American Journal of Community Psychology, Vol. 35, Nos. 3/4, June 2005 (© 2005)
DOI: 10.1007/s10464-005-3397-z

Getting the Big Picture in Community Science: Methods That Capture Context

Douglas A. Luke[1]

Community science has a rich tradition of using theories and research designs that are consistent with its core value of contextualism. However, a survey of empirical articles published in the *American Journal of Community Psychology* shows that community scientists utilize a narrow range of statistical tools that are not well suited to assess contextual data. Multilevel modeling, geographic information systems (GIS), social network analysis, and cluster analysis are recommended as useful tools to address contextual questions in community science. An argument for increased *methodological consilience* is presented, where community scientists are encouraged to adopt statistical methodology that is capable of modeling a greater proportion of the data than is typical with traditional methods.

KEY WORDS: community science; context; multilevel modeling; GIS; social network analysis; cluster analysis.

INTRODUCTION

A generation ago, Allen Barton had this to say about social science research:

> For the last thirty years, empirical social research has been dominated by the sample survey. But as usually practiced, using random sampling of individuals, the survey is a sociological meatgrinder, tearing the individual from his social context and guaranteeing that nobody in the study interacts with anyone else in it. It is a little like a biologist putting his experimental animals through a hamburger machine and looking at every hundredth cell through a microscope; anatomy and physiology get lost, structure and function disappear, and one is left with cell biology . . . . If our aim is to understand people's behavior rather than simply to record it, we want to know about primary groups, neighborhoods, organizations, social circles, and communities; about interaction, communication, role expectations, and social control. (Barton, 1968 as reported in Freeman, 2004)

Although this statement came a few years after the Swampscott Conference, it is almost as if Barton were talking to the group of community scientists who were in Massachusetts inventing a new field. As Kelly notes, "The conference was an occasion to acclaim that beyond conventional methods and, *with a focus beyond the individual*, there were valid activities and meaningful roles for a new kind of psychologist, the community psychologist (Kelly, 2002, p. 44, emphasis added).

Thus, community scientists have put context front and center as one of the core values of community psychology. Shinn and Rapkin (2000) advised that "...a central tenet of community psychology is that human behavior must be understood in context." So, for example, community scientists study domestic violence using methods and theories that are consistent with the view that domestic violence is not just an individual behavior, but a complex process shaped by historical, social, financial, and legal contexts. The types of groups that community scientists work with (e.g., domestic violence victims, young mothers, gays and lesbians, drug users, inner city residents, etc.) are of interest not because of any defining psychological characteristics, but because these groups have been affected in specific ways by the economic, social, cultural, and physical situations in which they are embedded.

[1]To whom correspondence should be addressed at Saint Louis University School of Public Health, 3545 Lafayette Avenue, Saint Louis, MO 63104; e-mail: dluke@slu.edu.

However, although community scientists frequently employ theories, models, and frameworks that take context into account, they seem to be less likely to employ contextual *methods* in their work. Community psychology does have a long tradition of analytic innovation (Revenson et al., 2002), but the majority of empirical work in community science utilizes a remarkably narrow array of analytic approaches.

In an earlier set of articles examining science and community psychology, Kelly's (2003) first suggestion for vitalizing scientific community psychology was to "demythologize statistics." He argued that not only do we need to move beyond traditional methods such as ANOVA, regression, and factor analysis; but that community scientists need to be in control of the quantitative methods, and not vice versa. This paper is one small response to his call.

The goal of this paper is to provide examples of a number of useful analytic methods that can be used to capture community context. It is my hope that if community scientists have a wider variety of analytic tools in their toolbox, they will be more likely to get away from Barton's sample survey approach to doing social science. To support this argument, the next section presents a review of statistical practices during six years of articles published in the *American Journal of Community Psychology*. The bulk of the paper describes four statistical methods appropriate for assessing context, and provides examples drawn from community science. The paper concludes with a discussion of the need for the development of *consilient* statistical methods in community science.

STATISTICAL PRACTICE IN COMMUNITY SCIENCE

To provide context for the following discussion, I conducted a review of the types of statistical methods used in the *American Journal of Community Psychology* (AJCP) during two time periods: from 1981 to 1983, and 20 years later from 2001 to 2003. By reviewing six years of AJCP articles, we can get a good sense of the types of statistical practices used by community scientists. Also, by examining two different time periods we can see how statistical practice has changed over a 20 year period. For the purposes of this paper, this review can help ascertain the extent to which contextual methods have been used in the past, or are being used currently by AJCP authors. Although AJCP is not the only place that community-oriented empirical work appears, it is one of the more important settings for work that aims to advance community science.

A total of 304 articles were reviewed; 144 in the early period (1981–1983) and 160 in the current period (2001–2003). Table I shows how the types of articles published in AJCP have changed during this time. In the early period, the majority (88%) of published articles were traditional empirical studies where original data were presented. The current AJCP publishes a wider variety of articles. Almost half of all articles (48%) came from themed special issues. More room was also made for Presidential and award addresses. Finally, on a percentage basis four times as many articles with qualitative content are being currently published compared to 20 years earlier (17% vs. 4%).

Figure 1 shows how often particular types of statistical analyses were used in the 215 empirical papers published in the early 1980s and from 2001 to 2003. (These percentages add up to more than 100%, because most papers used multiple types of statistical analyses.) An empirical article is one that presents original data that have been analyzed either quantitatively or qualitatively. We can see that there have been some changes in statistical practice over the past two decades. Well more than half (59.5%) of all empirical articles presented correlations in the early 80s; currently a little over a third (34.8%) of the empirical articles present correlations. The use of ANOVA methods shows a similar drop, from 59.5 to 37.1%. Use of structural equation modeling (SEM) and associated methods such as path analysis increased dramatically during this same period (from 2.4 to 11.2%). This is not surprising given that sophisticated SEM software did not appear until the 1990s.

However, despite these changes, this figure supports Kelly's contention that as a field we

Table I. AJCP Article Characteristics

Years	Total articles	Empirical	Qualitative	Special issue	Address
1981–1983	144	126 (88%)	6 (4%)	10 (7%)	5 (4%)
2001–2003	160	89 (56%)	27 (17%)	76 (48%)	19 (12%)
Totals	304	215 (71%)	33 (11%)	86 (28%)	24 (8%)

