Title
Simulating social network-based interventions for adolescent cigarette smoking

Journal
Social Science & Medicine, 380

ISSN
0277-9536

Authors
Lakon, Cynthia M
Wang, Cheng
Hipp, John R
et al.

Publication Date
2025-09-01

DOI
10.1016/j.socscimed.2025.118196

Simulating Social Network-Based Interventions for Adolescent Cigarette Smoking

Cynthia M. Lakon

Cheng Wang

John R. Hipp

Carter T. Butts

May 10, 2025

Post-print. Published in Social Science & Medicine Vol 380 Sep 2025 Article 118196

Word count: 7,925

Word count (including references): 9,020

Simulating Social Network-Based Interventions for Adolescent Cigarette Smoking

Abstract

Social network-based adolescent substance use interventions have demonstrated potential for reducing adolescent cigarette smoking. This approach is premised upon leveraging youths' social networks for the diffusion of peer influence. Determining which adolescents to select in network interventions for reducing smoking is a major consideration. We utilize a simulation approach that first estimates Stochastic Actor-Oriented models (SAOM) of adolescent smoking using data from two of the largest schools from the longitudinal saturation sample of the National Study of Adolescent to Adult Health (Add Health) (n = 3,154). We then conduct Agent-Based Simulation models which mimic the consequences of intervention strategies selecting adolescents in network positions and structures that are salient for smoking and the diffusion of peer influence within school-based networks, and we select adolescents smoking at different levels. Our findings indicate that selecting adolescents occupying central network positions yielded the greatest reductions in the number of smokers in a school, one year post intervention. Moreover, our findings indicate that in the school with the higher smoking prevalence, there was a beneficial network multiplier effect one year later, which resulted in more non-smokers than those smokers initially intervened upon. When examining the effects of varying the magnitude of peer influence, we find that targeting central positions in networks led to even greater decreases in smoking in schools with higher levels of peer influence. Our findings highlight interdependence and sensitivity of peer influence to network position and have implications for informing school-based network interventions for adolescent smoking.

Simulating Social Network-Based Interventions for Adolescent Cigarette Smoking

Introduction

Social network interventions have demonstrated potential for reductions in adolescent substance use.[1-3] Coined "network interventions,"[2] particularly those operating through peer networks, have shown reductions in adolescent substance use.[1,4] These interventions leverage the diffusion of peer influences throughout adolescent peer networks.[2] Peer influence may affect adolescent substance use via mechanisms including observational learning[5] and repeated peer reinforcement for talk regarding antisocial acts and attitudes.[6] Moreover, Differential Association Theory suggests that the more delinquent behavior one's peer group members engage in, the more likely an adolescent will be to adopt these "definitions" of the appropriateness of such risky behavior.[7] [8] Peer influence can be risk promotive or risk protective for adolescent smoking, likely depending on the nature of the pervasive norms sanctioning or discouraging smoking in a school. Numerous studies indicate that peer influence positively relates to adolescent smoking.[9-12] Moreover, there is evidence from studies that peer influence differs across school contexts.[13]

Existing studies have utilized Stochastic Actor Oriented Model (SAOM) and Agent-Based Simulation modeling approaches to simulate adolescent peer network-based interventions, including examining peer influence effects for adolescent substance use. In important early work, Schaefer et al. 2013[14] simulated the effects of altering both smoking based popularity and peer influence simultaneously and found that increasing peer influence and smoking based popularity jointly increased school smoking prevalence. Simulation studies have since demonstrated that increasing the magnitude of peer influence *decreases* adolescent smoking among youth in National Study of Adolescent to Adult Health samples,[15] a counterintuitive finding that provided insight into the value of simulation studies for understanding peer influence effects on school smoking prevalence. Moreover, other research indicates that the relationship between peer influence and adolescent smoking can vary across contexts, such as the smoking prevalence in high schools.[16] Furthermore, other research examined intervening on youth in school-based peer network positions and found this approach was more effective at reducing alcohol use compared to non-network interventions.[3] In addition, other research indicates that selecting on network position, and extending to other segmentation approaches, may be even more effective in reducing long term alcohol use among adolescent youth in contexts with high levels of peer influence.[17]

Building on this prior research, the current study extends previous work as we examine selecting youth in theoretically salient network positions and those embedded in key network structures for adolescent smoking for intervention, and those who smoke at varying levels, to examine the effects on smoking. We also vary the magnitude of peer influence throughout the study period to observe its effect on smoking across school contexts. We utilize an Agent-Based Simulation modeling strategy[18] to investigate the effects of our interventions and of altering the strength of peer influence for adolescent smoking. An innovation of the current study is that we examine whether there is interdependence in peer influence, network position and structure, for adolescent smoking and as such, we examine the sensitivity of peer influence to network position and structure as they relate to adolescent smoking. Agent-Based Simulation modeling provides a low-cost approach prior to going into the field for gaining insights into the consequences of interventions for behavior change and maintenance. In addition, this approach allows for altering one intervention condition at a time, which is not feasible in most real world settings.

Most importantly, this approach allows for testing the effects of upper and lower bounds of peer influence as it affects adolescent smoking.

Network-Based Selection Strategies

Centrality Measures

One important consideration for network interventions is which social positions within peer networks are most beneficial to select to decrease adolescent smoking. Studies indicate the importance of centrality measures for adolescent smoking.[19,20] In-degree centrality, a measure which reflects the number of incoming nominations - and indicates popularity - has been found to be positively associated with adolescent smoking.[20,21] Popular youth are trendsetters, who are uniquely positioned to widely transmit peer influences, norms and other information throughout their networks. Prior studies suggest that popular students can be seeds to enhance the spread of network interventions in schools.[22,23] Adolescents with high out-degree centrality regard many as friends, making them susceptible to the direct influence of many peers.[20] As friendship ties are highly reciprocal, in-degree and out-degree are positively correlated, youth high on both in-degree and out-degree have direct connections to others involving them in exchanges of information, support, and normative sanction with the potential to impact their entire network.[24, 22,23]

While degree is a critical dimension of social position, other forms of centrality are also potentially important for adolescent smoking. We investigate the impact of selecting adolescents based on closeness centrality (i.e., a measure of the extent to which individuals have short paths to others in the network),[24] and betweenness centrality (i.e., a measure of the extent to which an individual lies on the shortest paths between others in the network).[24] Another measure, eigenvector centrality, is related to both prominence (i.e., having many ties to others who themselves are well-connected) and the ability to robustly diffuse information or influence throughout the network[25] and is thus of great relevance to peer influence.

Bridging Measures

In addition to conventional centrality measures, another important basis for selecting youth for intervention may be their role as brokers or bridges between otherwise locally disconnected peers.[26] Although some forms of centrality (e.g., betweenness) are related to bridging, individuals may occupy local bridging roles without being central in other respects and, conversely, individuals may be highly central in other ways without serving as a bridge; these roles have hence attracted attention as being important youth to select for intervention. Valente and Fujimoto (2010) proposed a measure that reflects the degree to which a node occupies a brokerage position, grounded in Granovetter's work on the strength of weak ties.[27] This was extended by Everett and Valente,[28] who proposed a modified bridging measure, using betweenness centrality as a foundational concept. We employ both of these bridging measures herein. Substantively, an adolescent in a bridging position serves as the conduit between groups or individuals that are not otherwise tightly connected within the network, a role that can confer power and prominence via the ability to govern the flow of information between groups, broker exchanges, or simply supply novel information of resources.[29] Such individuals have been hypothesized to be well-positioned to transmit and enforce norms, making them potentially important youth for intervention.[26,28]

Group Measures

While certain types of power and influence may accrue from bridging roles, settings involving trust, loyalty, and normative conflict can favor individuals who are instead *cohesively embedded* in social groups.[30] As with bridging roles, such local embeddedness is related to, but distinct from, standard centrality measures: an individual may be highly central in many respects without being cohesively embedded within a peer group, and by turns such an individual who is cohesively embedded in for instance a small group may not be particularly central in other respects. Herein, we consider two forms of local embeddedness, which are measures of network structure, which describe linkages between individuals in a network. First, we consider the individual's *core number*. A *k*-core is a maximal set of individuals such that every member of the set is tied to at least *k* others in the set; the higher the core number (*k*) of a group, the more strongly it is connected by redundant local connections, and the more difficult it is to remove its members. An individual's core number is the number of the most cohesive (highest *k*) core to which they belong and is hence a measure of their local embeddedness. Deep core membership (high core number) may confer advantage in contagion-like influence processes,[31] and may be particularly important for complex contagion[32] in which multiple exposures from distinct individuals are required to trigger behavioral change. A distinct notion of local embeddedness is the local clustering coefficient, which measures the density of their egocentric network (i.e., the network consisting of a focal individual, those to whom they are tied, and the ties among those individuals). Substantively, an adolescent's clustering coefficient can be interpreted as the fraction of their friends who are friends with each other (conventionally taken to be zero for those with 1 or fewer friends). A local clustering coefficient of zero indicates a position in which none of one's friends are friends with each other (no local cohesion) and a coefficient of 1 indicates a position in which all of one's friends are friends with each other (complete local cohesion). Crucially, friendships for an individual with a high clustering coefficient are embedded in triadic relations with third parties (i.e., the focal individual and their friends generally have friends in common), a key condition for the development and maintenance of trust.[33] While such an individual may not be otherwise central - and may not belong to a larger cohesive group - the local embeddedness of their friendships within shared partnerships may make them better conduits of normative influence than other sorts of ties.[24]

Another important consideration for selecting adolescents for cigarette smoking interventions is the levels at which youth smoke. Adolescents who smoke at varying levels may differ from each other in numerous ways, including their personality, depressive symptoms, and other individual differences.[34] Therefore, there may be differences in intervention effectiveness depending on whether high or low smoking youth are intervened on.[1,14,35,36]

In sum, our study will examine the consequences, over time, for the school smoking prevalence when selecting adolescent youth for intervention based on the following: 1) occupying certain network positions; 2) being embedded in network structures; 3) heavy, moderate, or light smoking. We will also examine the effect of varying peer influence throughout the study duration across school contexts and with simultaneously selecting network position and structural measures for intervention.

