

# **Literature Review and Analysis: Public Health and Public Perspectives on Measuring Alcohol Policy Enforcement and Compliance**

**Michael Klitzner, Ph.D.**  
Principal Social Scientist  
The CDM Group, Inc.

**September 2002**

Paper prepared for the National Institute on Alcohol Abuse and Alcoholism (NIAAA), National Institutes of Health (NIH), U.S. Department of Health and Human Services (HHS), as part of the Alcohol Policy Information System (APIS) under Contract No. N01AA12009 to The CDM Group, Inc. (primary contractor) and Pacific Institute for Research and Evaluation (subcontractor).

The non-U.S. Government information, opinions, data, and statements contained herein are not necessarily those of the U.S. Government or the National Institute on Alcohol Abuse and Alcoholism (NIAAA), National Institutes of Health (NIH) and should not be interpreted, acted on, or represented as such. Information about data sources for the measurement of enforcement and compliance of alcohol-related policies is provided here as a convenience for researchers. The list of sources is not comprehensive and no inference should be drawn from the inclusion or exclusion of a particular source.

## INTRODUCTION

Within the public health/public policy literature, enforcement and compliance fall under the general rubric of implementation (i.e., the extent to which the conditions intended and/or necessary to obtain an impact are met). In the early '70s, Pressman and Wildavsky (1973) noted that any definition of policy is pauperized when implementation is not addressed. They further sought to illuminate by case example “how difficult it is to make the ordinary happen,” and thus, that implementation cannot be assumed. By the late '80s, this sentiment was reflected in the public health and policy literature (see Moskowitz, 1989) and has continued to gain acceptance (see Holder, 2000).

Public health and public policy are not primary theoretical fields. Rather, like medicine, engineering, and other applied disciplines, the main concern of public health and public policy is to develop strategies to address specific problems. These strategies are developed based on available knowledge, empirical findings, theory (from other fields), and practical considerations related to costs, feasibility, and acceptability to important constituencies. Similarly, public health and public policy research tends to focus on specific problems and is designed either to inform the development or choice of strategies (surveys of various types, analysis of social indicators, descriptive studies of population or programs) or to test the feasibility and efficacy of strategies once they are implemented (outcome studies of various types and sophistication). There is no preemptive concern with theory in this research. Rather, theory is one of many tools used to achieve largely practical goals.<sup>1</sup>

Accordingly, questions deemed central by the academic researcher (e.g., by what mechanism does enforcement or compliance affect health or social outcomes) may be of less concern to the policy or public health researcher. Although researchers in these fields do not doubt Lewin's dictum that there is nothing so practical as a good theory, most who work in the field also believe that there is nothing so useful as a robust result. For example, the clear finding that properly used child restraints reduce mortality and morbidity has direct implications for public policy and public health practice. This utility is independent of any understanding of the physics of crashes or the anatomy and physiology of babies and young children. Of course, such understanding may lead to the design of better restraints, but that activity is outside the scope of the two disciplines considered in this paper. Similar arguments could be made for findings concerning hypertension control and heart disease, speed limits and gasoline consumption, asbestos and lung cancer, ground-fault interrupt circuits and electrocution, and so on. Thus, many public health and public policy interventions are non-teleological and reports of studies concerning them may have little discussion of theories or mechanisms of effect.

This is not to imply that mechanisms are viewed as “black boxes” (although this does occur). Rather, mechanisms are either considered of secondary importance or are so obvious that they do not warrant much discussion. For example, this paper will later

---

<sup>1</sup> One may contrast the medical sociologist studying provider–patient interactions in order to test theories of social control and the public health researcher studying the same interactions in order to develop a curriculum to improve provider communication skills.

discuss an article on adherence to drug regimens for tuberculosis in which there is no discussion of the fact that antibiotic drugs kill tubercle bacillus or the mechanism by which this occurs. This information is simply considered too obvious to mention in a medical journal.

Given the above discussion, it seems appropriate to speak of *perspectives* (or *approaches*) rather than theoretical frameworks for the current analysis. Six such perspectives/approaches will be discussed, three related to etiology and intervention design and three related to measurement.