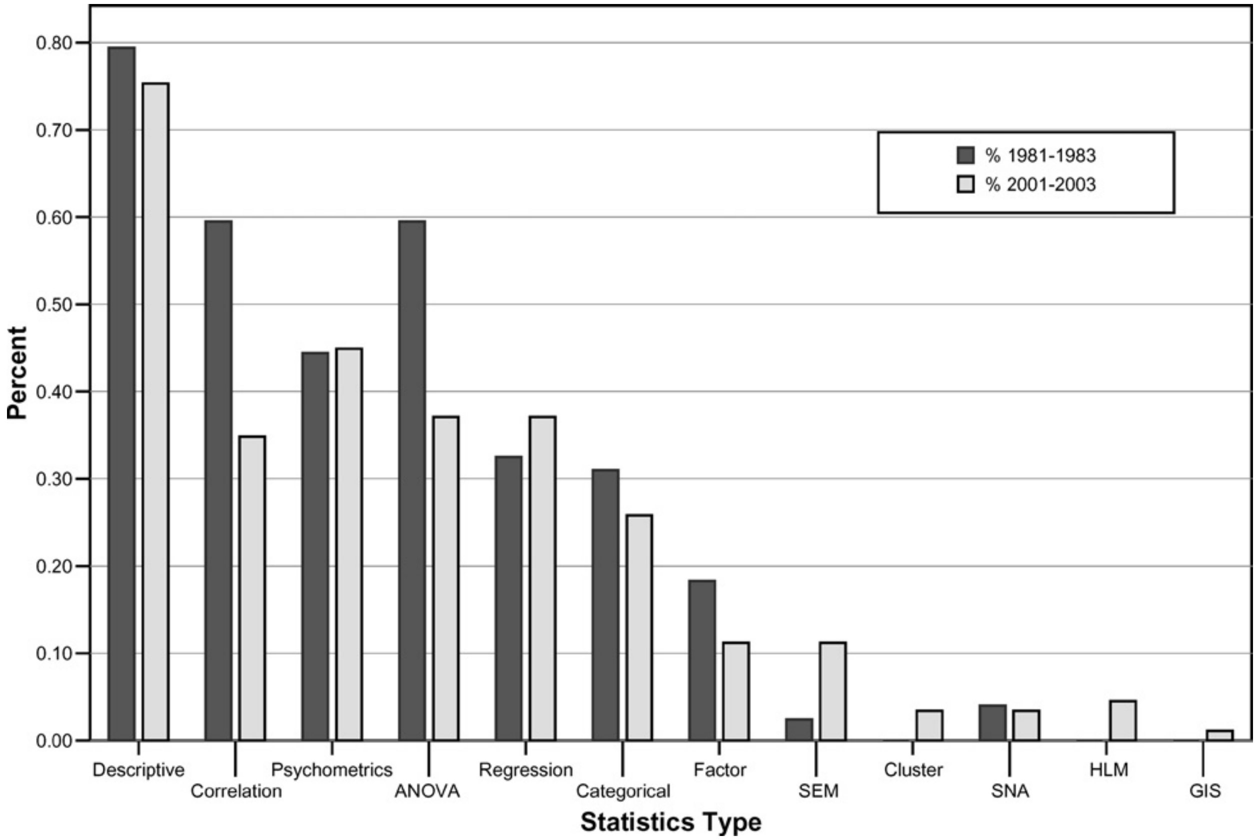


Fig. 1. Statistics usage in the *American Journal of Community Psychology* from 1981 to 1983 ($N = 126$ empirical articles) and from 2001 to 2003 ($N = 89$ empirical articles).

predominately use traditional statistical approaches such as ANOVA, regression, psychometrics, correlations, and categorical statistics (e.g., chi-square). More specialized techniques such as cluster analysis, social network analysis (SNA), multilevel modeling (HLM), meta-analysis, non-linear modeling, and geographic information systems (GIS) are used rarely. Of the four techniques that are the topic of this paper (i.e., cluster analysis, social network analysis, multilevel modeling, and geographic information systems) only network analysis was used at all in the early 3-year period. Furthermore, none of these contextual techniques appeared in more than four studies from 2001 to 2003. Although one would expect that a very general modeling technique such as ANOVA would be used more often than, say, network analysis, the generality of the technique is not the entire story. For example, structural equation modeling, multilevel modeling, and nonlinear modeling are also very general statistical approaches, but they have not been used widely in the community sciences.

One possible interpretation of the results from this methodologic review is that community science simply does not need to use nontraditional statistical tools to "...make a positive difference in community living and activities..." (Sarason, 2003, p. 209). However, the main point I hope to make in the following sections of this paper is not that we should use different tools because they are newer, more innovative, or more popular. Rather, I believe that the quantitative methods that we historically rely on are not always the most appropriate tools if community scientists are seriously interested in understanding how the physical, social, organizational, cultural, economic, and political context shape human behavior and health. Although this paper will talk much about statistical practice, it is not a statistics paper. At its heart, I believe this argument is more philosophical and political than technical. Put simply, the decisions we make about the tools we use in community science say something about the values we hold as community scientists.

CONTEXTUALISM AND COMMUNITY SCIENCE

> I should venture to assert that the most pervasive fallacy of philosophic thinking goes back to neglect of context.
>
> – John Dewey

As part of its mission statement, the *Society for Community Research and Action* (SCRA) recognizes that "Human competencies and problems are best understood by viewing people within their social, cultural, economic, geographic, and historical contexts" (SCRA, 2004). This core value of contextualism can be seen in many ways, including the close ties of community psychology to ecological psychology, the use of explicitly contextual theories such as Behavior Setting Theory (Wicker, 1992), and the recognition that effective interventions based on community science can and should be aimed at the extra-individual level (Tseng et al., 2002). That is, changing families, neighborhoods, schools, churches, and organizations can be a more effective way to enhance health than simply intervening at the person level.

The long history of contextualism in community science puts us at the leading edge of a wave of change in how human health and behavior should be studied. One very prominent example is the 2001 report issued by the National Institutes of Health (NIH) entitled *Toward higher levels of analysis: Progress and promise in research on social and cultural dimensions of health* (Office of Behavioral and Social Sciences Research, 2001). This report presented a new agenda for NIH research focusing on two goals: (1) expanding health-related social sciences research at NIH and (2) integrating social science research into interdisciplinary contextual and multilevel studies of health. Particularly relevant for this discussion is the recommendation to:

> . . . support the development of state-of-the-art social science methods. Challenges include measurement at the group, network, neighborhood, and community levels; the further development of methods for longitudinal research; multi-level research designs that integrate diverse qualitative and quantitative approaches . . . ; and the development of improved methods for data collection and analysis (p. 3).

The empirical work published in AJCP shows the importance and prevalence of contextual approaches. Notable examples in the past few years include interventions to enhance the quality of life of battered women through a community-based advocacy approach (Bybee & Sullivan, 2002); a description of the relationship of ethnicity to family networks (Hirsch, Mickus, & Boerger, 2002); identification of supportive organizational characteristics for rape victim advocates (Wasco, Campbell, & Clark, 2002), and a study examining the effects of organizational downsizing on the health of employees (Kivimäki, Vahtera, Elovainio, Pentti, & Virtanen, 2003). These (and many other) examples utilize explicitly contextual frameworks. However, closer examination of these studies reveals that although they use a contextual framework, the actual data collection and analyses are restricted to the person-level.

This is a general pattern. One hundred twenty of the empirical studies reviewed above incorporate some type of contextual framework or construct to understand individual-level behavior or characteristics. However, less than one in four (22.5%) of these studies collected data directly from the extra-individual levels or settings, or analyzed these data using appropriate methods for multiple levels. In other words, although community scientists value contextual *thinking*, we are much less likely to actually employ contextual *methods*.

There are a number of possible explanations for this disconnect between our theories and our methods. First, community psychology is still a relatively young field, and many practitioners have received their training from people and places still rooted in the person-centered traditions of clinical and general psychology. Many of the research design and analytic approaches which are appropriate for contextual studies are more commonly used in other disciplines such as organizational behavior, sociology, urban planning, etc.

Another reason for this disconnect that I would like to discuss at more length has to do with the *Rule of the hammer*—if the only tool you know how to use is a hammer, every situation looks like it needs to be hammered. The traditional statistical tools used most often by community scientists (i.e., regression, ANOVA, factor analysis, etc.) have been developed to be used with a single level of analysis. Although it is common to use predictor variables in ANOVA and regression that give contextual information (e.g., group membership, state of residence, classroom), these models actually work by disaggregating group-level information to the individual level so that all predictors in a model are tied to the individual unit of analysis. This leads to at least two problems. First, all of the un-modeled contextual information ends up pooled into the single individual error term of the model (Duncan, Jones, & Moon, 1998). This is problematic because individuals belonging to the same context will presumably have correlated errors, which violates one of the basic assumptions of the general linear model. The second problem is that by ignoring context the model assumes that the regression coefficients apply equally to all contexts, "...thus propagating the notion that processes work out in the same way in different contexts" (Duncan et al., 1998, p. 98).

So, since we mainly know how to use statistics that are limited to individual-level analyses, it is not surprising that we end up not collecting or analyzing true contextual, multilevel data.

STATISTICAL APPROACHES CONSISTENT WITH CONTEXTUALISM

The good news is that there are a number of quantitative tools available that are consistent with the value of contextualism, and that are extremely useful for describing or modeling the influence of ecological, environmental, or group-level factors on individual-level behavior or health. In the next sections of this paper I would like to highlight four such methods: (1) multilevel modeling; (2) geographic information systems (GIS); (3) social network analysis; and (4) cluster analysis.