DATA AND METHODS

The current study addresses these research questions using Stochastic Actor-Oriented models (SAOM).[37] As in prior simulation studies,[3,14-16,38] we will first estimate SAOMs with data from the longitudinal sociometric network oversample of the National Longitudinal Study of Adolescent to Adult Health (Add Health).[39] Next, we will use the parameters from the SAOMs to conduct our Agent-Based Simulation models. The current study was reviewed and granted approval under exempt review by the institutional review board at the university where the study was conducted.

Data

Data for this study are from the longitudinal, school-based network component of the large, nationally representative Add Health Study,[39] a 16-school saturated sample in which all students self-reported their smoking behavior and nominated up to five female and five male best friends, at three time points. The two largest schools provide enough statistical power for the current study, and we therefore construct network samples from these two schools: one, a rural Midwestern public high school (n = 976) referred to as "Jefferson High" has a relatively higher smoking prevalence,[40] and the other, a suburban Northeastern public high school (n = 2,178), referred to as "Sunshine High" has a relatively lower smoking prevalance.[41] There are three time points of data utilized in our study: t_1 is the In-School Survey conducted in 1994, t_2 is the Wave I in-home survey conducted in 1995, and t_3 is the Wave II in-home survey conducted in 1996.

Measures

Two time-varying dependent variables are constructed for each school sample. One captures smoking behavior and is based on the original survey questions measuring smoking behavior, (i.e., "During the past twelve months, how often did you smoke cigarettes?") utilized during the In-School Survey and (i.e., "During the past 30 days, on how many days did you smoke cigarettes?”) utilized during the Wave I and Wave II in-home surveys. Consistent with prior research,[9] we recategorize smoking behavior into four levels based on cigarette use activity over the previous 30 days, with 0 = "never", 1 = "1~3 days" (light smoker), 2 = "4~21 days" (moderate smoker), and 3 = "22 or more days" (heavy smoker) for the three time points. The second dependent variable captures the network ties of adolescents, based on questions asking each student to nominate up to five female and five male best friends. We construct an adjacency matrix at each time point for each school.

We specified our SAOM to predict smoking behavior and network ties by building on previous research[9] and incorporating the transitive reciprocated triplets effect.[42] Detailed information on the constructed variables can be found in the Supplemental Materials. At time 1, the proportion of non-smokers in Jefferson High was 42%, whereas heavy smokers constituted 28% of students. The level of smoking was considerably lower at Sunshine High, where 69% of the students were non-smokers and 9% were heavy smokers at time 1.

Methods

Model estimation

We first estimate the parameters of the SAOM, using the R-based Simulation Investigation of Empirical Network Analysis (RSiena) software package[43] to estimate a SAOM[37] for each school. The SAOM strategy is based on a latent, continuous time process in which an actor (or focal individual) *i* makes decisions at random times regarding the state of his or her behavior (specifically, considering a unit change in the level of smoking behavior) and outgoing ties (specifically, the state of an actual or potential tie to a random alter or selected individual *j*). The decisions are represented in a logistic choice model that combines a set of objective functions $f_i(\beta, x) = \sum_k \beta_k s_{ik}(x)$, where β_k is the estimated parameter for the *k*th actor-specific effect $s_{ik}(x)$, and *x* is the feature of the joint network/behavioral state. A positive parameter suggests the direction of change in the network/behavioral state, while a negative parameter suggests the avoidance of this change. The rate function $\lambda_i(\alpha, x)$ indicates the number of opportunities the actor *i* will change his or her behavior or outgoing ties across consecutive time periods. We used Method of Moments (MoM) to estimate the objective and rate functions simultaneously, yielding one equation for the smoking variable and another equation for the network tie variable. Model convergence is assessed using criteria of both *t* statistics for deviations from observed statistics (ideally less than 0.10 for each parameter) and the overall maximum convergence ratio (ideally less than 0.25).[43] The goodness-of-fit (GOF) is assessed with a Monte Carlo test of Mahalanobis Distance statistics,[16] and our model fit satisfactorily.

Simulation strategy

The general strategy for our simulation models is as follows: 1) select adolescents in the network at the first time point (t_1) based on their network position and/or smoking level, 2) simulate the intervention between t_1 *and* t_2 which results in some percentage of selected participants becoming non-smokers by t_2, by altering the smoking level of these students to zero to mimic the consequences of the end of a successful intervention, and 3) simulate forward the network and behavior to t_3, one year later, to assess whether behavior change was maintained post intervention based on school smoking prevalence. For each condition, we conducted 1,000 simulations and reported the average of the results. We also conduct simulations in which we alter the strength of the peer influence parameter to observe whether changes in the magnitude of the peer influence effect results in changes in the school smoking prevalence. Using the estimated SAOM parameters, we conduct a series of simulations adopting the following strategies.

Intervention 1:Random heavy smokers. We randomly select 20% of heavy smokers (i.e., those with a value of 3 on the smoking variable at t_1). In Jefferson High, 270 (or 27.7%) of the 976 students at t_1 are heavy smokers, and thus we select 54 for the intervention. In Sunshine High, 195 (or 9%) of the 2,178 students are heavy smokers at t_1, and thus we select 40 students for the intervention. We selected these same numbers of students for all subsequent interventions we describe.

Intervention 2: Network-based heavy smokers. For each school, we select 20% of heavy smokers at t_1 based on having the highest values on one of the nine network properties listed in Table 1, change their smoking levels to 0 at t_2, and simulate the network forward to t_3. We then repeat this strategy by selecting adolescents with the highest values on each of the remaining eight network properties. This strategy tests whether an intervention will be more effective when selecting heavy smokers who occupy particular network positions, rather than selecting heavy smokers at random. Given the importance of being central in a school-based network for smoking, we include five centrality measures: in-degree, out-degree, closeness (using Gil-Schmidt power index for disconnected networks), betweenness, and eigenvector.[24] To account for the possible importance of small group cohesion, we computed the local clustering coefficient.[24] To capture the possible importance of linkage to highly connected subgroups in the broader school network, we constructed a *k*-core measure[24] indicating the subgraph containing *i* with the largest *k* such that all members of the subgraph have at least *k* friends within the subgraph. Finally, considering that individuals in bridging positions may be important as they can link together various highly connected regions of the network, we computed two bridging measures created by Valente et al.[26,28] Given the possible concern that missing data might bias the centrality measures, we conducted simulation experiments using the 84 schools from the Add Health public use data, and report in the Supplemental Materials that the three measures exhibit considerable robustness (Figures A1 and A2).

<<<Table 1 about here>>>

Intervention 3: Light or moderate smokers. We modified Interventions 1 and 2 by sequentially selecting light smokers (i.e., value of 1 on the smoking variable), or moderate smokers (i.e., value of 2 on the smoking variable) at t_1, and preferentially selected those with the

highest values on each of network measures. We then simulated the network and behavior forward in time and computed the levels of smoking in the schools at t_3.

Intervention 4: Network-emphasized smokers. We altered Intervention 2 by selecting adolescents with the highest values on each of the network properties and who smoked at any level (i.e., 1 to 3 on the smoking variable).

We calculate the degree of effectiveness of the intervention with three different calculations. First, we compute the *initial effect* (IE), which is the percentage point increase in non-smokers that would be observed if all adolescents who initially transitioned to non-smokers due to the intervention remained non-smokers by *t3*. This is calculated as (NS+I)/N, where NS is non-smokers at *t3*, I is the number intervened on, and N is the school size. We computed the *long term effect* (LTE) as the difference between the percentage point increase in non-smokers in a given intervention compared to when there is no intervention. This is calculated as $\%NSI_3 - \%NS0_3$, where $\%NSI_3$ is the percent non-smokers at *t3* after an intervention, and $\%NS0_3$ is the percent non-smokers at *t3* with no intervention. Finally, we compute the *Long-term to initial effect ratio* (LTIER), which is the ratio of LTE to IE (LTE/IE).

Modifying the size of the peer influence parameter across different school settings throughout the study duration. For these simulations, we modify Interventions 1 and 2 by assessing the effectiveness of the selection strategies when simultaneously altering the strength of the peer influence parameter at all three time points and then simulating the data forward in time. We assessed five peer influence estimates: 1) zero; 2) 50% of the actual estimate; 3) the estimate from the SAOM; 4) 50% larger; and 5) 100% larger.[a] The summary statistics for the network measures and smoking levels are shown in Table 2.

<<<Table 2 about here>>>

RESULTS

We begin by showing the correlations among the network measures we use for determining whom to select for the intervention, as well as their correlations with smoking at time 1 (Table 3). In the top panel, we see that the correlations in Jefferson High are somewhat high, with correlations in the range of .7 to .8 for some of the network measures. The bottom panel shows that the correlations in Sunshine High are much lower. The correlations of these network measures with smoking levels are quite low in both schools, and always below 0.1.