*Etiologic perspectives* provide the framework for the measurement of enforcement and compliance. Historian of science S. J. Gould (2002) notes that definition of variables in science must follow G.K. Chesterton's comment on art: The essence of any painting lies within the confines of its frame. Without the limitations imposed by a framework, variables are condemned to "mushy indefiniteness and consequent vacuity" (Gould, 2002, page 765). Such frameworks also allow accurate interpretation of data. Lord Kelvin's famously inaccurate measure of the age of the earth came not from imprecision in his data (applying his 19th century measures within the proper framework yields estimates that are startlingly accurate by modern standards), but rather from a flawed understanding of the geothermal processes that produced the heat he measured.

Similarly, the successful *measurement* of enforcement and compliance, including choice of measures and their analytic treatment, depends heavily on the presumed relationships over time between underlying causes of social or health problems, phenotypic manifestations of the problem in an individual or population, the policy or intervention implemented to impact the problem, the enforcement mechanisms used to support the policy, the type of compliance required, contextual factors, and so on. Absent at least a preliminary set of assumptions about these relationships, enforcement and compliance (or any other variable) can generally not be indexed in any meaningful way, as we will examine later in this paper.

Thus, etiologic and intervention perspectives/approaches provide guidance in developing necessary frameworks for measurement. They do not suggest *specific* measures, but in some cases do have concrete implications for the measurement or analysis process, and these will be discussed.

The three measurement approaches to be discussed *do* suggest specific measures. These three approaches can be linked to the three etiologic and intervention perspectives, although no simple one-to-one relationships are implied. Rather, each measurement approach can be associated with one or more of the discussed perspectives (relationships that will be obvious to the reader). It would be wrong to assert that the measurement approaches necessarily grew out of these perspectives. Rather, all six perspectives have developed in parallel and in interaction with one another over time, and have all been influenced by predominant schools of thought in the social and health sciences in the 20th century.

## **Definitions of Enforcement and Compliance**

There appears to be no accepted or consistently used definition of *enforcement* in the public policy or public health literatures. Rather, this term is part of the large corpus of common language terms that scientists use in the same way as lay persons and that are intermingled with technical terms and jargon in scholarly discourse and writing. At an operational level (e.g., models building/measurement), enforcement is commonly defined by some measure of enforcement activity (varying in type, scope, intensity, or any combination). Enforcement is most commonly used with reference to policing or monitoring activities and less commonly used with reference to the processing of violations or the imposition of sanctions.

*Compliance* is also generally used as common language, i.e., acquiescence to a standard. At the operational level, compliance is generally measured as some percentage of a group or population that does or does not follow a given rule. In public health, compliance also has a technical meaning, defined as either (1) the extent to which a patient follows a medical regimen (e.g., a course of antibiotics) or (2) the extent to which a practitioner follows a prescribed set of guidelines in diagnosis or treatment. At the operational level, this technical use of compliance is reflected in measured deviations between the prescribed regimen, procedure, etc., and the actual behavior of patients or practitioners (e.g., six of seven daily doses completed; occult blood screening offered to 60 percent of patients over 40).

## **Measurement Terminology in Public Health and Public Policy**

The public health/public policy literature generally distinguishes between *measures* and *indices*, and between the activities of *measuring* and *indexing*. *Measures* are generally used to refer to numbers (or categories) derived from a single data collection activity, and aggregated across cases. Thus, a radar survey of speed (*measuring* speed; see Hingson et al., 1996) yields a single datum for each car that can be aggregated across cars to yield a *measure* of speed (or compliance with speed limits). A measure may rely on a number of items on a questionnaire or survey that are combined (usually according to either a simple additive model or an empirically developed set of weights) to yield a single datum per case.

