tobacco-use behavior is path-dependent (hysteretic)
- Start using tobacco when opinion exceeds initiation threshold (solid line)
- Quit using tobacco when opinion falls below cessation threshold (dashed line)



Opinion to Behavior Mapping with Addiction

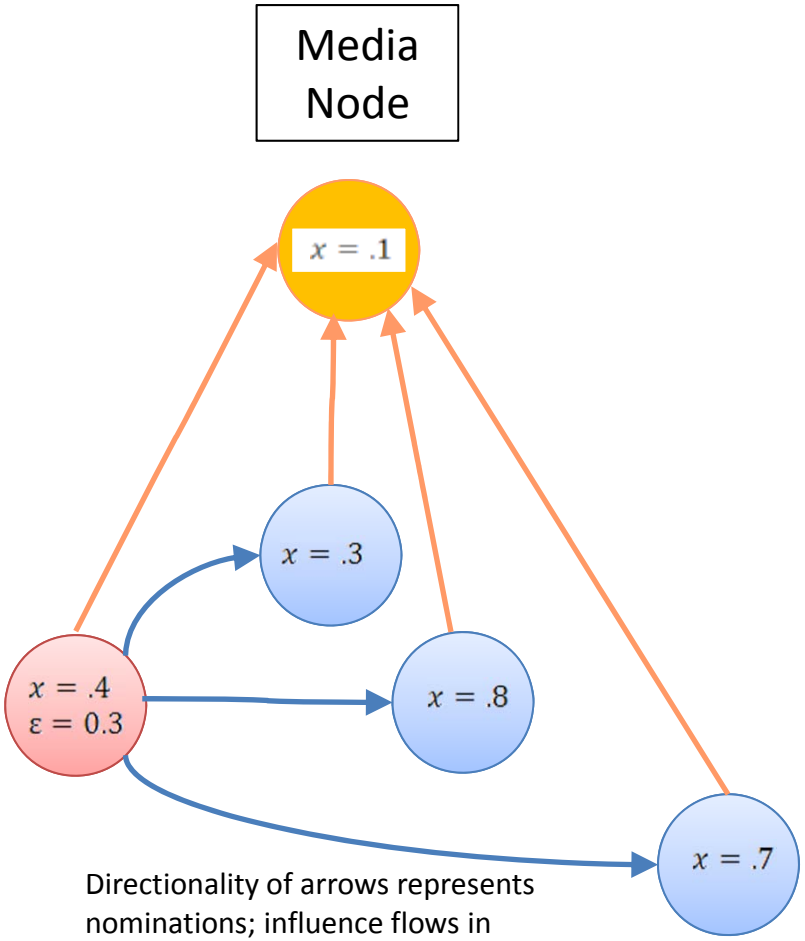
- Individual tobacco-use behavior is path-dependent (hysteretic)
- Start using tobacco when opinion exceeds initiation threshold (solid line)
- Quit using tobacco when opinion falls below cessation threshold (dashed line)
- Magnitude of the hysteresis effect represents strength of addiction



Modeling Media and Education

Information sources modeled as social network nodes broadcasting messages to individuals

- Pro-tobacco information sources represented as high-opinion messages, values closer to 1.0
- Public-health information sources represented as low-opinion messages, values closer to 0.0



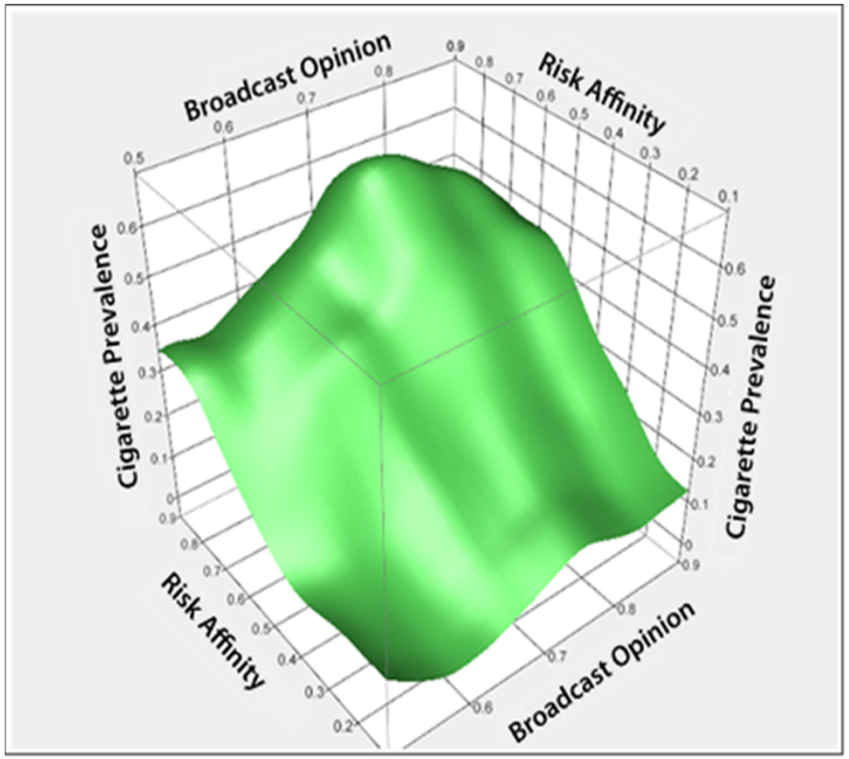
Directionality of arrows represents nominations; influence flows in opposite direction