<<<Table 3 about here>>>

Estimates from the Stochastic Actor-Oriented models

Herein, we briefly highlight the key results from our estimated SAOMs, as our simulation results are the focus of the current study, and a more detailed discussion of the SAOMs is presented in the Supplemental Materials. Our SAOMs indicated a peer influence effect in both schools, as students tended to adopt the smoking levels of their friends. There is also evidence of a homophilous selection effect, as students in both schools were more likely to form friendships with others at similar smoking levels. In both schools, students tended to nominate other students as friends who displayed relatively high levels of smoking, implying that some popular students engaged in high levels of smoking. The goodness-of-fit (GOF) assessment is good,[44] and the Monte Carlo test of Mahalanobis distance statistics yields a *p*-value greater than 0.05,

[a] For example, the estimated influence parameter for Jefferson High is 0.76, indicating that the odds of an adolescent adopting the smoking behavior of their peers is increased just over 100% for a one unit increase ((exp(.76)-1)=1.14), and a 50% increase in the parameter to 1.16 implies an increase just over 200% ((exp(1.14)-1)=2.13).

indicating that the model adequately reproduces the smoking level distribution in the observed data for each school.

Agent-Based Simulation Models

Using the parameters from our SAOMs, we next conducted our Agent-Based Simulation models, to observe how our intervention conditions affected smoking behavior over time and post intervention.

Interventions 1 (random heavy smokers) and 2 (network-based heavy smokers). Figure 1 shows the percent non-smokers at t_3 in the higher smoking prevalence school (Jefferson High) based on several different selection strategies. The first box plot shows the results with no intervention along with the posterior interval based on the simulations and is referred to as the "no intervention" condition, with 48.6% (or n=474) non-smokers at t_3, this value is based on simulating the school using the estimated parameters from the SAOMs. Given that 20% of heavy smokers (n=54) were intervened upon in the various strategies, this implies that there would then be 54% [(474 + 54)/976] non-smokers total in the school by t_3 if they all remained non-smokers. This indicates a 5.5 percentage point (54/976) increase in non-smokers in this school, which is the *initial effect* (IE). However, it is possible that some of these intervened upon students will transition back to smoking by t_3, or that other students in the school will change their smoking status by t_3, due to peer influence effects.

<<<Fig. 1 about here>>>

The remaining box plots in this figure show the percentage of non-smokers in the school at t_3 based on the simulated networks from different intervention selection strategies. For example, the second box plot indicates that if 20% of heavy smokers are selected *at random* for the intervention (i.e., random heavy smokers intervention strategy #1), then the average simulation results in 52.3% non-smokers at t_3, which is a 3.7 percentage point improvement over the no intervention value (48.6%) and is the LTE. The ratio of LTE to IE gives us the *long-term to initial effect ratio* (LTIER), which is 67% (3.7%/5.5%= 67.3%) in this case. Thus, only 67% as many non-smokers still exist at t_3 in the random selection condition, which occurs because of peer influence effects inherent within this school network that impact smoking behavior by t_3.

The remaining box plots in Figure 1 show the results when selecting adolescents engaging in high levels of smoking for the intervention based on various social network properties (network-based heavy smokers intervention strategy #2). The third and fourth box plots from the left (20% out-degree and 20% in-degree) select adolescents based on having high out-degree or in-degree values, respectively. When selecting based on out-degree, there is an improvement over the random approach as the percent non-smokers at t_3 rises to 54.2%, and when selecting based on in-degree (a proxy for popularity of students), the percentage non-smokers rises even further to 55.5% at t_3. Thus, although 510 students were non-smokers at t_3 in the random intervention, 538 students were non-smokers at t_3 when selecting based on in-degree. The implication is that whereas 57 students are intervened upon in each strategy, there are 28 more non-smokers at t_3, on average, when intervening based on high in-degree rather than random selection. The next three box plots show strong results when selecting adolescents based on all three of these centrality measures, selecting based on eigenvector centrality is the most advantageous and does just as well as selecting based on in-degree, yielding 55.5% non-smokers by t_3. The remaining four strategies are not as effective: selecting based on the local clustering coefficient (20% lcc) or selecting based on strong linkage to the network more broadly based on *k*-core value (20% k-core) only perform modestly better than the random strategy, whereas the two bridging measures (the two right-most box plots) perform similarly to a random approach.

There are statistically significant differences between all of these network-based interventions and the random intervention strategy, as computed by dividing the posterior intervals by root N.

Thus, selecting based on either in-degree or eigenvector centrality appears to have the best results in the higher smoking prevalence school (Jefferson High). Although the differences compared to the other network selection strategies are narrow, they are statistically significant. Note as well that there is a beneficial network multiplier effect observed, as the LTIER of 125% (6.9%/5.5% = 125.5%) indicates that selecting based on in-degree or eigenvector centrality yields 25% *more* non-smokers at t_3 than were intervened upon. This may imply beneficial effects of peer influence in this model. Furthermore, the best results occur when selecting based on in-degree or eigenvector centrality regardless of the number of heavy smokers selected for the intervention, see Figure 2a which plots the percent non-smokers based on selecting between 10% and 90% for intervention (based on 10% increments), and shows that these two strategies always outperform a random selection strategy.

<<<Fig. 2 about here>>>

Turning to the lower smoking prevalence school (Sunshine High), the intervention selection strategies have a similar ranking in this school compared to the higher smoking prevalence school (see Figure 3). Without an intervention, the first box plot in this figure shows that 72.8% are non-smokers at t_3. Selecting 40 (or 20%) heavy smokers for the intervention implies an increase of 1.8 percentage points (i.e., 40 out of 2,178) in non-smokers to 74.6%, and this 1.8 is the initial effect (IE). The long term effect (LTE) is just 0.4% (73.2%-72.8%=0.4%) when randomly selecting adolescents for the intervention. However, the results can be improved by selecting based on certain network positions. Similar to Jefferson High, we see that one of the best strategies is selecting on high eigenvector centrality (73.6% non-smokers at t_3).

Whereas high in-degree was one of the best strategies for Jefferson High, out-degree is one of the best strategies for selecting adolescents for the intervention in the lower smoking prevalence school (also 73.6% non-smokers at t_3). Moreover, selecting students based on high *k*-core (73.5%) is nearly as effective. Although these are small differences for the network strategies in this school compared to random selection—even the most effective strategy selecting on eigenvector centrality only yields 9 more non-smokers at t_3 compared to randomly selecting heavy smokers—they are statistically significant. Furthermore, there is no beneficial network multiplier effect in this school with lower substance use, as the highest LTIER value is 44% (0.8%/1.8%=44.4%), indicating that there are *fewer non-smokers* at t_3 than were intervened upon, even in the most effective strategies. Selecting students based on closeness or betweenness centrality, the clustering coefficient, or the two bridging measures shows little improvement over a random selection strategy. Furthermore, the rankings are similar regardless of the percent heavy smokers selected for the intervention, as Figure 2b shows that in-degree and eigenvector centrality always outperform the random selection strategy.

<<<Fig. 3 about here>>>

Simulation Intervention 3 (Light or moderate smokers). In simulation intervention 3, we modify simulation intervention 2 by instead intervening on 54 light smokers, and in a separate simulation model, we select 54 moderate smokers for intervention in the higher smoking prevalence school (Jefferson High). There is strong evidence that selecting light smokers rather than heavy smokers for the intervention has much weaker effects. Whereas there were 55.5% non-smokers at t_3 when selecting heavy smokers with either high in-degree or high eigenvector centrality, selecting light smokers yields a value closer to 53% (see Figure A7 in the supplemental material). When selecting moderate smokers, the intervention results are closer to those when selecting heavy smokers, with about 54% non-smokers at t_3 when selecting based on these two network positions (see Figure A8 in the supplemental material). The pattern was similar in the lower smoking prevalence school (Sunshine High): there were 73.6% non-smokers at t_3 when selecting heavy smokers with either high out-degree or high eigenvector centrality, these numbers fell to 73.2% when selecting moderate smokers and 73% when selecting light smokers (see Figures A9 and A10 in the supplemental materials).

Simulation Intervention 4 (Network-emphasized smokers). For our final intervention strategy, we modified Simulation Intervention 2, which prioritized selecting heavy smokers, and in Intervention 4 we instead prioritize the network measures and select for intervention those 54 students with the highest values on the particular network measure, with the only condition being that they engage in smoking *at any level*. This strategy yields results even slightly better than the earlier strategy focused entirely on selecting 54 heavy smokers (20%) who also happen to have relatively high values on the network measures. The results for the higher smoking prevalence school (Jefferson High) are presented in Figure A11 of the supplemental materials. Although most of the network selection strategies do well, the best results are obtained when selecting adolescents for the intervention based on in-degree (55.7% non-smokers at t_3) or eigenvector centrality (55.6%). Selecting primarily based on the local clustering coefficient is the least effective network-based strategy. The results for the lower smoking prevalence school (Sunshine High) are shown in Figure A14 of the supplemental materials. In this school, selecting primarily based on social network measures does not do quite as well as selecting primarily heavy smokers, as the best selection techniques end up with 73.2% non-smokers by t_3.

Assessing direct and indirect effects

Although we have found that there are more nonsmokers at t_3 after these interventions, a question is whether this is due to direct effects—the adolescents intervened on are more likely to not smoke at t_3—or due to indirect effects, in which other adolescents who were not intervened on stop smoking by t_3 due to network peer influence effects.[17] First, for all heavy smokers who were selected for the intervention, we compute their level of smoking by t_3 for each of the simulations. Second, we conducted a similar number of simulations in which there is no intervention, and we compute the level of smoking by t_3 for these same heavy smokers across these simulations. The difference in smoking behavior between the first and second set of simulations just described is shown in Figure 4 as the 'selected' adolescents. We conduct a similar two-step process for all heavy smokers who were *not* selected for intervention. First, we compute their level of smoking by t_3 across the simulations in which an intervention was performed. Second, we compute their level of smoking by t_3 in simulations with no interventions. The difference in these two values are shown as the 'unselected' adolescents in Figure 4. We plot these for three intervention strategies: 1) random; 2) in-degree centrality; 3) eigenvector centrality.