An *index* refers to a higher-level aggregation of data across measures that is either hypothesized or known to provide a more accurate measure of a construct than can be gained from any single datum available for each case. Thus, Morrissey, et al. (1985) and Cohen and Kennedy (1983) use a number of measures of interorganizational communication and cooperation (number of referrals, number of meetings, understanding of each other's organizational goals) to assess the extent to which health and mental health systems conform to a "systems of care" model. Here, no single measure is deemed adequate to capture the construct of a system of care. Rather, an index is needed. To *index* a construct is to create, implement, and combine a series of relevant measures. Thus, all enforcement and compliance indices follow the same general formula:

$$\text{Index}_{E/C} = f(m_1, m_2, m_3, \dots, m_n)$$

In public health, and to some extent in public policy, the term *indicator* is used to refer to a source of data collected for some reason other than research. Thus, when arrest rates, which will be discussed below, are used by researchers to measure enforcement or compliance, the actual numbers used are most often compiled by official agencies for their own internal purposes. Similarly, mortality and morbidity statistics, census data, and other information collected by public agencies can be used in various ways to construct measures of enforcement and compliance. They are not, however, primarily collected for use by researchers. The designation of a measure as an indicator carries an implicit red flag: Such data can be highly misleading and may fail to reflect reality for a variety of reasons. As Sir Josiah Stamp, quondam head of the Inland Revenue Department, U.K., (1896–1919), famously noted:

The Government are very keen on amassing statistics—they collect them, add them, raise them to the nth power, take the cube root and prepare wonderful diagrams. But what you must never forget is that every one of those figures comes in the first instance from the village watchman, who just puts down what he damn well pleases.

## **ETIOLOGIC AND INTERVENTION PERSPECTIVES/APPROACHES**

It is possible to contrast three public health and public policy perspectives/approaches; *catchment-area*, *traditional*, and *systems*.

### **Catchment-Area**

In the *catchment-area* approach (a public health construction), prevention or treatment interventions are provided to selected, high-risk individuals (Holder, 1998). The sociologic concept of secondary deterrence (Ross, 1984) and the juvenile justice continuum (prevention, intervention, adjudication, supervision) of the Office of Juvenile Justice and Delinquency Prevention (OJJDP) provide public policy expressions of the same idea.

The catchment-area approach was originally designed to address infectious illness and has been successfully applied to public health problems such as community heart disease and cancer prevention. Under the catchment-area approach, enforcement and compliance can straightforwardly be considered as aspects of the “treatments,” contributing to their strength and integrity.<sup>2</sup> By definition, they are measured at the level of high-risk individuals since these are the exclusive concern of this approach. Examples of the catchment-area approach to compliance measurement in the public health/public policy literature range from measures of patient compliance with drug regimens for tuberculosis and HIV/AIDS (Lucas, 2001), rates of smoking cessation among heart disease patients (Rigotti et al., 1994), and rates of condom use among sexually active adolescents (American Academy of Pediatrics, 2001) to offender compliance with terms of probation, and recidivism and relapse rates. An example of enforcement measurement comes from Huang (2002), which studied the impact of child support enforcement on non-marital and marital births. This study indexed enforcement using archival data on the nature of the

---

<sup>2</sup> On closer examination, of course, the definition of enforcement and compliance as treatment variables is overly simplistic. In the case of tuberculosis, for example, enforcement of and compliance with treatment by affected individuals reduces the probability that additional individuals will be infected.

enforcement policy and the budget allocated to its implementation. Other examples will follow in a revision of this paper.

The catchment-area approach has clear limitations when applied to public health/public policy challenges such as injury prevention, violence, traffic safety, HIV/AIDS prevention, alcohol problems, and drug abuse. As noted by Holder (1998), although high-risk individuals experience the highest rates of problems, their absolute numbers are so small that they contribute only modestly to aggregate rates of mortality and morbidity considered across populations (Edwards et al., 1994). Thus, even highly successful strategies based on the catchment-area approach cannot produce more than modest impacts on some health and social problems. The limited effects of prevention efforts in a variety of health areas can be traced to their reliance on the catchment-area approach (Klitzner, 2000).

In contrast to the catchment-area approach, *traditional* and *systems* approaches share a focus on populations rather than individuals. Interventions are generally targeted at populations, and effects are measured by risk reduction or behaviors aggregated at the population level. Traditional and systems approaches also include an extensive focus on the environments in which individuals behave.