Modeling Risk to Address Multiple Products

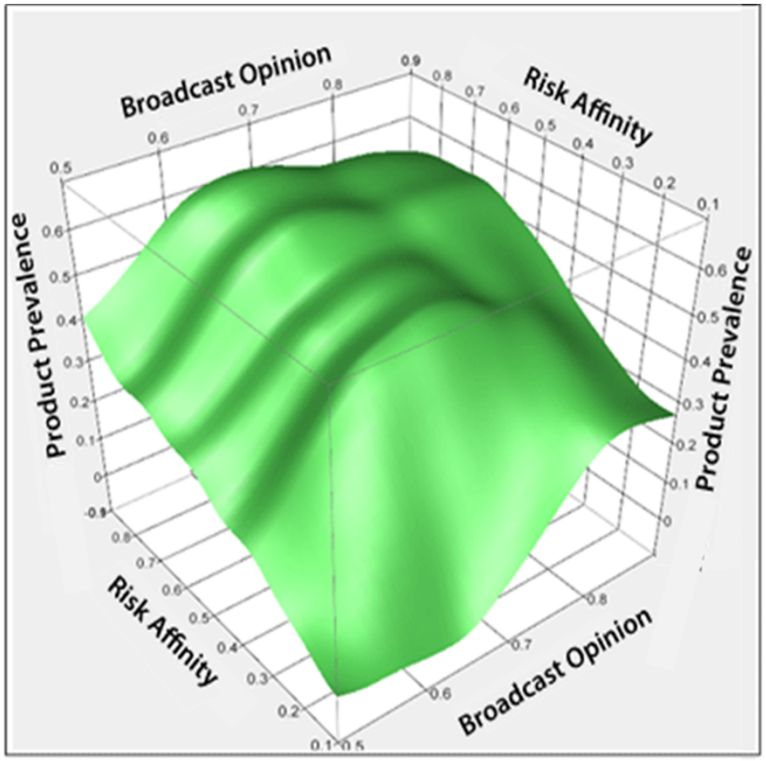
- Risk can be a key component in modeling behavior change in multiple-product regimes
  - Risk affinity: Property of individuals
  - Risk perception: Property of a class of products
- Assumptions:
  - Risk-tolerant individuals neglect potential harm and will use products with high perceived risk
  - Risk averse individuals will use only products with low perceived risk
- Incorporating risk can provide comparative values for use of multiple products

Sandia National Laboratories

Preliminary Results: Effect of Perceived Risk on Modeled Prevalence



Cigarettes



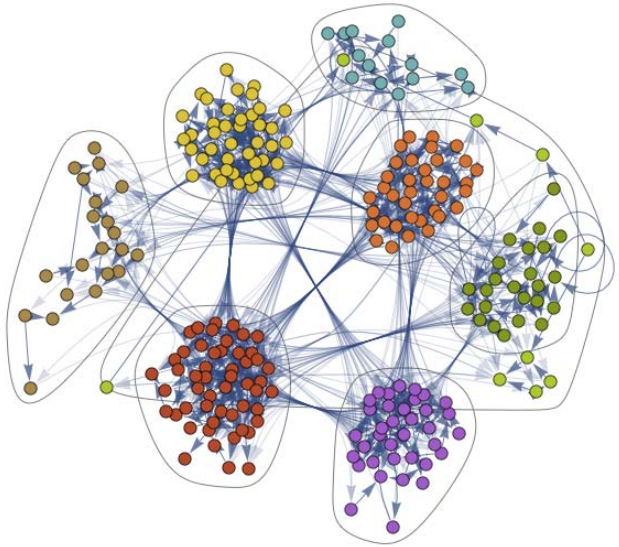
Lower Perceived Risk Product

Lower perceived risk increases modeled prevalence for the same ranges of risk affinity and broadcast opinion (greater volume under the surface)