In general, the biggest increase in non-smokers by t_3 occurs among the selected adolescents, with only modest evidence of indirect effects. In the random intervention at t_3 there were 25.1% more non-smokers among the selected compared to how they did in the non-intervention scenario, and 25.2% fewer heavy smokers. Among the unselected group (heavy smokers at t_1 who were not selected), there were just 3.8% more non-smokers from the interventions compared to the non-intervention scenario and 4.1% fewer heavy smokers. For the in-degree intervention, the percent non-smokers among the selected rises to 31.3% more than in a non-intervention scenario, whereas the indirect pattern is a bit stronger here as 6.1% of those not selected are non-smokers by t_3 during an intervention compared to otherwise. When selecting based on eigenvector centrality, the percent non-smokers at t_3 among the selected is slightly higher yet at 35%, whereas the indirect effect is again around 6%. Thus, it appears that it is mostly the selected adolescents who are more likely to become non-smokers by t_3, and not the unselected.

<<<Figure 4 about here>>>

Consequences of varying the magnitude of peer influence

In the final set of simulation results, we assess the effectiveness of the different selection strategies when simultaneously varying the strength of the peer influence effect. We modified the size of the peer influence parameter and then simulated the data forward. We summarize the results for the higher smoking prevalence school (Jefferson High) with an ANOVA model, wherein the outcome is the average percent non-smokers at the end of each simulation run. The predictors include indicator variables for the different intervention selection strategies and indicator variables for the different peer influence parameter values. This model explains nearly 98% of the variance, indicating little evidence of interaction effects between the peer influence parameter values and these intervention strategies.

In Figure 5, the top left panel shows the average number of non-smokers across each of the intervention selection strategies (each of the lines) at different values of the peer influence parameter (measured on the x-axis). The value of 0 on the x-axis shows the average results of the model with the peer influence parameter set to zero; the value of 0.5 shows results for the model with the influence parameter set to ½ the value of the actual model; the value of 1 shows the results for the simulation already shown earlier (the actual value of the peer influence parameter); and values of 1.5 and 2.0 are for models with the peer influence parameter set to

50% and 100% larger than the actual model parameter, respectively. We observe that, in all cases, the number of non-smokers at the end of the simulation is higher for larger values of the peer influence parameter, demonstrating that peer influence has a protective effect for smoking behavior. All intervention selection strategies are more effective than the no intervention condition, which is the bottom line in the figure. The relative ranking of the different strategies is similar to what we showed earlier; nonetheless, the number of non-smokers at the end of the simulation increases monotonically with an increasingly larger peer influence parameter. There is generally a linear effect of a stronger influence parameter on the number of non-smokers for influence values between 50% and 200% of the original value, although there is a nonlinear bend as peer influence approaches zero. This nonlinear bend as influence approaches zero is stronger for some of the network measures, specifically eigenvector, in-degree, and out-degree centrality. Nonetheless, the steepness of the lines is even greater for network measures near the top of the graph, indicating a synergistic relationship between peer influence and selecting youth in central network positions.

<<<Figure 5 about here>>>

In the top right panel of Figure 5 we show the results for each intervention selection strategy compared to the no intervention condition based on the difference in the number of nonsmokers at t_3, thus representing the Average Treatment Effects (ATE) when the influence parameter was set to the value observed in our sample. We also include 95% posterior intervals in the whiskers of this plot. The bottom left panel displays the average treatment effects compared to the strategy of selecting the intervention participants *randomly*, showing that all of these strategies significantly differ from random selection.

The bottom right panel combines information about each intervention selection strategy and the strength of the peer influence parameter. This panel shows the ATEs for each intervention strategy based on a one-unit increase in the influence parameter (e.g., going from the level of influence in our sample to double that value). This plot highlights that the most beneficial strategies (in-degree and eigenvector centrality) are even more effective as the level of peer influence becomes stronger. For example, whereas a one-unit increase (100%) in the influence parameter results in an ATE of 85 students for the random intervention condition, it results in an ATE of almost 120 for the indegree and eigenvector centrality intervention strategies. These findings suggest that such strategies would be even more successful in a network with very strong peer influence effects.

DISCUSSION

Our findings highlight the importance of selecting youth who occupy certain network positions, smoke at different levels, and the importance of leveraging interdependence in peer influence and network position for decreasing smoking. We found that peer influence is sensitive to the network selection strategies under study. Overall, each of our selection strategies resulted in more non-smokers compared to having no or random intervention across both schools, with eigenvector, out-degree, and in-degree centrality generally performing the best. Nonetheless, the differences between some of the network strategies were relatively narrow in some cases, suggesting that several of these network intervention strategies are more effective than a simple random selection strategy. We observed a beneficial network multiplier effect in the higher smoking prevalence school, wherein selecting based on in-degree or eigenvector centrality resulted in *more* non-smokers at the third time point than were intervened upon, indicating a negative and nonlinear effect on the number of smokers. The effects of peer influence alone and jointly with selecting youth in specific network positions and structures both led to decreases in smoking behavior.

In the higher smoking prevalence school, selecting on all of the centrality measures reduced the number of smokers, with eigenvector or in-degree centrality being the most effective strategies. Youth with a high in-degree centrality are popular and may act as role models for smoking behavior; thus, selecting popular youth likely facilitates diffusion of smoking related peer influences throughout peer networks. Moreover, youth with high levels of eigenvector centrality are also able to diffuse peer influences broadly throughout the network. This pattern persisted regardless of the percent selected for intervention, suggesting the potential for scalability of this intervention approach.

We observed a beneficial network multiplier effect in the higher smoking prevalence school, as the long-term to initial effect ratio (LTIER) indicated that selecting based on in-degree or eigenvector centrality yields *more* non-smokers at the third time point than were intervened upon. On the one hand, when selecting upon these centrality measures, these youth were more likely to remain non-smokers throughout the study. On the other hand, the fact that there were more non-smokers at t_3 than initially intervened upon may suggest a protective effect of the diffusion of peer influences via popular youth to their peers. Similarly, other research found that simulating increases in peer influence in school-based peer networks resulted in decreases in school smoking prevalence.[15]

Regarding our simulations based on the lower smoking prevalence school (Sunshine High), the effectiveness of network-based selection was much weaker. Interestingly, in this school, we did not observe a beneficial network multiplier effect as there were fewer non-smokers at the third time point than were intervened upon, even with the most effective strategies. This pattern may be due to the relatively fewer youth smoking at any level in this school. Perhaps the predominant school smoking norms do not favor smoking in this school, and as such, youth who smoke may have less of an effect on changing the smoking prevalence.

Selecting light or moderate smokers rather than heavy smokers for the intervention in Jefferson High (high prevalence) resulted in weaker outcomes compared to selecting heavy smokers. Given that Jefferson High has more youth smoking at higher levels relative to Sunshine High, selecting light smokers may lead to less diffusion of peer influence effects if popular students are heavier smokers. Perhaps light smokers were not perceived as popular and, as such, did not act as role models for other youth to emulate, or they were not friends with heavier smokers and therefore not in a position to influence them.

In our final intervention strategy, we selected youth with the highest values on the specific network measures among those who smoke at *any level.* This strategy yielded results slightly more efficacious in the higher smoking prevalence school. These findings indicate that selecting youth high on in-degree or eigenvector centrality, regardless of their smoking levels, is the most efficacious strategy for reducing the number of smokers in these schools. These findings may also suggest the applicability of this approach to youth who smoke at various levels.

Regarding our findings relating to peer influence, as the strength of peer influence increases, selecting youth occupying specific network positions becomes more effective in increasing the average number of non-smokers at the last time point. In all cases, the number of non-smokers at the end of the simulation is higher for larger magnitudes of peer influence, demonstrating that influence has a protective effect for smoking. The relative ranking of the different strategies remained unchanged, although the number of non-smokers at the end of the simulation is monotonically greater with an increasingly larger influence parameter. There is a generally linear effect of a stronger influence parameter on the number of non-smokers as long as the influence parameter is at least 50% of that observed in the study. However, there is some evidence of a nonlinear bend as influence approaches zero, particularly for eigenvector, in-degree and out-degree centrality. These are notable findings, as the SAOM is highly nonlinear, and yet we observe a mostly linear relationship between increasing peer influence and the corresponding increase in non-smokers.

Overall, our findings regarding peer influence indicate the sensitivity of peer influence to the various network selection strategies and indicate interdependence in network position and peer influence, as they impact adolescent smoking. Most notably, eigenvector and in-degree are the most sensitive to peer influence, which may be useful for informing interventions. Moreover, which network selection strategy to utilize in an intervention also depends on the strength of the peer influence in a network, as we see that at zero levels of peer influence, these network selection strategies are equally inconsequential.

Implications. Our findings highlight the merit of selecting youth in theoretically important network positions for adolescent smoking and the peer influence they transmit. Our study underscores the importance of selecting youth based on their position in the social network, most notably selecting those with high levels of eigenvector, in-degree, or out-degree centrality for decreasing adolescent smoking. Moreover, we found that peer influence is sensitive to network position, especially eigenvector and in-degree centrality. These findings indicate the importance of using these selection strategies in networks with high levels of peer influence to maximize the effectiveness of the network selection strategy.

Another important implication of our findings is that selecting only a relatively small percentage of students in a school can lead to relatively large gains in decreasing smoking prevalence. Lastly, our findings suggest the importance of understanding the nature of the peer influence and pervasive school-based norms regarding smoking prior to intervention, which can help determine whether or not popular youth are likely to smoke, and thus whether or not to select them for intervention.