Multilevel Modeling

Multilevel modeling is a general regression-based statistical tool that can build statistical models of data that are multilevel in nature. Multilevel models are more statistically appropriate for multilevel data than are traditional regression or ANOVA techniques (Snijders & Bosker, 1999). For example, if the data are truly multilevel in nature, then the level-1 units are clustered within the level-2 units, and the level-1 error terms are not likely to be independent, thus violating one of the basic assumptions of the general linear model. More importantly for us, with multilevel modeling we can build statistical models that match our conceptual models. That is, there will no longer be the disconnect between our thinking and our methods.

Multilevel modeling, also known as hierarchical linear modeling, mixed-effects modeling, or growth-curve modeling, has been underutilized in community science compared to developmental psychology, education, sociology, or political science. Hedeker, McMahon, Sason, and Salina (1994), in a study of a worksite smoking cessation program, gave an early demonstration of the utility of these methods; focusing on how the clustering of the person-level data can be appropriately adjusted for in a hierarchical model. Coulton, Korbin and Su (1999), used multilevel methods in an examination of how neighborhood factors influence child maltreatment. They convincingly demonstrated that the parameter estimates for the predictor variables were different than would be expected if only individual or neighborhood data had been used. Specifically, adverse neighborhood conditions weakened the effects of previously established individual risk factors, such as violence in the family of origin. Finally, Mankowski, Humphreys, and Moos (2001) develop a multilevel model that shows that person–environment fit of treatment belief systems is an important predictor of the extent of 12-step group involvement.

The following more detailed example may help demonstrate the utility of multilevel modeling for community science. The data and analyses come from work that we have done in the area of tobacco control policy (Luke & Krauss, 2004). The main goal of this study was to identify the important influences on voting on tobacco-related legislation by members of Congress from 1997 to 2000. The unit of analysis is an individual Congress member.

A typical single-level regression model that could be developed for these data might look like this:

$$\text{VotePct}_i = \beta_0 + \beta_1(\text{Party})_i + \beta_2(\text{Money})_i + r_i$$

where we want to predict the percentage of time that a member of Congress voted in a pro-tobacco industry direction based on his or her political party, and the amount of money received from a tobacco industry political action committee (PAC).

However, members of Congress are not randomly selected, and it is reasonable to expect that the data will exhibit cluster effects. That is, Senators or Representatives within the same state such as Massachusetts are likely to be more similar to each other on any number of important characteristics than Congress members from two different states. Figure 2 illustrates this—we can see that the percentage of time that a Congress member votes for a bill in a pro-tobacco industry direction varies from state to state.

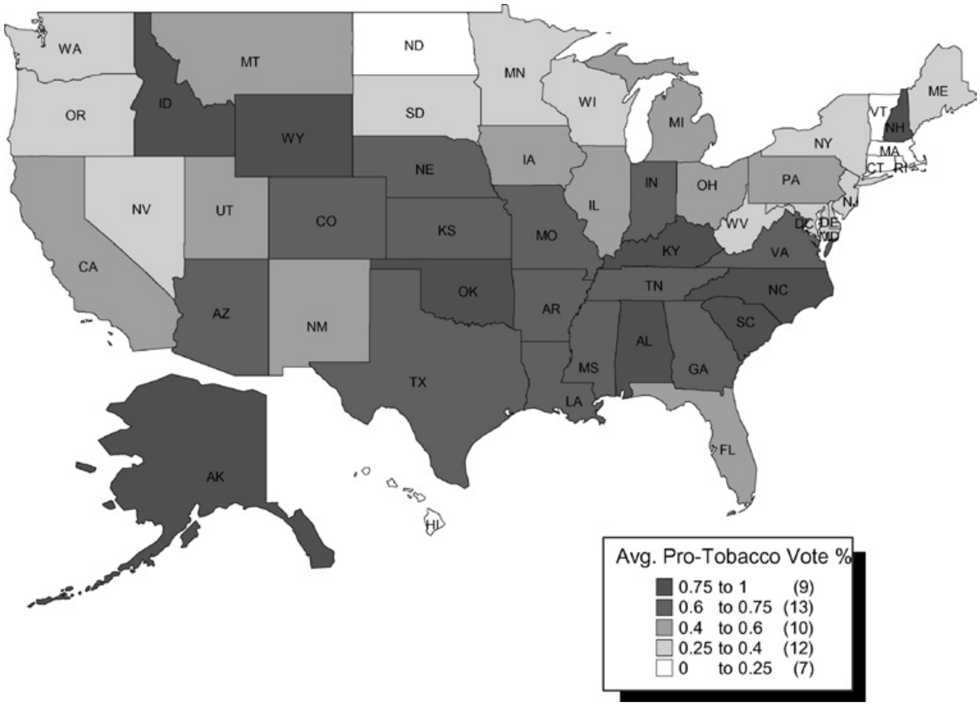


Fig. 2. Average pro-tobacco vote percentage by Congress members – 1997–2000.

The single-level regression model above cannot handle the clustering of the data (and the concomitant non-independence of error terms). More important than this statistical problem is the fact that as community scientists we would like to build a more contextual model. In particular, we would like to account for state-level characteristics that may influence voting behavior apart from the individual party and money received. The size of the tobacco economy in a state, operationalized as size of tobacco harvest, is a type of contextual variable that might influence Congressional voting.

The following set of equations show how a multilevel statistical model is structured:

$$\begin{aligned}
\text{VotePct} &= \beta_{0j} + \beta_{1j}(\text{Party})_{ij} + \beta_{2j}(\text{Money})_{ij} + r_{ij} \\
\beta_{0j} &= \gamma_{00} + \gamma_{01}(\text{Acres})_j + u_{0j} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}(\text{Acres})_j + u_{1j} \\
\beta_{2j} &= \gamma_{20} + \gamma_{21}(\text{Acres})_j + u_{2j}
\end{aligned}$$

Although this statistical model looks quite a bit more complicated than the earlier single-level model, it is mainly doing two important things for us. First, the beta-coefficients at level-1 become dependent variables at level-2—indicating that the level-1 intercepts and slopes may vary from state to state. Second, this type of model clearly shows which predictor variables are acting at level-1 (i.e., political party, PAC money) and which at level-2 (amount of tobacco acreage produced in a state).

Table II presents the results of fitting this multilevel model. We can see that at the Congress member level being a Republican ($\hat{\gamma}_{10} = .5066$) and receiving more PAC money ($\hat{\gamma}_{20} = .0049$) are associated with voting more often in the pro-tobacco industry direction. Specifically, for every \$10,000 received by a member of Congress from a tobacco PAC, there is approximately a 5% greater chance to vote in the pro-tobacco industry direction on a tobacco-related bill. However, we also see that there is a significant state-level contextual effect—members from states with a greater tobacco economy are also more likely to vote pro-tobacco ($\hat{\gamma}_{01} = .0027$). The relatively large and significant random variability component for the intercept term (u_{0j}) suggests, moreover, that there may be other important contextual effects that have not been included in the model. This relatively simple multilevel model shows us that it is important to take both individual and contextual effects into account when modeling tobacco policy behavior in Congress. For more details on how to

Table II. Hierarchical Model Estimates of the Effects of Political Party, Total Contributions, and State-Level Tobacco Acreage on Percentage of Pro-Tobacco Votes

Fixed effects	Coefficient	*SE*	*T*-ratio	*p*
For intercept (β_{0j})				
Intercept (γ_{00})	.1828	.0205	8.90	.000
Acres (γ_{01})	.0027	.0005	5.10	.000
For PARTY slope (β_{1j})				
Party (γ_{10})	.5066	.0215	23.53	.000
Acres (γ_{11})	−.0016	.0004	3.60	.001
For MONEY slope (β_{2j})				
Money (γ_{20})	.0049	.0005	8.80	.000
Acres (γ_{21})	−.00002	.0000	5.50	.000
Random effects	*SD*	Var. Comp.	χ^2	*p*
Intercept (u_{0j})	.0978	.0096	84.1	.000
Party Slope (u_{1j})	.0705	.0050	54.8	.023
Money Slope (u_{2j})	.0009	.0000	29.2	>.50
Level-1 (e_{ij})	.1628	.0265		
Model Fit	Deviance	Parameters	AIC	BIC
	−353.8	13	−327.8	−272.3

Notes: Political party is coded 0 for Democrat and 1 for Republican. Contributions are recorded in thousands of dollars and Acres are recorded in thousands of acres.

fit and interpret multilevel models, see Hox (2002), Luke (2004), or Snijders and Bosker (1994).