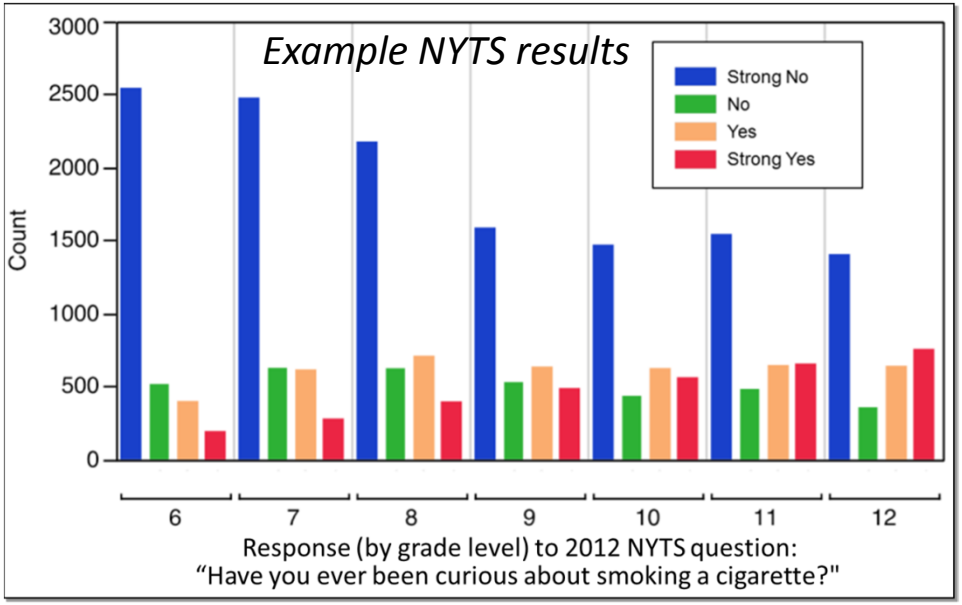
Sandia National Laboratories

Ongoing Validation of SnapDragon

- Agent based models notoriously difficult to validate
  - Abstract software representation of complex real-world psychological and sociological phenomena
  - Survey data are often incomplete and noisy
- Compare results against real-world datasets
- Validating individual assumptions contained within the model