Limitations. This study has several limitations. Most notably, we focused only on two schools, resulting in limited heterogeneity in terms of school type and levels of substance use. Another limitation is that we only have three time points and thus do not know what happened after the third time point to the smoking status of these students. Moreover, the individual networks were limited to 10 ties, so it is not clear how our results would have changed if we had individual networks that were not truncated.

The Add Health data were collected in the mid 1990's and since then there have been changes in rates of adolescent smoking, adolescents have turned to vaping, and smoking related norms have shifted. However, these network data still remain among the richest adolescent sociometric datasets collected in the US, to date. Most importantly, in-person school-based networks of adolescent youth, even prior to the aforementioned shifts, continue to be potent socialization and developmental contexts for understanding adolescent smoking, smoking related peer influences and the evolution of peer networks, given the salience of in-person interactions for the processes of observational learning,[5] attitude formation,[45] the formation and adoption of normative beliefs[45] and the formation of expectancies, including outcome and generalized expectancies, for behavior in a social setting.[5]

We also acknowledge that we built into our simulations the assumption that adolescents could quit smoking. Extant studies have examined interventions for smoking cessation and quit attempts among adolescents, indicating quit rates at or above 20% in US high school samples, including school-based interventions.[46] One study indicated that among the treatment group, 52% of students were smoke-free, versus 20% in the control group.[47] Another study indicated that at 5.2 months post-intervention, 21.7% of the treatment and 12.6% of the control group reported having quit smoking.[48,49] Moreover, among high risk youth in continuation (i.e., alternative) high schools with eighty five percent daily smokers in the sample, their post intervention quit rate was 17%.[50] Our study is based on youth drawn from a nationally representative sample, so the assumption that 20% of heavy smokers could be intervened upon to quit smoking is plausible. Moreover, other published studies[14] have examined how social network–based interventions affect population-level smoking prevalence, initiations, and cessations. Future studies could use the same simulation approach as we have herein to examine whether youth first initiate a quit attempt or multiple attempts towards eventual cessation, to account for the likelihood of relapse.

We also acknowledge that the smoking intervention we conduct in this simulation study is based on quit attempts, while previous intervention research has focused more on primary prevention interventions. That said, there is an existing literature on quit attempts and smoking cessation among adolescents which calls for more research in this area.[46] Moreover, there is value in the tertiary prevention of adolescent smoking, which is particularly important in high risk high school settings, including continuation (alternative) high schools with high prevalences of smoking and use of other substances.

Another limitation of this study is that our modeling approach contains some assumptions that may pose challenges for implementation in real high school settings. One potential challenge is that those who are delivering the intervention may be able to ascertain the smoking status of adolescents. Consequently, measures should be taken to secure the confidentiality of students' smoking status and ensure students have consented to having their smoking status revealed. Another potential challenge is in the implementation of the intervention strategies, which may require a network analyst team to create the network selection properties, which could raise costs. Our findings, however, indicate that in-degree centrality would be a reasonable strategy, thus lessening the need for more nuanced approaches.

Conclusion. Overall, our findings indicate that selecting youth high based on various social network positions, particularly in-degree, out-degree, and eigenvector centrality, decreased smoking prevalence, one year post intervention. Our findings also indicate the importance, protective effect, and nuanced interdependence of peer influence and network position for decreasing adolescent smoking, and that interventions could be even more efficacious when combining the selection of youth occupying certain network positions with increases in peer influence. Overall, our findings highlight the importance of understanding, before intervening, which network positions the most influential students occupy with regards to the transmission of smoking related peer influence and the nature of the predominant smoking-related peer influences in a school, to most effectively amplify the diffusion of anti-smoking peer influences throughout school-based networks.

REFERENCES

1. Valente TW, Ritt-Olson A, Stacy A, Unger JB, Okamoto J, Sussman S. Peer acceleration: effects of a social network tailored substance abuse prevention program among high-risk adolescents. *Addiction.* 2007;102(11):1804-1815.
2. Valente TW. Network Interventions. *Science.* 2012;6(337):49-53
3. Wang C, Hipp JR, Butts CT, Lakon CM. Insight into Selecting Adolescents for Drinking Intervention Programs: a Simulation Based on Stochastic Actor–Oriented Models. *Prevention Science.* 2022;23(1):48-58.
4. Valente TW, Hoffman BR, Ritt-Olson A, Lichtman K, Johnson CA. Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in school. *American Journal of Public Health.* 2003;93(11):1837-1843.
5. Bandura A. *Social learning theory.* Englewood Cliffs, NJ: Prentice-Hall; 1977.
6. Granic I, Dishion TJ. Deviant talk in adolescent friendships: A step toward measuring a pathogenic attractor process. *Social development.* 2003;12(3):314-334.
7. Bruinsma GJ. Differential association theory reconsidered: An extension and its empirical test. *Journal of Quantitative Criminology.* 1992;8:29-49.
8. Sutherland EH, Cressey DR, Luckenbill DF. *Principles of criminology.* Altamira Press; 1992.
9. Lakon CM, Wang C, Butts CT, Jose R, Timberlake DS, Hipp JR. A Dynamic Model of Adolescent Friendship Networks, Parental Influences, and Smoking. *Jounal of Youth and Adolescence.* 2015;44(9):1767-1786.
10. Lakon CM, Hipp JR, Timberlake DS. The social context of adolescent smoking: a systems perspective. *American Journal of Public Health.* 2010;100(7):1218-1228.
11. Mercken L, Steglich C, Sinclair P, Holliday J, Moore L. A Longitudinal Social Network Analysis of Peer Influence, Peer Selection, and Smoking Behavior Among Adolescents in British Schools. *Health Psychology* 2012;31(4):450–459.
12. Mercken L, Candel M, Willems P, De Vries H. Disentangling social selection and social influence effects on adolescent smoking: the importance of reciprocity in friendships. *Addiction.* 2007;102(9):1483-1492.
13. Ragan DT. Peer beliefs and smoking in adolescence: a longitudinal social network analysis. *The American journal of drug and alcohol abuse.* 2016;42(2):222-230.
14. Schaefer DR, Adams J, Haas SA. Social networks and smoking: exploring the effects of peer influence and smoker popularity through simulations. *Health Education & Behavior*. 2013;40(1_suppl):24S-32S.
15. Lakon CM, Hipp JR, Wang C, Butts CT, Jose R. Simulating dynamic network models and adolescent smoking: the impact of varying peer influence and peer selection. *American journal of public health.* 2015;105(12):2438-2448.
16. Adams J, Schaefer DR. How initial prevalence moderates network-based smoking change: estimating contextual effects with stochastic actor-based models. *Journal of health and social behavior*. 2016;57(1):22-38.
17. McMillan C, Schaefer DR. Comparing targeting strategies for network-based adolescent drinking interventions: A simulation approach. *Social Science & Medicine.* 2021;282:114136.
18. Macy MW, Willer R. From factors to actors: Computational sociology and agent-based modeling. *Annual review of sociology.* 2002;28(1):143-166.

19. Alexander C, Piazza M, Mekos D, Valente T. Peers, schools, and adolescent cigarette smoking. *Journal of Adolescent Health.* 2001;29(1):22-30.
20. Valente TW, Unger JB, Johnson CA. Do popular students smoke? The association between popularity and smoking among middle school students. *Journal of Adolescent Health.* 2005;37(4):323–329.
21. Littlecott HJ, Moore G, McCann M, et al. Exploring the association between school-based peer networks and smoking according to socioeconomic status and tobacco control context: a systematic review. *BMC Public Health.* 2022;22(1):142.
22. Badham J, Kee F, Hunter RF. Effectiveness variation in simulated school-based network interventions. *Applied Network Science.* 2019;4:1-14.
23. Badham J, Kee F, Hunter RF. Simulating network intervention strategies: Implications for adoption of behaviour. *Network Science.* 2018;6(2):265-280.
24. Wasserman S, Faust K. *Social network analysis: methods and applications.* New York: Cambridge University Press; 1994.
25. Friedkin NE. Theoretical foundations for centrality measures. *American Journal of Sociology.* 1991;96(6):1478-1504.
26. Valente TW, Fujimoto K. Bridging: locating critical connectors in a network. *Social Networks.* 2010;32 (3):212–220.
27. Granovetter MS. The strength of weak ties. *American Journal of Sociology*. 1973;78(6):1360-1380.
28. Everett MG, Valente TW. Bridging, brokerage and betweenness. *Social networks.* 2016;44:202-208.
29. Burt RS. *Structural holes: The social structure of competition.* Harvard university press; 1992.
30. Krackhardt D, Nohria N, Eccles B. The strength of strong ties. *Networks in the knowledge economy.* 2003;82.
31. Lin J-H, Guo Q, Dong W-Z, Tang L-Y, Liu J-G. Identifying the node spreading influence with largest k-core values. *Physics Letters A.* 2014;378(45):3279-3284.
32. Centola D, Macy M. Complex contagions and the weakness of long ties. *American journal of Sociology.* 2007;113(3):702-734.
33. Burt RS, Knez M. Kinds of third-party effects on trust. *Rationality and society.* 1995;7(3):255-292.
34. Audrain-McGovern J, Rodriguez D, Tercyak KP, Cuevas J, Rodgers K, Patterson F. Identifying and characterizing adolescent smoking trajectories. *Cancer Epidemiology Biomarkers & Prevention.* 2004;13(12):2023-2034.
35. Valente TW. Network interventions. *science.* 2012;337(6090):49-53.
36. Dishion TJ, McCord J, Poulin F. When interventions harm: peer groups and problem behavior. *American Psychologist.* 1999;54(9):755-764.
37. Steglich C, Snijders TA, Pearson M. Dynamic networks and behavior: Separating selection from influence. *Sociological methodology.* 2010;40(1):329-393.
38. Wang C, Hipp JR, Butts CT, Jose R, Lakon CM. Peer influence, peer selection and adolescent alcohol use: a simulation study using a dynamic network model of friendship ties and alcohol use. *Prevention science.* 2017;18(4):382-393.
39. Harris KM, Halpern CT, Whitsel EA, et al. Cohort profile: The national longitudinal study of adolescent to adult health (add health). *International journal of epidemiology*. 2019;48(5):1415-1415k.