Geographic Information Systems

The second useful quantitative tool for exploring context is geographic information systems (GIS). GIS is a set of database, mapping, and statistical tools that allow visual and quantitative assessment of geographic information. (Geographic in the broad sense, meaning any type of information that has a physical location.) Although the production of maps is a common end result of GIS methods, GIS can also be used to construct contextual variables and statistically analyze spatial relationships (Haining, 2003). GIS methods emerged as a computing technology over the last several decades and are used extensively in a wide variety of areas including geology, meteorology, urban planning, public safety, marketing, political science, and public health (Mark, Chrisman, Frank, McHaffie, & Pickles, 2004). There has been a growing movement towards a 'community GIS' where GIS methods are used for community development, mapping of community assets, community health assessments, and eco-development (Carver, 2001). However, community scientists have been relatively slow to take advantage of GIS techniques.

GIS has been used most often by health and social scientists to examine patterns of criminal and risky health behaviors. For example, Wieczorek and Hanson (1997) used GIS methods to examine patterns of driving-while-intoxicated (DWI) and reveal the geographic context of this behavior. The authors use GIS maps including contour plots which show that DWIs are not distributed randomly around the metropolitan area. This type of analysis leads to policies and interventions that can be aimed more precisely, leading to lower costs and hopefully enhanced effectiveness.

A second example comes from the area of tobacco control policy. Figure 3 is a map showing the distribution of tobacco billboards in the St Louis metropolitan area, shortly before tobacco billboards were eliminated as a provision of the 1998 Master Settlement Agreement (Luke, Esmundo, & Bloom, 2000). We used GIS to collect the billboard data and analyze the location patterns of tobacco advertising. In the figure, the billboards are coded by the type of image found on the billboard, and the underlying map shows the population mix by census tract. By combining the billboard and census data, we can see that billboards with African American images on them tended to be concentrated in neighborhoods with higher proportions of African American residents. We used this evidence to support an argument that the tobacco industry was targeting African Americans for their products.

A final example of the utility of GIS is provided by Goldstein et al. (2003) in a case study of

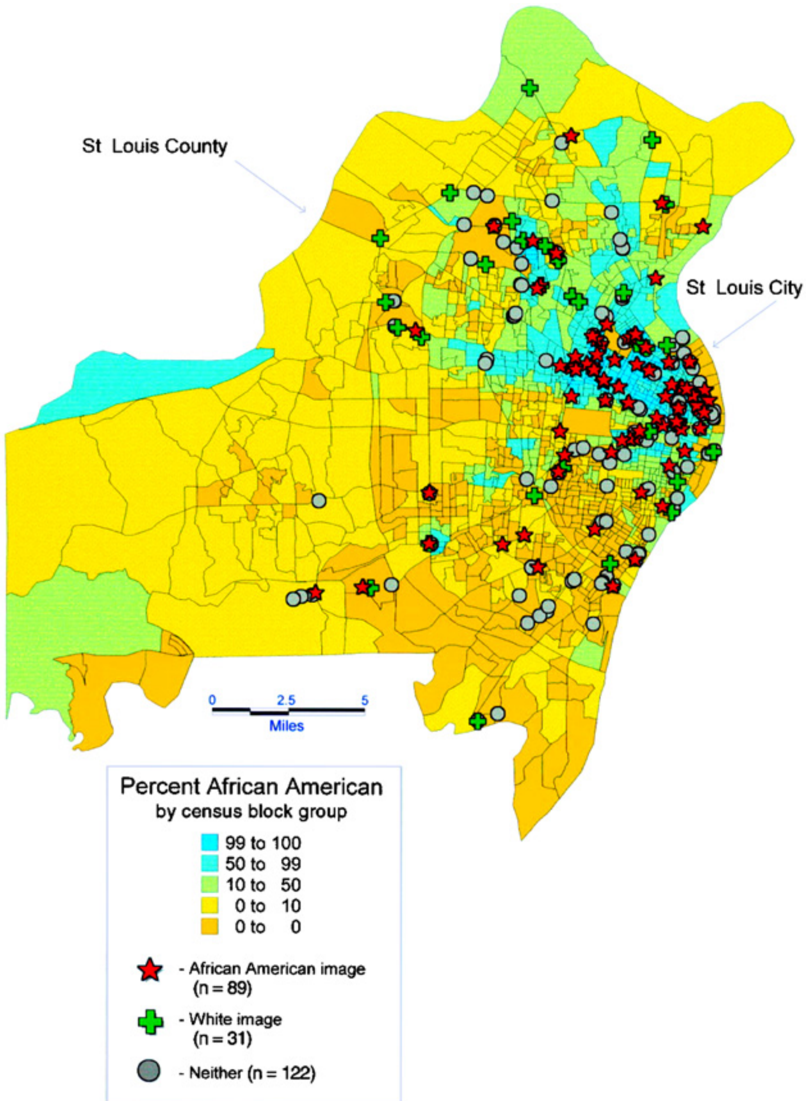


Fig. 3. Targeted placement of tobacco billboards with African American images in 1998. Symbols indicate type of image found on the tobacco billboard. Census block groups are shaded according to percentage of African American residents. Taken from Luke, Esmundo, & Bloom (2000).

the adoption of 100% tobacco-free school policies in North Carolina. In this study the authors used key informant interviews to identify important factors influencing the successful adoption of tobacco-free policies in 14 North Carolina school districts. Although key informants suggested that the local tobacco economy had little direct influence on policy adoption, GIS analyses revealed that all school districts passing policies were located in counties with relatively little tobacco production (Fig. 4). This map not only reveals an important pattern that can be the subject of further research, it allows dissemination of this finding in a way that is clear, powerful, and easy to understand.

Network Analysis

Social network analysis is a broad set of methods for the systematic study of social structure (Degenne & Forsé, 2004). Network tools are based on the analysis of *relational* data—information about the connections among a set of actors, be they

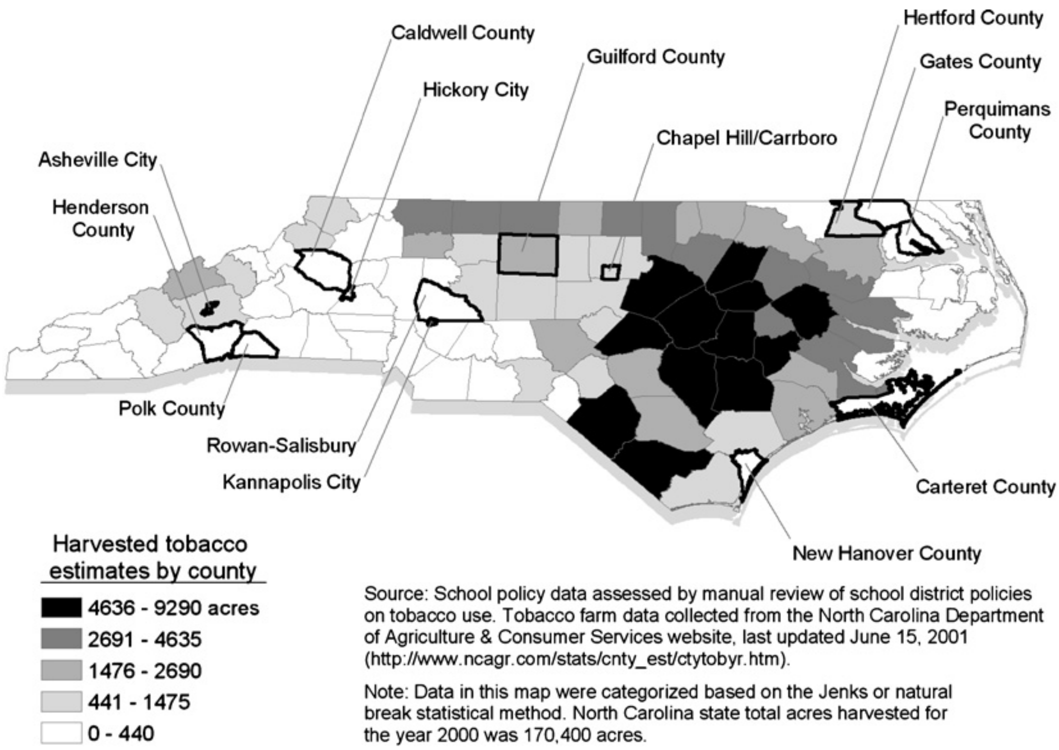


Fig. 4. Location of 14 North Carolina school districts adopting 100% tobacco-free policies. Taken from Goldstein et al. (2003).