High school social network from Valente et al., in press



NYTS: National Youth Tobacco Survey

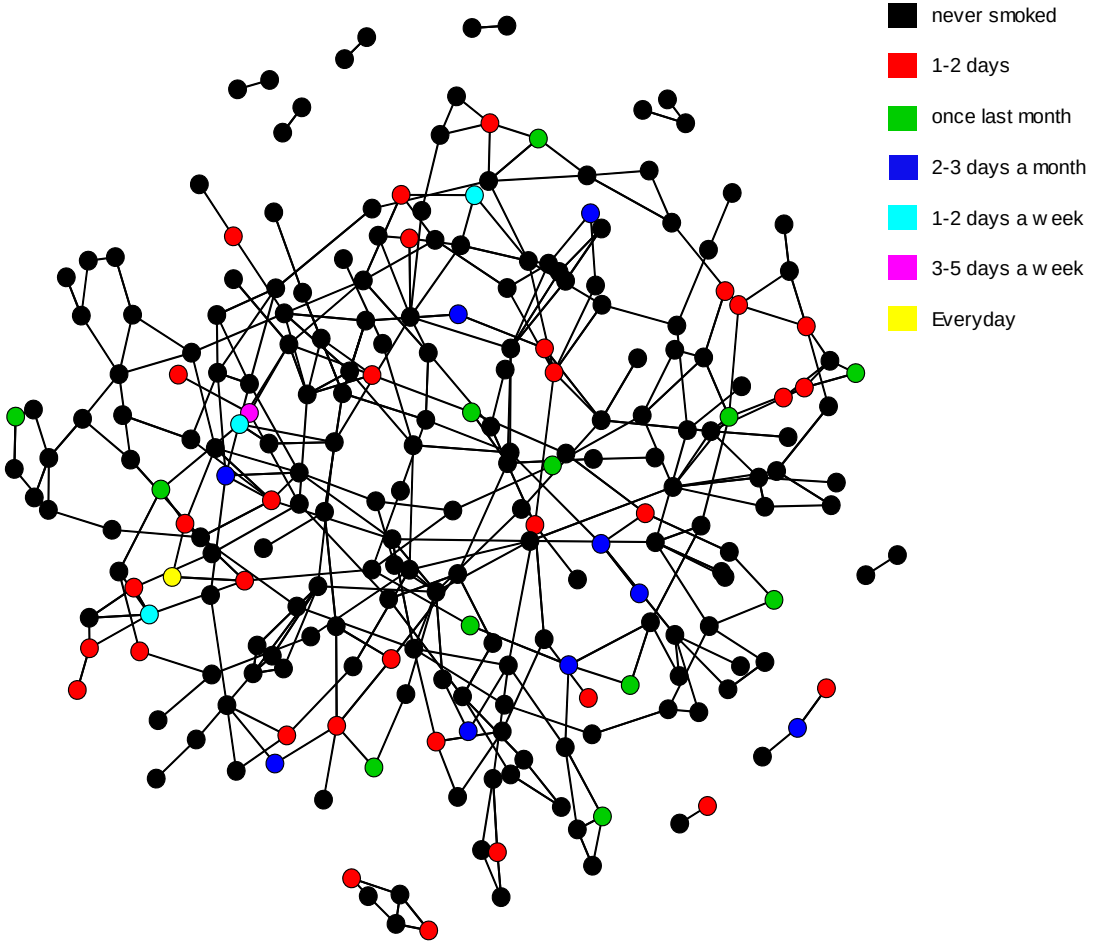
Example Statistical Validation of Model Assumption

Adolescent tobacco users preferentially associate with other tobacco users.

Validation Data

Social Network (links):
Friendship nominations within senior class at a Los Angeles area High School

Tobacco Use (colors):
Response to question:
How many times have you smoked in the last 12 months?



Sandia National Laboratories

* From Variations in network boundary and type: A study of adolescent peer influences, Thomas W. Valente et al., [in press: *Social Networks, an international journal of structural analysis].* Support for this research was provided by NIH/NCI grant #CA157577-02S1 (Valente, PI)

Network metric of interest: Assortativity Coefficient

- Assortativity is bias in favor of connections between network nodes with similar characteristics
- Traditionally measured with Assortative Mixing Coefficient (Newman 2003)
  - Correlation between two nodes at end of every graph edge
  - Analogous to Pearson's Correlation Coefficient (R) applied to network
  - Defined for continuous, not interval data
  - No rigorous test of statistical significance
- Ideal method would be:
  - More flexible to data type and relationships
  - Robust to outliers and skewed distributions
  - Clearly defined significance test