40. Bearman PS, Moody J, Stovel K. Chains of affection: the structure of adolescent romantic and sexual networks. *American Journal of Sociology.* 2004;110(1):44-91.
41. Shoham DA, Tong L, Lamberson PJ, et al. An actor-based model of social network influence on adolescent body size, screen time, and playing sports. *PLoS ONE.* 2012;7(6): e39795.
42. Block P. Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks.* 2015;40:163-173.
43. Ripley RM, Snijders TAB, Boda Z, Vörös A, Preciado P. Manual for SIENA version 4.0 (version February 16, 2017). 2017; http://www.stats.ox.ac.uk/siena/.
44. Wang C, Butts CT, Hipp J, Lakon CM. Model adequacy checking/goodness-of-fit testing for behavior in joint dynamic network/behavior models, with an extension to two-mode networks. *Sociological Methods & Research.* 2022;51(4):1886-1919.
45. Azjen I. The theory of planned behavior. *Organizational Behavior and Humnan Decision Processees.* 1991;50:179-211.
46. Garrison MM, Christakis DA, Ebel BE, Wiehe SE, Rivara FP. Smoking cessation interventions for adolescents: a systematic review. *American Journal of Preventive Medicine.* 2003;25(4):363-367.
47. Adelman WP, Duggan AK, MS PH, Joffe A. Effectiveness of a high school smoking cessation program. *Pediatrics.* 2001;107(4):e50-e50.
48. Dino G, Horn K, Goldcamp J, Fernandes A, Kalsekar I, Massey C. A 2-year efficacy study of Not On Tobacco in Florida: An overview of program successes in changing teen smoking behavior. *Preventive Medicine.* 2001;33(6):600-605.
49. Dino GA, Horn KA, Goldcamp J, Maniar SD, Fernandes A, Massey CJ. Statewide demonstration of not on tobacco: a gender-sensitive teen smoking cessation program. *The Journal of School Nursing.* 2001;17(2):90-97.
50. Sussman S, Dent CW, Lichtman KL. Project EX: outcomes of a teen smoking cessation program. *Addictive Behaviors.* 2001;26(3):425-438.

Table 1. Network properties for selecting adolescents

Network property	Operational Definition
Out-degree	Number of friend(s) an adolescent nominated.
In-degree	Number of other adolescents who nominated an adolescent as a friend.
Closeness centrality	The inverse of the sum of the shortest distances from a focal adolescent to all the other adolescents.
Betweenness centrality	The sum of proportions that a focal adolescent is on the short path(s) between any pair of other adolescents over all shortest path(s) between this pair of adolescents.
Eigenvector centrality	Adolescents with higher eigenvector centrality values have friends also with higher eigenvector centrality values.
Local clustering coefficient	The proportion of existing friendship tie(s) out of all possible friendship ties among an adolescent's friends.
k-core	An adolescent has *k* friends, who also have at least *k* friends amongst those *k* persons.
Valente-Fujimoto version (2010) bridge measure	The ratio of sum of changes in the peer network cohesion after removing an adolescent's each friendship tie to the number of the adolescent's total friendship ties. Adolescents with higher bridge values are more likely to disconnect the peer network if they are removed.
Everett-Valente version (2016) bridge measure	The ratio of the sum of an adolescent's doubled betweenness centrality value and the number of other adolescents in the peer network to the number of an adolescent's total friendship ties. Adolescents with higher bridge values are more likely to disconnect the peer network if they are removed.

Note: All network variables (except out-degree and in-degree) were created using symmetrized ties, to account for incoming and outgoing ties.

Table 2. Smoking behavior and friendship network descriptive statistics

	Higher smoking prevalence school (Jefferson High) (*n* = 976)			Lower smoking prevalence school (Sunshine High)_ (*n* = 2,178)		
Smoking (past 30 days, %)	t_1	t_2	t_3	t_1	t_2	t_3
0 = never	42.0	53.2	45.4	68.6	78.3	71.8
1 = 1-3 days	21.3	9.1	11.7	17.5	7.4	9.4
2 = 4-21 days	9.0	11.6	10.6	4.9	7.1	8.9
3 = 22 or more days	27.7	26.1	32.4	8.9	7.2	10.0
Network statistics						
Out-going ties	6,063	3,713	2,484	5,685	4,201	2,296
Reciprocity index	0.34	0.35	0.35	0.30	0.28	0.25
Transitivity index	0.18	0.19	0.20	0.19	0.19	0.14
Jaccard index	0.22	0.21		0.16	0.16	
Limited nominations (%)	0	4.82	0.41	0	3.49	1.52
Out-degree, mean (SD)	6.21(3.36)			2.61(3.08)		
In-degree, mean (SD)	6.21(4.79)			2.61(2.55)		
Closeness centrality, mean (SD)	0.29(0.03)			0.00(0.00)		
Betweenness centrality, mean (SD)	2237.96(2296.38)			6537.50(9807.08)		
Eigenvector centrality, mean (SD)	0.02(0.02)			0.01(0.02)		
Local clustering coefficient, mean (SD)	0.22(0.16)			0.23(0.25)		
k-core value, mean (SD)	11.57(4.07)			5.17(3.21)		
Valente-Fujimoto version bridge measure (X 1000), mean (SD)	0.016 (0.039)			0.023 (0.032)		
Everett-Valente version bridge measure, mean (SD)	289.21(151.27)			1267.40(1128.26)		

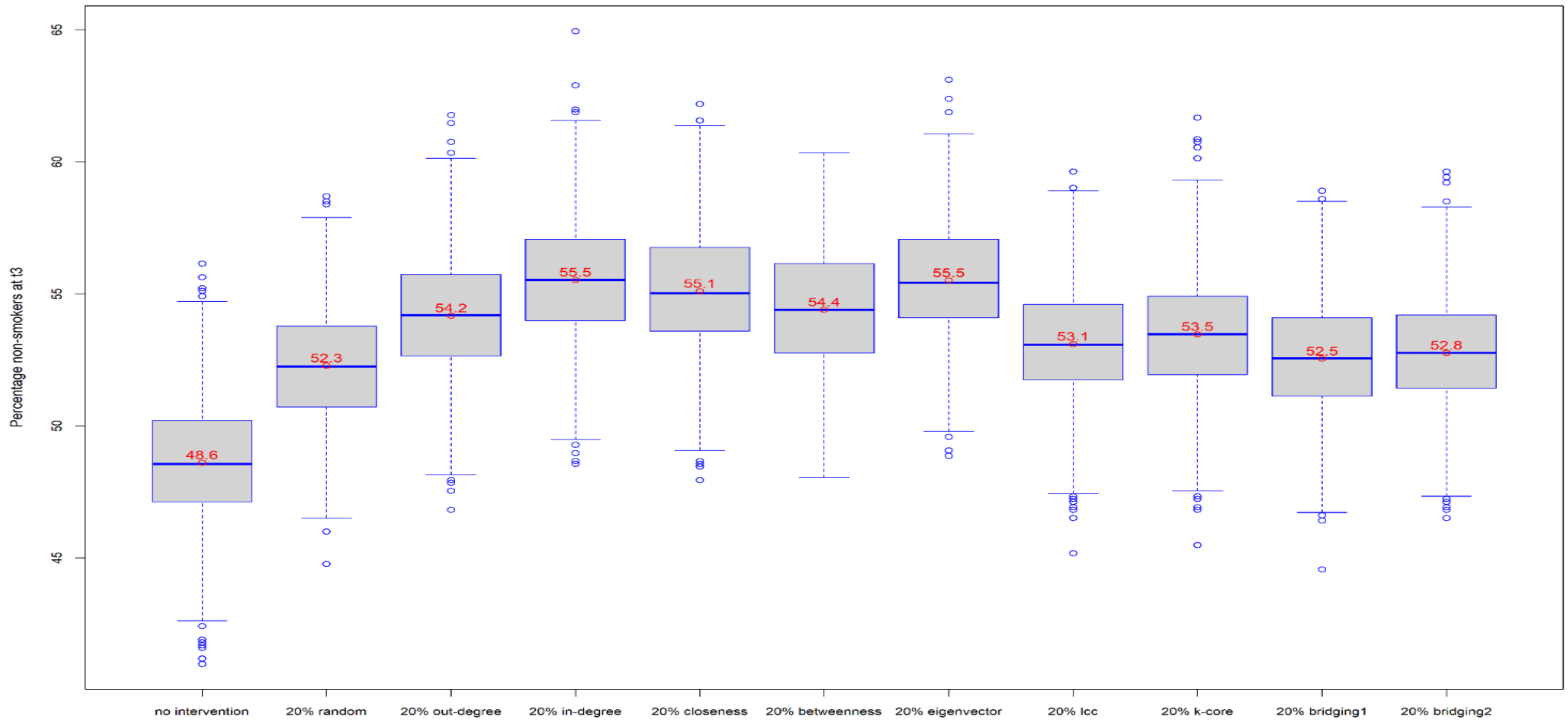
Table 3a. Correlation coefficients among network properties and smoking level at t_1 in higher smoking prevalence school **(**Jefferson High) (n = 976)