persons, agencies, etc. This distinguishes network analysis from much of the rest of community science where *attribute* data are typically collected and analyzed. The relations can be almost any type of information about a connection between actors—friendship, recognition, money exchange, kinship, information exchange, and respect. The relation can be very specifically defined, such as sharing needles among a network of drug users.

Network analysis has been broadly used in the physical, social, and health sciences. Network methods have been used to model disease transmission (Rothenberg et al., 1998), job-seeking (Granovetter, 1973), the diffusion of innovations (Valente, 1995), inter-organizational behavior (Mintz & Schwartz, 1981), social capital and community development (Lin, 2000), and teen smoking (Alexander, Piazza, Mekos, & Valente, 2001), to name just a few.

The use of network analysis in community science grew out of the work on social support. Social support was thought to have a positive effect on a wide variety of important individual characteristics and behavior, including coping, bereavement, suicide, general physical and mental health (Cohen, Underwood, & Gottlieb, 2000). Early on, social support tended to be measured by simply asking individuals to rate how much support they received from others. A network analysis approach to social support became attractive as community scientists distinguished between perceived social support and the relational or structural aspects of social support. That is, researchers began to 'move beyond the individual' (Felton & Shinn, 1992) and see that social support was not just a psychological characteristic but an outcome of being embedded in a social network of friends, family, and other support providers.

We started seeing studies that had participants identifying their own support networks. This allowed investigators to determine the size of the supportive networks, distinguish support by source, and even get some very simple measures of network structure. For example, in an innovative study of social support satisfaction, Stokes (1983) had each participant identify up to 20 people who were important in their lives and with whom they had monthly contact. After identifying the network member, each participant drew lines connecting each pair of people in their lists who were significant in each other's lives. This allowed Stokes

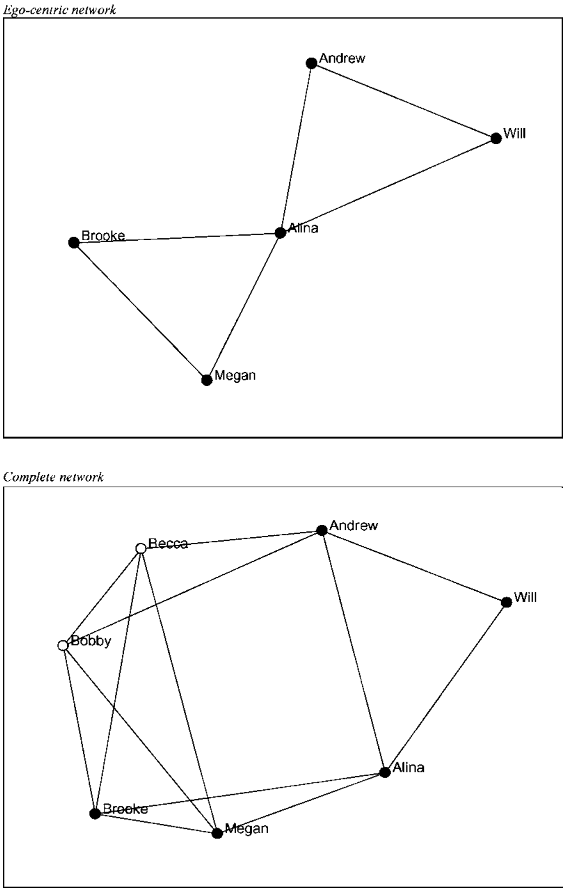


Fig. 5. Comparison of ego-centric friendship network to complete friendship network.

to not only look at source of support and satisfaction with support, but to also calculate the network measures of size and density. Density is the proportion of ties observed in a network relative to the total possible number of ties.

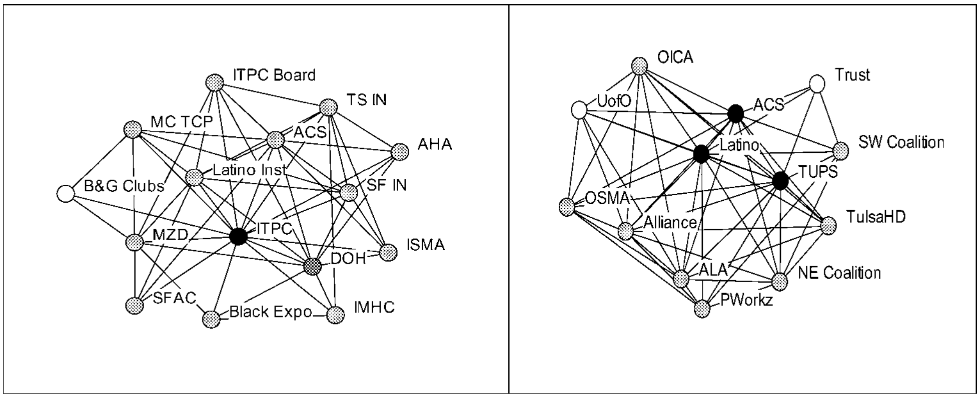
This type of network is known as an *ego-centric* network, because it is determined solely by the perspective of a single member. To see the limitations of ego-centric networks, consider Fig. 5. On the top is the type of network that you could get if you asked Alina who her friends were, and then asked her if her friends were friends with each other. However, this ego-centric network may not be representative of any real, complete friendship network. If you were to observe the children in action, or talked to all of the children, you might find out that you need to add two more members to the network (see bottom of Fig. 5). Becca and Bobby are not directly friends of Alina, but they are friends with Alina's friends. The characteristics of complete networks may or may not be similar to the ego-centric network. In particular, only direct relationships are examined in ego-centric networks. For example, you could not examine the indirect friendship support that Alina might receive from Bobby and Becca using only an ego-centric network.

The limitations of ego-centric networks are another illustration of Barton's 'meatgrinder survey approach.' If social networks are important because they influence people, then we need to observe and analyze complete networks that are not defined from the perspective of one member. Furthermore, most of the powerful network analysis techniques are only usable with complete networks.

Community scientists have started to recognize the utility of moving beyond ego-centric networks to examine complete networks, especially in the study of organizational relations and behavior. In an early classic example, Tausig (1987) used network methods as an assessment tool for a community mental health service system in a New York county. Using the presence or absence of interactions between all 45 mental health service agencies, Tausig was able to identify different types of "cracks" in the service system, including missing links and conflicted links.

More recently, Pennie Foster-Fishman, Deborah Salem, and their colleagues have used network methods in a series of their organizational studies (Foster-Fishman, Salem, Allen, & Fahrbach, 1999, 2001; Salem, Foster-Fishman, & Goodkind, 2002). In a particularly good example of using network tools on a complete network, they looked at the exchange relations within and between two interorganizational alliances within a single county. By collecting data on the complete interorganizational network, the authors were able to determine that organizational membership on both alliances was associated with a broader variety of exchange relations.