### **Traditional**

The *traditional* approach derives from what Claude Levi-Strauss termed *structuralism* (Levi-Strauss, 1963). Structuralist analysis begins with the categorization of relevant constructs into parts. In public policy, the traditional approach is reflected in what Almond called *structural-functionalism* (Almond and Coleman, 1960), while in public health it can be found in any number of variants of what is often called the *public health model*.

Structural-functionalism divides the policy-making process into input functions (political socialization and recruitment, interest articulation, interest aggregation, and political communication) and output functions (rule-making, rule application, and rule adjudication). Enforcement and compliance will fall into the adjudication category. Structural functionalism uses a simple open system as its organizing principle. Open systems models, which derive from engineering and manufacturing, translate inputs (needs and requirements) into outputs (designs and products) through the application of engineering and manufacturing processes (throughputs).<sup>3</sup> Structural functionalism and other open systems approaches suggest a linear measurement model in which a set of independent variables is hypothesized to affect one or more dependent variables—the canonical paradigm for not only policy and public health research, but all research in the social sciences. As discussed beginning on page 11, the linear measurement model has been widely applied but has limitations.

In its original form, the public health model posited that hosts and agents interact within their environments to determine vectors of disease transmission. It thus constitutes a rudimentary non-linear model (i.e., one that allows feedback and multiple/reciprocal

---

<sup>3</sup> For an example from weapons system design and manufacture, see Hanratty et al. 1999.

causation). This model has been adapted to a wide variety of behavioral health and behavior problems (including problems that are traditionally considered social such as violence). Such adaptations have required redefinition and expansion of the terms *hosts*, *agents*, and *environments*. These redefinitions and expansions have, in turn, led to misspecification of problems, and transformation of the public health model into no more than an extended metaphor in many instances (Gruenewald et al., 1997). As Gruenewald et al. (1997) note, the redefined and expanded public health model fails to provide clear criteria for classifying such social factors as hosts, agents, or environments. It has become “three empty conceptual boxes into which various social phenomena may be grouped” (page 20). To our knowledge, the public health model has not been used to develop measures of enforcement or compliance, although studies such as the Lucas (2001) examination of compliance with drug regimens clearly rely on an approach to disease prevention that is based on the public health model.<sup>4</sup> However, it is clear that the model itself offers little guidance in either increasing compliance or in measuring it.

Thus, the public health model fails at the level of specification of variables and structural-functionalism fails at the level of interaction among variables. Although the traditional perspective addresses conceptual limitations of the catchment-area model, its utility as a guide to measuring enforcement and compliance is limited.

## Systems

In response to the limitations of the catchment-area and traditional approaches, public health practitioners and researchers began to explore the applicability of systems dynamics to public health problems. The term *community* appears frequently in the public health and public policy systems literature, perhaps because some of the earliest work in this area was done by community psychologists and because communities as colloquially understood remain an important focus of systems-driven research in public health and public policy.<sup>5</sup> However, systems analysis can be applied to any definable social system (e.g., a hospital, other work settings, a school or school system, a public safety department, etc).

As noted by Holder (1998), *systems* approaches owe a significant debt to Lewin (1947), who introduced the interactionist perspective to social science, and to Saronson (1974), who conceptualized the community as a variety of institutions which are formally and informally related to each other. Current applications in public health and public policy reflect Stoneall's (1983) network-exchange perspective—a view in which a social system is defined by economic and social exchange, social control, participation, and support. Under this definition, human organizational and institutional arrangements are sufficiently homogeneous to allow reasonably precise specifications of community

---

<sup>4</sup> This dim view of the expanded public health model contrasts sharply with its clear utility for its originally intended purpose. When one wishes to understand the spread of a well-defined pathology (e.g., small pox), the traditional perspective provides significant traction and precision. Unlike its application to social pathologies, the criterion of well-specified causal relationships among clearly defined variables is met.

<sup>5</sup> We would suggest that “community” as used in the systems literature generally refers to a grouping of individuals and/or organizations within a defined geographic boundary and/or a defined group. So for example, a community might subsume a single housing unit, a city, a State, or an entire nation. It might also subsume all doctors (the medical community), all people of faith (the faith community), and so on.

systems related to specific health and social problems. The systems approach has much in common with behavioral ecology—the exploration of the individual, social, and environmental dynamics that underlie human behavior (Smith and Winterhalder, 1992).