	1	2	3	4	5	6	7	8	9	10
1. Out-degree	1.000									
2. In-degree	0.347	1.000								
3. Closeness centrality	0.722	0.683	1.000							
4. Betweenness centrality	0.479	0.691	0.746	1.000						
5. Eigenvector centrality	0.591	0.745	0.787	0.585	1.000					
6. Local clustering coefficient	-0.292	-0.216	-0.440	-0.406	-0.233	1.000				
7. *k*-core	0.724	0.507	0.820	0.417	0.582	-0.318	1.000			
8. Valente-Fujimoto version (2010) bridge measure	-0.183	-0.131	-0.208	0.060	-0.157	-0.155	-0.352	1.000		
9. Everett-Valente version (2016) bridge measure	-0.213	-0.087	-0.114	0.388	-0.192	-0.246	-0.455	0.748	1.000	
10. Smoking level	-0.001	0.014	0.048	0.048	0.040	-0.059	0.013	-0.031	0.007	1.000

Table 3b. Correlation coefficients among network properties and smoking level at t_1 in lower smoking prevalence school **(**Sunshine High) (n = 2,178)

	1	2	3	4	5	6	7	8	9	10
1. Out-degree	1.000									
2. In-degree	0.111	1.000								
3. Closeness centrality	0.157	0.147	1.000							
4. Betweenness centrality	0.567	0.452	0.208	1.000						
5. Eigenvector centrality	0.263	0.346	0.108	0.151	1.000					
6. Local clustering coefficient	-0.157	0.002	0.042	-0.315	0.048	1.000				
7. *k*-core	0.543	0.570	0.162	0.398	0.474	0.018	1.000			
8. Valente-Fujimoto version (2010) bridge measure	-0.024	-0.145	-0.091	0.193	-0.077	-0.289	-0.333	1.000		
9. Everett-Valente version (2016) bridge measure	0.193	0.070	0.084	0.738	-0.067	-0.460	-0.049	0.587	1.000	
10. Smoking level	-0.091	-0.015	0.012	-0.023	-0.016	0.028	-0.072	0.020	-0.011	1.000

Figure 1. Simulation results showing percent non-smokers at t_3 in higher smoking prevalence school **(**Jefferson High) when selecting 54 (or 20%) heavy smokers in Jefferson High for intervention based on network properties at t_1 and including no intervention and random conditions (n=976)



Note: 20% random – Randomly selecting 20% heavy smokers; 20% out-degree – Selecting 20% heavy smokers with highest out-degree values; 20% in-degree – Selecting 20% heavy smokers with highest in-degree values; 20% closeness – Selecting 20% heavy smokers with highest closeness centrality values; 20% betweenness – Selecting 20% heavy smokers with highest betweenness centrality values; 20% eigenvector – Selecting 20% heavy smokers with highest eigenvector centrality values; 20% lcc – Selecting 20% heavy smokers with highest local clustering coefficient values; 20% k-core – Selecting 20% heavy smokers with highest *k*-core values; 20% bridging1– Selecting 20% heavy smokers with highest Valente-Fujimoto version (2010) bridge measure values; 20% bridging2 – Selecting 20% heavy smokers with highest Everett-Valente version (2016) bridge measure values.

Figure 2. Percent non-smokers at time 3, based on varying number of heavy smokers intervened upon. Comparing three selection strategies

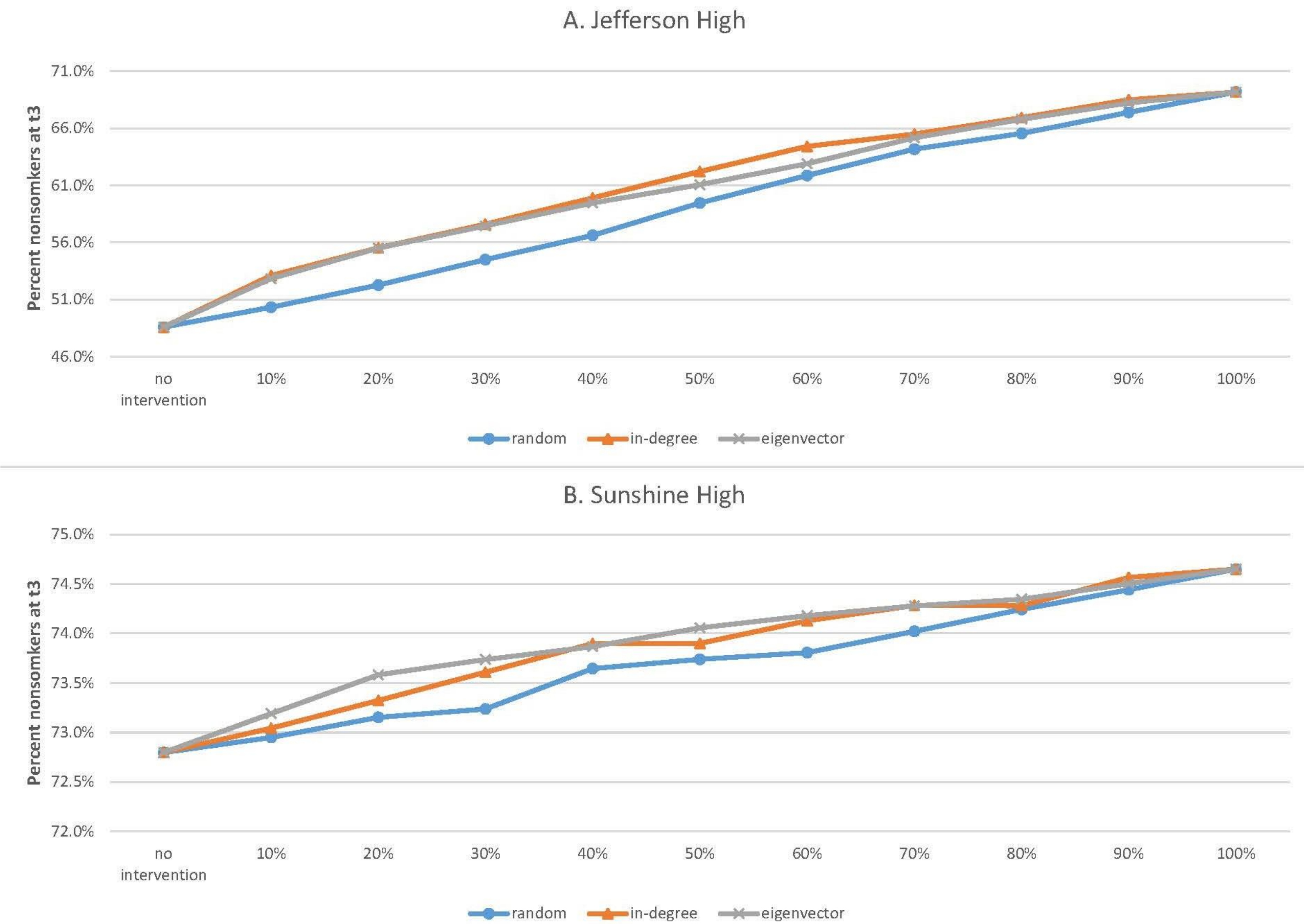
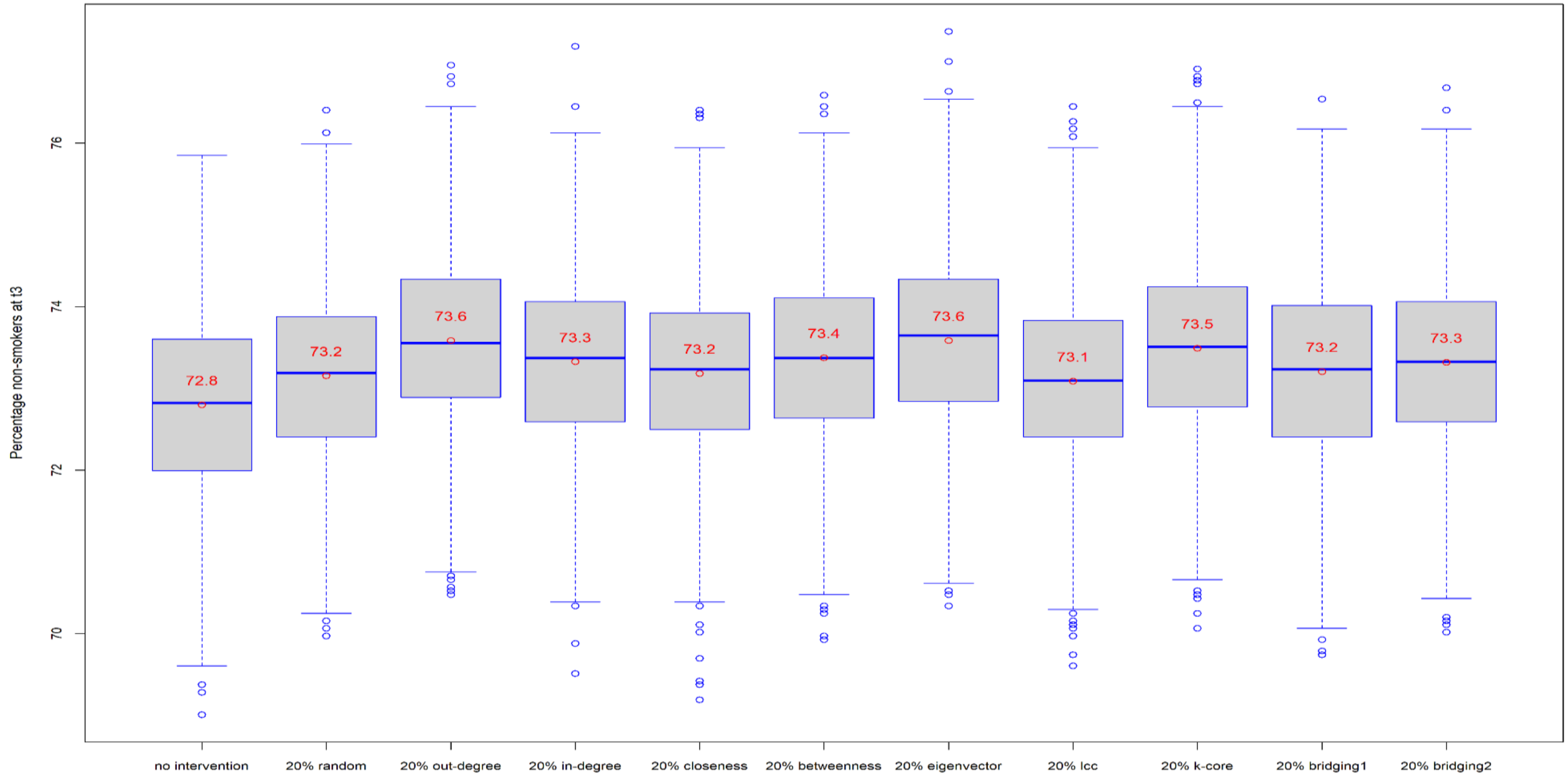