For a more detailed network example, consider Fig. 6. These data come from an evaluation study of 10 comprehensive state tobacco control programs conducted from 2002 to 2004 (Krauss, Mueller, & Luke, 2004). In conducting the evaluations we took a systems perspective (Best et al., 2003), and viewed a state tobacco control program not as one activity directed by one lead agency, but rather a set of interrelated activities coordinated by an organized network of tobacco control agencies and organizations. The network analysis data were collected to help us understand the structure of these complex state tobacco



Supportive Climate - Indiana (Centrality Index 22.7%)

Challenging Climate - Oklahoma (Centrality Index 6.6%)

Fig. 6. Comparison of contact networks for two states with divergent financial and political climates.

control programs, to identify other state characteristics related to program structure, and to determine if changes in program funding and political support were associated with changes in program structure.

Fig. 6 presents the contact networks for Indiana and Oklahoma as they existed in 2002. Each of the nodes represents one of the core tobacco control groups in the state. For example, in Indiana AHA is the American Heart Association, DOH is the Indiana Department of Health, and ITPC is Indiana Tobacco Prevention and Cessation, the lead agency for the state program. Two groups are connected by a line if they have frequent contact, defined as contact at least once a month via meetings, phone, or e-mail.

One of the things that is apparent from these networks is that the individual agencies vary in how "central" they are in the communication structure. For example, the Boys and Girls Club of Indiana has monthly contact with only three other groups in the state, whereas ITPC meets monthly with every member of the tobacco network. Thus, ITPC is more central in the network, and can be interpreted as being a gatekeeper, or having more control over the flow of information. This centrality is measured using Freeman's betweenness index (Wasserman & Faust, 1994); and is indicated in these graphs with color coding. Nodes with the highest degree of betweenness centrality are black.

Freeman's actor betweenness index is defined as follows—for any individual actor *i*, the actor's centrality is the proportion of times that actor *i* is included in the geodesics of actors *j* and *k*, summing across all *j*,*k* pairs. A geodesic is the shortest network path between any two actors. Thus, betweenness centrality measures how often an individual actor *i* is involved in the communication between other pairs of actors in the network. Betweenness centrality can range from 0 to 1. The individual actor betweenness scores can be combined into an aggregate network centralization index. Network centralization can range from 0, when all members have the same betweenness scores, to 1 (or 100%) when one member has maximum betweenness, and all other members have no betweenness. High-network centralization scores indicate hierarchical communications structures, whereas low-centralization scores indicate flat communication structures.

We have used this centrality information to assess the extent to which state financial and political climates influence the structure of state tobacco control networks. States with positive climates are those that have high levels of per capita spending on tobacco control, have high cigarette tax rates, have numerous tobacco control "champions" in the state, have actively supportive governors and legislators, and have a low tobacco industry presence. Indiana and Oklahoma's networks are presented here because they have remarkably different climates. In 2002 Indiana had a very positive climate for tobacco control, while Oklahoma's was very challenging.

The pattern we observed across the states that we evaluated was that states who had a strong positive climate tended to have communication networks that were relatively hierarchical, with the lead agency always being the most central actor in the network. Indiana's network is a good example of this. ITPC is the most central node, and the overall network centralization index of 22.7% indicates a moderately

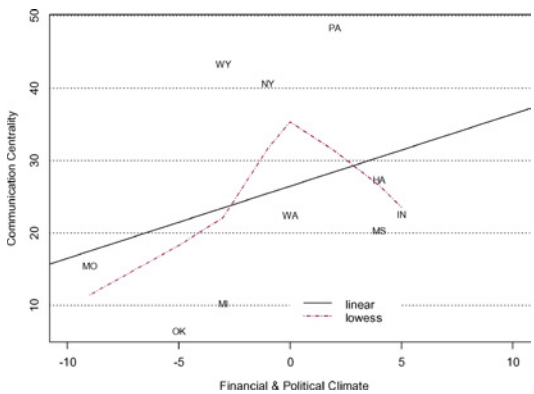


Fig. 7. Relationship of climate score to network centrality for 10 state tobacco control programs in 2002. (*Note*: the lowess line is a smoothed local regression line (Cleveland, 1993)).

hierarchical structure. States that had challenging climates, on the other hand, tended to show a different pattern. The communication structure was flatter, and the lead agency was not always the most central member. Oklahoma illustrates this—TUPS is the lead agency, but two other agencies, the Latino Agency and the American Cancer Society, also are highly central. Also, the centrality index for Oklahoma is only 6.6%, indicating a flat communication structure. Fig. 7 shows the relationship between state climate and centrality scores for all 10 states; as climate improves, the centrality scores get higher, indicating greater centralization of communication. We are interpreting this finding as suggesting that state networks are adapting to their environments. When times are good, a lead agency takes a central role in guiding activities and distributing financial resources. When times are not as good, there is less money flowing into the program, and other non-lead agencies step up and become more actively involved.

Network analysis is the only tool that can provide this type of structural information. From a policy perspective, this information is important because it can help us understand how to most effectively sustain important health and social programs even when the environment is challenging. Although network analysis requires a very different "non-survey" approach to data collection and analysis, it is a very useful component of a contextual toolbox for a community scientist.

Cluster Analysis

The bulk of the methods most often used by community scientists work by telling us information about how *variables* are related to each other (e.g., correlations, regression, factor analysis, etc.). However, there is also a class of analytic tools that works by telling us how *cases* are related to each other. These methods include cluster analysis, multidimensional scaling, network analysis, and various aspects of data mining.

A little over a decade ago, Rapkin and I suggested that cluster analysis should be of interest to community psychologists because it could reveal diversity and heterogeneity within our data (Rapkin & Luke, 1993). Since then, cluster analysis has been used sporadically by community scientists; for example, clustering was used in three empirical articles in AJCP during 2002 and 2003 (Fleishman et al., 2003; Salem et al., 2002; Zapert, Snow & Tebes, 2002). Cluster analysis works by grouping cases (often people) based on their similarities and dissimilarities across any number of variables. Cluster analysis is an exploratory technique that can be used to reveal unknown heterogeneity. This stands in contrast to traditional modeling where heterogeneity is handled by using covariates to control or adjust for these individual differences. Controlling heterogeneity in this way has two problems. First, controlling heterogeneity implies that the individual differences are not of scientific interest. Data heterogeneity may, in fact, be the most interesting part of the study. Second, this approach can sometimes lead to a type of methodological laziness, where certain covariates (e.g., age, SES, race, gender) are always included in a model because past studies have always included these covariates. Controlling heterogeneity is a way to decontextualize the data. By removing important individual differences, we are once again adopting Barton's meatgrinder approach.

What may not be as clear is that cluster analysis can be a useful tool for uncovering and describing contextual patterns. By grouping cases together on the basis of patterns of similarity, cluster analysis can be used to uncover naturally occurring groupings, social structures, or social types. Whereas network analysis and GIS methods are often used when the contexts are already defined (e.g., the classroom or the neighborhood), cluster analysis can be used to discover new examples of social contexts.

In a notable example of this, Kuhn and Culhane (1998) used cluster analysis on administrative data on public shelter use by single adults in New York and Philadelphia. They were able to identify three clear types of homeless persons on the basis of patterns of shelter usage. Transitional homeless were persons who entered the shelter system once and stayed for only a short period. The episodically homeless are persons who are frequently in and out of the shelter system. Finally, the chronically homeless used the shelter system less frequently, but stayed for much longer periods of time. Not only did these findings counter stereotypes of homelessness (the transitional group made up 80% of shelter use, and the chronic group constituted only 10%), but they have influenced subsequent social policy at the federal level (Trutko & Barnow, 2003). The new typology of homelessness identified by Kuhn and Culhane is a useful description of the context of homelessness that helps us understand the interplay between poverty, housing, and social services.

Another recent example from AJCP shows the advantages of cluster analysis. In a longitudinal study of adolescent substance use, Zapert, Snow, and Tebes (2002) used cluster analysis to identify subgroups of teenagers on the basis of their patterns of drug use from 6th to 10th grade. They were able to identify six distinctive clusters of substance users. These six clusters are shown in Fig. 8, which is based on the data presented in Table I in the original article. The graphs show that the clusters capture the complexity of teen substance use, with distinctive patterns across time and across different types of substances. These patterns may represent different classes of drug usage patterns. As such they represent contextual information that may have implications for community practice and policy. In particular, interventions can be focused more specifically on the right type of high-risk adolescent at the right time in their developmental trajectory given these types of data.