The systems perspective requires the specification of a dynamic system (elements and their interactions over time). This approach was originally developed to model industrial and economic processes (Forrester, 1961, 1968), but has been widely used for other systems including social and health problems (Alfred and Graham, 1976; Holder, 1989; Levin et al., 1972, 1975). The application of large-scale systems models in the social sciences has been predominantly within the domain of macroeconomics (Klein, 1991; MacRae, 1985). Sociologists, political scientists, and criminologists have begun to develop macrosociological models of social and criminal behavior (e.g., Maltz, 1984; Tuma and Hannan, 1984), but these models are currently restricted to focal areas (e.g., political structure, social diffusion processes, recidivism) in which the global complexities of social processes can be ignored.

A systems dynamics approach extends conventional analytic methodologies by locating health and social problems within a complex network of variables that interact with one another in an iterative fashion over time (Mitroff and Sagasti, 1973). Using the systems approach, Holder (1998) conceptualizes social and health problems as “the accumulative result of the structure and interactions of complex social, cultural, and economic factors within the community system” (p. 12). Thus, the systems approach organizes and synthesizes what Holder (1998) refers to as a “hodgepodge of fragments” into a coherent and analyzable structure. Holder notes that the systems perspective renders the apparently chaotic behavior of complex systems amenable to analyses, which is not possible with classical deterministic approaches. Gruenewald et al. (1997) put the same idea this way: “It is in frank recognition of the contingent nature of social behavior that the systems perspective gains its power and flexibility” (p. 22).

The systems approach has clear implications for the measurement of enforcement and compliance (and, for that matter, any variable). First, the location of enforcement and compliance within a larger system is a prerequisite for developing measures of these variables. This is because any given variable in a system is affected by and affects other variables in the system, and these relationships must be explicated if empirical estimates of effects are to be made (Gruenewald et al., 1997).<sup>6,7</sup> In other words, circumstances under which enforcement or compliance can be accurately indexed through single measures of isolated variables will be rare (although not impossible, as will be discussed later in this paper). Second, there is little basis for assigning enforcement and compliance to the traditional categories of independent and dependent variables. Almost all variables in a systems model are both causes and effects. In this regard, Gruenewald et al. (1997) note that indicators are products of—and embedded in—the systems they’re designed to

---

<sup>6</sup> Holder (1998) notes that one popular way to represent community systems is through lines that connect “everything to everything.” He further notes that such “systems” reflect either naïveté or ignorance concerning the relationships they actually obtain.

<sup>7</sup> Both Holder (1998) and Gruenewald et al. (1997) suggest that the ability to build a mathematical (computer) simulation that mimics changes in a real community over time is one powerful test of the underlying systems model.



study. This embeddedness raises analytic issues that cannot be separated from issues of measurement (see **Synthetic**, beginning on page 13).

Finally, the systems perspective suggests that smaller levels of aggregations (e.g., communities, counties) are generally to be preferred over larger levels of aggregation (States, the nation). This is because systems of enforcement are often local, and aggregation to higher levels may mask important differences.<sup>8</sup> However, Gruenewald et al. (1997) note that higher levels of aggregation are often the only feasible option and can yield useful information on enforcement effects (see, for example, Ross, 1982).

## **MEASUREMENT PERSPECTIVE/APPROACHES**

Three approaches to the measurement of enforcement and compliance can be identified in the public health/public policy literature. For lack of better terms, these approaches may be called *descriptive*, *linear*, and *synthetic* (combining theory-building and measurement).

### **Descriptive**

A lack of process data hampered early attempts to evaluate the effectiveness of public health interventions. In his critical review of the primary prevention of alcohol problems, Moskowitz (1989) concluded:

Little information is communicated [in research reports] as to what the intervention was, let alone how well or how extensively it was implemented. The lack of data on mediating variables and the absence of process or implementation data severely hamper our ability to interpret research findings (p 77).