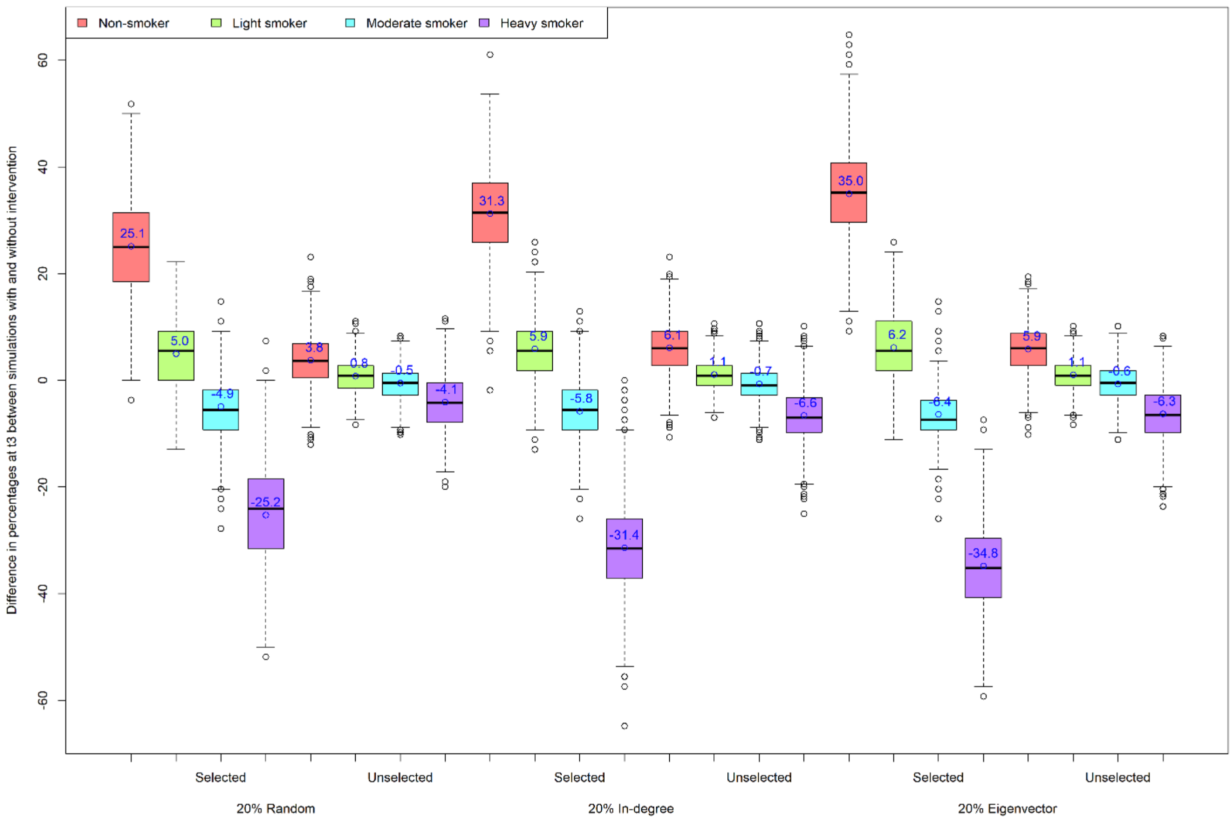
Figure 3. Simulation results showing percent non-smokers at t_3 when selecting 40 (or 20%) of heavy smokers in lower smoking prevalence school (Sunshine High) for intervention based on network properties at t_1 and including no intervention and random conditions (n=2,178)



Note: 20% random – Randomly selecting 20% heavy smokers; 20% out-degree – Selecting 20% heavy smokers with highest out-degree values; 20% in-degree – Selecting 20% heavy smokers with highest in-degree values; 20% closeness – Selecting 20% heavy smokers with highest closeness centrality values; 20% betweenness – Selecting 20% heavy smokers with highest betweenness centrality values; 20% eigenvector – Selecting 20% heavy smokers with highest eigenvector centrality values; 20% lcc – Selecting 20% heavy smokers with highest local clustering coefficient values; 20% k-core – Selecting 20% heavy smokers with highest *k*-core values; 20% bridging1– Selecting 20% heavy smokers with highest Valente-Fujimoto version (2010) bridge measure values; 20% bridging2 – Selecting 20% heavy smokers with highest Everett-Valente version (2016) bridge measure values.

Figure 4. Plotting percent at each smoking level (none, low, moderate, high) at t_3 compared to non-intervention scenario, for intervened on students (selected) and other heavy smokers (unselected) for three intervention strategies: 1) random; 2) in-degree centrality; 3) eigenvector centrality

A. Higher smoking prevalence school (Jefferson High)



B. Lower smoking prevalence school (Sunshine High)

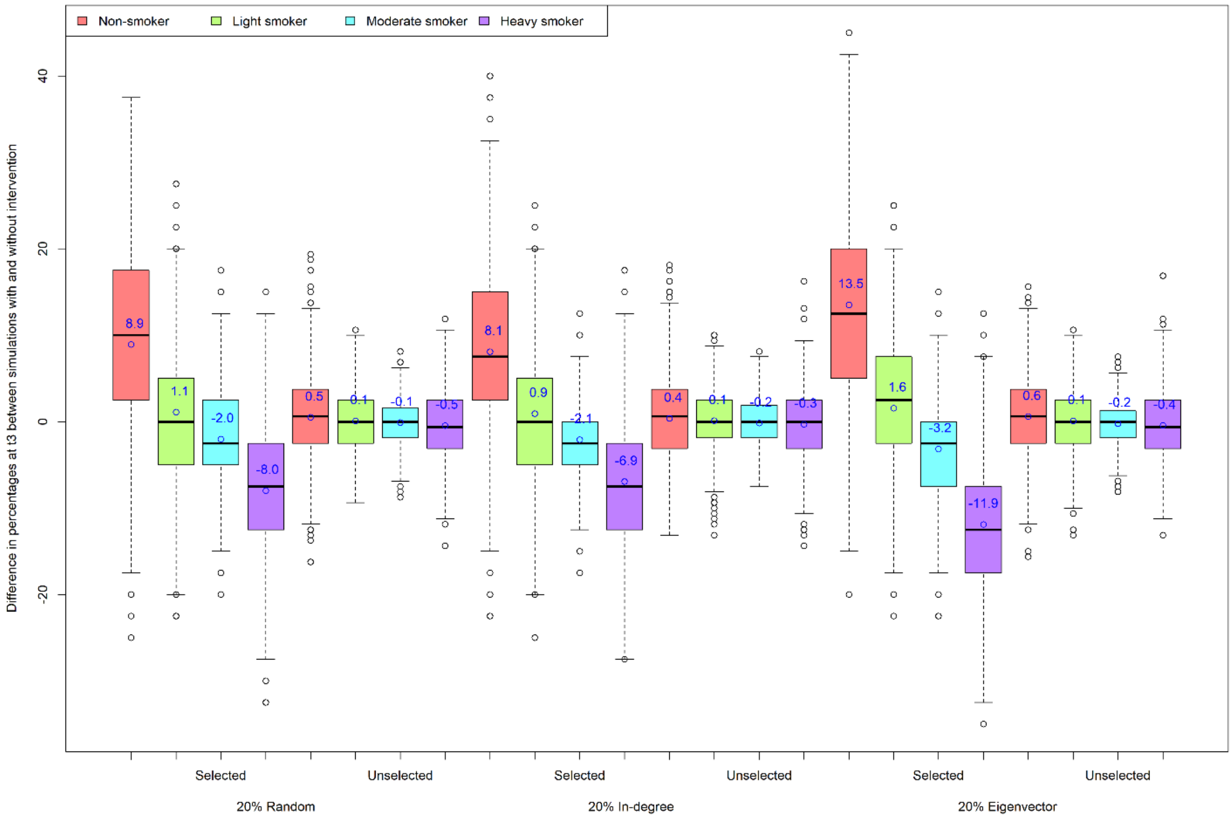


Figure 5. Effectiveness of network-based smoking interventions with varying levels of peer influence in the higher smoking prevalence school (Jefferson High) (n=976)