Incidentally, this type of graph uses the technique of "small multiples" (Tufte, 1983) to facilitate interpretation of the patterns across clusters, time, and substance. It is my experience that community and social scientists tend to use graphics merely to report statistical results, such as regression lines or line-plots of ANOVA interaction effects. This leads to graphs based on a small number of data points, or, in Tufte's terminology, a low "data-ink ratio." The power of graphics lies in their ability to reveal data

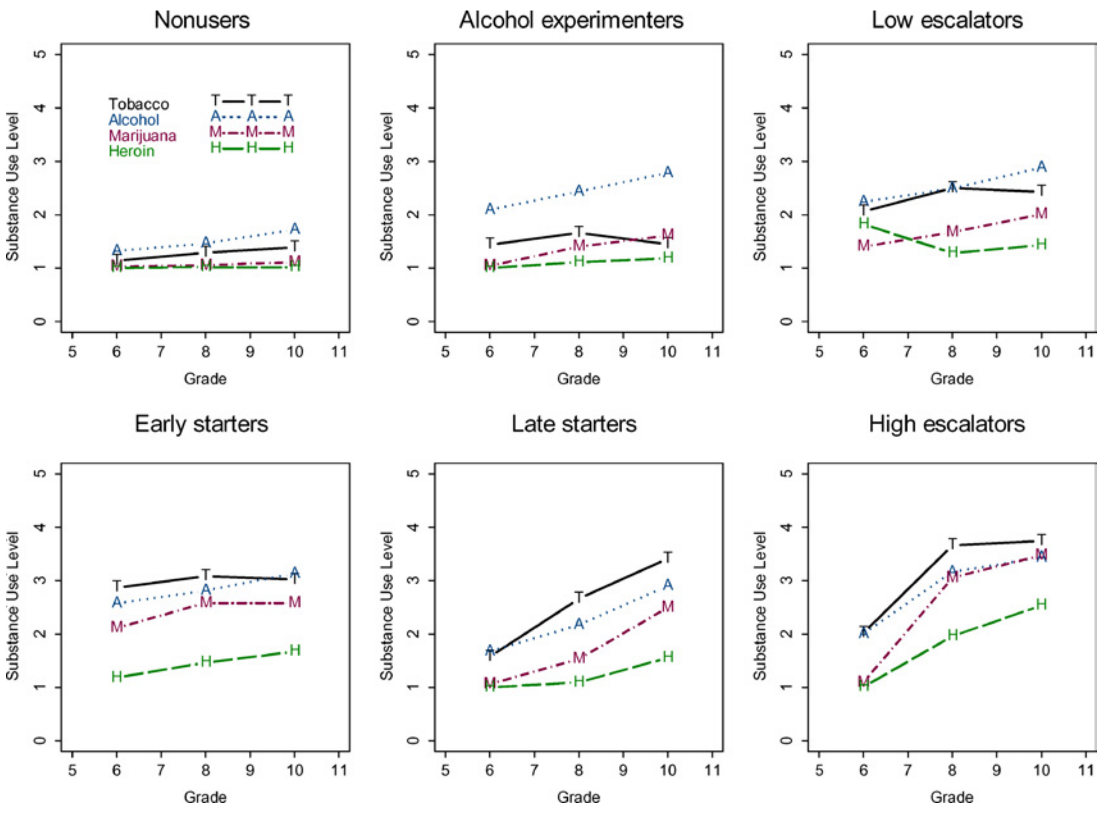


Fig. 8. Cluster profiles of teenager substance use patterns. Based on results presented in Zapert, Snow, and Tebes (2002).

complexity. Fig. 8 is a moderately complex graphic with three dimensions (i.e., cluster, grade, and substance) and is based on 72 data points. More importantly, it is much easier for the reader to see and understand the exciting structure of the data using this type of graph, than from a data table that simply lists the numbers.

DISCUSSION

In this paper I have argued that the uncritical use of traditional analytic methods can be inconsistent with core values and principles of community science. Specifically, traditional methods of the kind that Kelly refers to ignore or distort the *diversity* and *contextual* embeddedness of our data.

Toward Methodological Consilience in Community Science

Consilience has been defined as a "... measure of how much a theory explains, so that we can use it to tell when one theory explains more of the evidence than another theory" (Thagard, 1988). Thus, consilience is an epistemological construct that can be used to judge the adequacy of scientific theories. We can extend this idea to examine the adequacy of scientific methods for explaining the data. That is, *methodological consilience* is a measure of how much an analytic approach can explain.

A statistician usually thinks of a statistical model as a means for describing the variability in a dataset:

$$\text{Data} = \text{Modeled variability} + \text{Unmodeled variability}$$

Viewed this way, a model is successful when the *Modeled* variability is large relative to the *Unmodeled* variability. However, Tukey suggested a different way to think about modeling (Hoaglin, Mosteller, & Tukey, 1983):

$$\text{Data} = \text{Smooth} + \text{Rough}$$

Here, the purpose of modeling is to explore interesting structures and patterns (i.e., smooth) in the data. Also, modeling is viewed more iteratively—what is considered "rough" in one analysis may be the subject of another analysis to reveal previously unexamined patterns hidden in the rough.

By adopting Tukey's approach to modeling, we can see more clearly the relevance of methodological consilience for community science. Although he did not use these same terms, Roger Barker (1968) clearly understood that much of the "smooth" part of behavioral data that we want to understand as community scientists is accounted for by extra-individual settings, processes, and structures.

By using a wider array of statistical techniques such as cluster analysis, multilevel modeling, GIS, network analysis, and others, analysts are more likely to be able to explore and understand the complex data patterns of interest to community scientists. If context is a core value of community science, then our methods will be more consilient to the extent that they allow analysis of contextual effects.

Another way to frame this is to consider a lesson taught to me by my advisor in graduate school, Julian Rappaport. I remember a conversation in his office where he told me that a good community scientist needed three things: a phenomenon of interest, a theory, and a set of research methods. His purpose at the time, as far as I can recollect, was to gently remind a young graduate student who was increasingly interested in quantitative analysis that community psychology at its heart was always about something *real*. Community science, in other words, starts with the phenomenon of interest, or the social problem that the scientist wants to study and change. After many years of taking his lesson to heart, I would like to suggest a corollary—in order for our work to have meaning and effect real change, we need to make sure to use tools that are consistent with our values, and that can capture the big picture of community context.

ACKNOWLEDGMENTS

I would like to thank Jenine Harris for her help with collecting and coding the survey data used here. Thanks also to Nancy Mueller, Beth Shinn and the reviewers for comments that were extremely helpful in revising this paper.

REFERENCES

Alexander, C., Piazza, M., Mekos, D., & Valente, T. (2001). Peers, schools and adolescent cigarette smoking. *Journal of Adolescent Health*, *29,* 22–30.

Barker, R. G. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior.* Stanford, CA: Stanford University Press.

Best, A., Moor, G., Holmes, B., Clark, P. I., Bruce, T., Leischow, S., Bucholz, K., & Krajnak, J. (2003). Health promotion dissemination and systems thinking: Towards an integrative model. *American Journal of Health Behavior*, *27,* S206–S216.

Bybee, D. I., & Sullivan, C. M. (2002). The process through which an advocacy intervention resulted in positive change for battered women over time. *American Journal of Community Psychology*, *30,* 103–132.

Carver, S. (2001): *Participation and geographic information: A position paper*. ESF-NSF Workshop on Access to Geographic Information and Participatory Approaches using Geographic Information. Spoleto, December 2001.

Cleveland, W. (1993). *Visualizing data*. Summit, NJ: Hobart Press.

Cohen, S., Underwood, L., & Gottlieb, B. H. (2000). *Social support measurement and intervention: A guide for health and social scientists*. New York, Oxford University Press

Coulton, C., Korbin, J., & Su, M. (1999). Neighborhoods and child maltreatment: A multi-level study of resources and controls. *Child Abuse & Neglect*, *23,* 1019–1040.