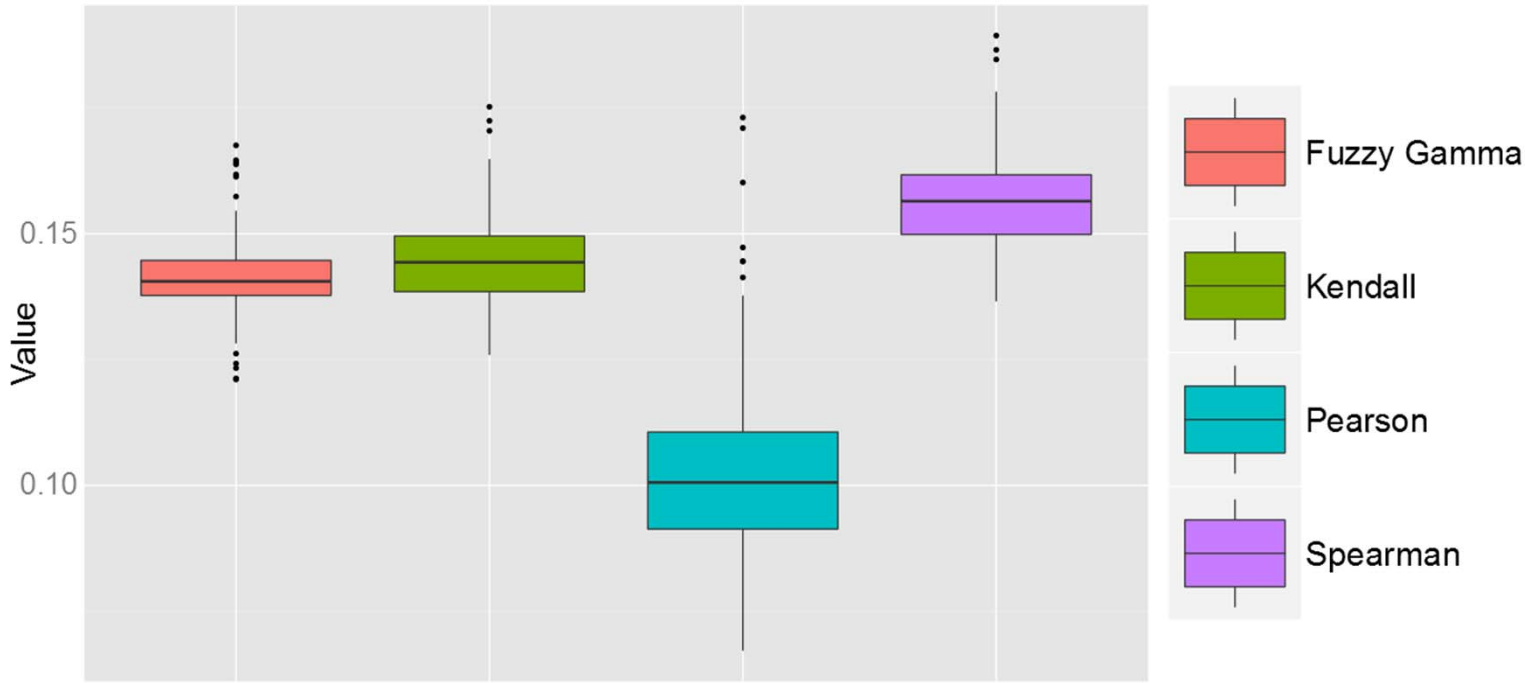
Possible Alternative Methods for Assortativity Coefficient

- Pearson product-moment correlation coefficient
  - Measure of linear statistical dependence between two continuous variables, not robust when normality is violated
- Spearman's rank correlation coefficient
  - Nonparametric measure of statistical dependence using a monotonic function. Ranked version of Pearson
- Kendall rank correlation coefficient (τ_b)
  - τ_b is a common rank correlation coefficient that takes into account ties, a pair $\{(x_i, y_i), (y_j, y_j)\}$ is said to be tied if $x_i = x_j$ or $y_i = y_j$
- Fuzzy gamma rank correlation coefficient
  - Gamma rank correlation coefficient is a well-known rank correlation measure frequently used between two ordinal variables
  - to increase the robustness of the coefficient when dealing with noisy data, a fuzzy order relation is used

Sandia National Laboratories

Comparing Robustness of Assortativity Methods

- Distributions after perturbing a single node from our network with an extreme value (outlier)
- Fuzzy gamma rank correlation coefficient most robust estimation of correlation based on smoking behavior



Sandia National Laboratories

Permutation Significance Test

- Permutation test advantages for assortativity coefficient
  - Distribution free approach
  - Independent of sample size.
- Randomize nodal attributes in the network.
- If there is no assortativity, then the distribution of the correlation coefficient should not change under random permutations
- H_0: There is no assortative mixing by smoking habits
- The p-value is the probability that assortativity under the null distribution exceeds the observed assortativity in the network

Permutation Test Results

- 1000 random permutations without replacement
- The standard assortativity method (Pearson) is the only method that is not statistically significant at p=0.01

Correlation Coefficient	Observed	*P*-value
Pearson	0.1105843	0.01389861
Kendall	0.1470971	0.0019998
Spearman	0.1589439	0.00239976
Fuzzy Gamma	0.1425991	0.00229977

Conclusions

- Snapdragon is an agent-based social-network model of tobacco use
- Model is currently under validation
- Example validation task showed
  - Standard method of network assortativity measurement is not well suited to interval-based validation data
  - Test of four possible assortativity metrics showed Fuzzy Gamma is most robust to outliers and skewed distributions
  - Permutation method useful for significance testing
  - Proposed alternative assortativity measurements have smaller p values than standard method.

Thank You