Degenne, A., & Forsé, M. (2004). *Introducing social networks*. London: Sage.

Duncan, C., Jones, K., & Moon, G. (1998). Context, composition and heterogeneity: Using multilevel models in health research. *Social Science and Medicine*, *46,* 97–117.

Felton, B. J., & Shinn, M. (1992). Social integration and social support: Moving social support beyond the individual level. *Journal of Community Psychology*, *20,* 103–115.

Fleishman, J. A., Sherbourne, C. D., Cleary, P. D., Wu, A. W., Crystal, S., & Hays, R. D. (2003). Patterns of coping among persons with HIV infection: Configurations, correlates, and change. *American Journal of Community Psychology*, *32,* 187–204.

Foster-Fishman, P. G., Salem, D. A., Allen, N. A., & Fahrbach, K. (1999). Ecological factors impacting provider attitudes towards service delivery reform. *American Journal of Community Psychology*, *27,* 785–816.

Foster-Fishman, P. G., Salem, D. A., Allen, N. A., & Fahrbach, K. (2001). Facilitating interorganizational collaboration: The contributions of interorganizational alliances. *American Journal of Community Psychology*, *29,* 875–905.

Freeman, L. C. (2004). *The development of social network analysis: A study in the sociology of science*. Vancouver, Canada: Empirical Press.

Goldstein, A. O., Peterson, A. B., Ribisl, K. M., Steckler, A., Linnan, L., McGloin, T., & Patterson, C. (2003). Passage of 100% tobacco-free school policies in 14 North Carolina school districts. *Journal of School Health*, *73,* 293–299.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78,* 1360–1380.

Haining, R. (2003). *Spatial data analysis: Theory and practice*. Cambridge: Cambridge University Press.

Hedeker, D., McMahon, S. D., Jason, L. A., & Salina, D. (1994). Analysis of clustered data in community psychology: With an example from a worksite smoking cessation project. *American Journal of Community Psychology*, *22,* 595–615.

Hirsch, B. J., Mickus, M., & Boerger, R. (2002). Ties to influential adults among black and white adolescents: Culture, social class, and family networks. *American Journal of Community Psychology*, *30,* 289–304.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (1983). *Understanding robust and exploratory data analysis*. New York: Wiley.

Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Hillsdale, NJ: Erlbaum.

Kelly, J. G. (2002). The spirit of community psychology. *American Journal of Community Psychology*, *30,* 43–64.

Kelly, J. G. (2003). Science and community psychology: Social norms for pluralistic inquiry. *American Journal of Community Psychology*, *31,* 213–217.

Kivimäki, M., Vahtera, J., Elovainio, M., Pentti, J., & Virtanen, M. (2003). Human costs of organizational downsizing: Comparing health trends between leavers and stayers. *American Journal of Community Psychology*, *32,* 57–68.

Krauss, M. J., Mueller, N. H., & Luke, D. A. (2004) Interorganizational relationships within state tobacco control networks: A social network analysis. *Preventing Chronic Disease, 1*, [serial online]. Available from: URL: http://www.cdc.gov/pcd/issues/2004/oct/04_0041.htm.

Kuhn, R., & Culhane, D. P. (1998). Applying cluster analysis to test a typology of homelessness by pattern of shelter utilization: Results from the analysis of administrative data. *American Journal of Community Psychology*, *26,* 207–232.

Lin, N. (2000). *Social capital*. Cambridge: Cambridge University Press.

Luke, D. A. (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage.

Luke, D. A., & Krauss, M. (2004). Where there's smoke there's money: Tobacco industry campaign contributions and U.S. Congressional voting. *American Journal of Preventive Medicine*, *27,* 363–372.

Luke, D. A., Esmundo, E., & Bloom, Y. (2000). Smoke signs: Patterns of tobacco billboard advertising in a metropolitan region. *Tobacco Control*, *9,* 16–23.

Mankowski, E., Humphreys, K., & Moos, R. H. (2001). Individual and contextual predictors of involvement in twelve step self-help groups after substance abuse treatment. *American Journal of Community Psychology*, *29,* 537–563.

Mark, D. M., Chrisman, N., Frank, A. U., McHaffie, P. H., & Pickles, J. (2004). *The GIS history project*. Retrieved February 2004, from http://www.geog.buffalo.edu/ncgia/gishist/bar_harbor.html.

Mintz, B., & Schwartz, M. (1981). Interlocking directorates and interest group formation. *American Sociological Review*, *46,* 851–869.

Office of Behavioral and Social Sciences Research. (2001). *Toward higher levels of analysis: Progress and promise in research on social and cultural dimensions of health*. NIH Publication No. 21-5020. Available online at: http://obssr.od.nih.gov/news/publications.html.

Rapkin, B., & Luke, D. A. (1993). Cluster analysis in community research: Epistemology and practice. *American Journal of Community Psychology*, *21,* 247–277.

Revenson, T. A., D'Augelli, A. R., French, S. E., Hughes, D., Livert, D., Seidman, E., Shinn, M., & Yoshikawa, H. (Eds.). (2002). *Ecological Research to Promote Social Change: Methodological Advances from Community Psychology*. New York: Plenum.

Rothenberg, R. B., Potterat, J. J., Woodhouse, D. E., Muth, S. Q., Darrow, W. W., & Klovdahl, A. S. (1998). Social network dynamics and HIV transmission. *AIDS*, *12,* 1529–1536.

Salem, D. A., Foster-Fishman, P. G., & Goodkind, J. R. (2002). Adoption of innovation in collective advocacy organizations. *American Journal of Community Psychology*, *30,* 681–710.

Sarason, S. B. (2003). The obligations of the moral-scientific stance. *American Journal of Community Psychology*, *31,* 209–211.

SCRA. (2004). *The SCRA mission*. Retrieved February 2004, from http://www.apa.org/divisions/div27/goals.html.

Shinn, M., & Rapkin, B. D. (2000). Cross-level research without cross-ups in community psychology. In J. Rappaport & E. Seidman (Eds.), *Handbook of community psychology* (pp. 669–693). New York: Kluwer Academic.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Stokes, J. P. (1983). Predicting satisfaction with social support from social network structure. *American Journal of Community Psychology*, *11,* 141–152.

Tausig, M. (1987). Detecting "cracks" in mental health service systems: Application of network analytic techniques. *American Journal of Community Psychology*, *15,* 337–351.

Thagard, P. (1988). *Computational Philosophies of Science*. A Bradford Book. Cambridge: MIT Press.
Trutko, J., & Barnow, B. (2003). *Core performance indicators for homeless-serving programs administered by the U.S. Department of Health and Human Services: Final Report*. Available online: http://aspe.hhs.gov/hsp/homelessness/perf-ind03.
Tseng, V., Chesir-Teran, D., Becker-Klein, R., Chan, M. L., Duran, V., Roberts, A., & Bardoliwalla, N. (2002). Promotion of social change: A conceptual framework. *American Journal of Community Psychology*, *30,* 401–428.
Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CN: Graphics Press.
Valente, T. W. (1995). *Network models of the diffusion of innovations*. Cresskill: Hampton.
Wasco, S. M., Campbell, R., & Clark, M. (2002). A multiple case study of rape victim advocates' self-care routines: The influence of organizational context. *American Journal of Community Psychology*, *30,* 731–760.
Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.
Wicker, A. W. (1992). Making sense of environments. In W. B. Walsh & K. H. Craik (Eds.), *Person-environment psychology: Models and perspectives* (pp. 157–192). Hillsdale, NJ: Erlbaum.
Wieczorek, W. F., & Hanson, C.E. (1997). New modeling methods: Geographic information systems and spatial analysis. *Alcohol Health and Research World*, *21,* 331–339.
Zapert, K., Snow, D. L., & Tebes, J. K. (2003). Patterns of substance use in early through late adolescence. *American Journal of Community Psychology*, *30,* 835–852.