Moskowitz's paper and many other similar critiques led to the adoption of process evaluation (implementation analysis) as an expected component of public health and, to a lesser degree, public policy research. Most often, implementation analysis has taken the form of *descriptive*—or qualitative—analysis, the approach originally used by Pressman and Wildavsky (1973) in their classic study of the Oakland Project. In the main, the descriptive approach to indexing enforcement involves gathering, through interviews, observation, and review of records, data on whether or not enforcement (or increased enforcement, or a new enforcement strategy) was implemented and (sometimes) with what level of strength and integrity.

For example, Homel (1988) was able to compare the effects of “weak” and “intense” implementations of random breath testing and random stopping programs in Australia. Here, the classification of these enforcement programs as “weak” or “intense” is justified through synoptic analyses of implementation rather than quantitative measurement. Homel provides little information on his “weak” program except to note that it was highly controversial and was opposed by two major dailies. However, his synoptic

---

<sup>8</sup> This problem is well known in social epidemiology, where national or State averages may mask significant local differences in health or social problems (alcohol use, teen pregnancy, HIV/AIDS, etc.).

analysis of one “intense” site gives a flavor for what can be accomplished with a descriptive analysis:

In 1985 more than 200,000 roadside tests were conducted out of a driving population of only 268,887! Three mobile breath analysis units are commonly transferred to at least five different sites in an eight-hour shift, and since they are equipped with breath analysis equipment as well as cooking, toilet and electricity generating facilities, testing can proceed uninterrupted over the shift. Occasional massive operations... are undertaken, and only after their completion are public announcements about them made by the authorities. It is thought that by this method the public will remain convinced that RBT is not an empty threat. Since Tasmania is a small state, publicity through the electronic media is not bought by the Government, but extensive (free) publicity is achieved through newspapers. In particular, a daily list of the names of convicted drink-drivers is published in the papers (page 235).

Note that although numbers are used, no ratio (e.g., tests/100,000 population) are presented. Rather, the author comments on the meaning of the numbers with his punctuation. Note also that the justification for categorizing the site as “intense” is implicit rather than stated. However, the specific criteria in the mind of the investigator may be imputed: (1) the number of citizen contacts is very high, (2) testing is uninterrupted over extended periods of time, (3) evidentiary testing can be easily done, and (4) the perception of threat is kept high. The contributions of these factors to the intensity of enforcement is left to the reader. It is clear from other sections of the article that publicity is deemed important and can be assessed from the information in the paragraph above. Note again, however, that no numbers are presented (number of PSAs aired, column inches devoted to RBT, etc.). Other studies using similar descriptive measures exist and are being reviewed for inclusion in a revision of this paper.

Descriptive techniques have also been used to index compliance. For example, pseudopatients (trained observers posing as patients), interviews, and observations are used to assess health provider compliance with recommendations for clinical practice (Greer, 1987; Greer and Greer 1991; Mechanic, 1978). Interviews and observation have also been used to index the fidelity of community organization activities to their stated goals (Klitzner et al., 1990), the use of safety belts by motorists (Hingson, 1996), and the compliance of patients with medical regimens (Lucas, 2001). In the alcohol area, Saltz and Stangherra (1997) used observation to assess the extent to which servers follow responsible serving practices.

Compliance with safety belt use offers few conceptual or data collection problems; the belt is either fastened or not. Although smoking is, in most of the health literature, an all-or-nothing variable, it can be difficult to measure because self-reports can be unreliable or systematically biased. Thus, in Rigotti et al. (1994), saliva cotinine assay was used to validate self-reports (this is standard procedure in smoking research).

Some compliance variables are not easily dichotomized. In Klitzner et al. (1990), the fidelity of implementation was defined as conformance with a set of criteria provided by a national umbrella organization. As would be expected, interview and observational data collection revealed that individual community organizations deviated from the criteria,

but in different ways. The investigators coded the organizations as “strong,” “moderate,” or “weak” based on theoretical considerations of the importance of each criterion to the effects the organizations sought to bring about. These sorts were found to match the general impressions of the national umbrella organization.

Measures of patient compliance can often be dichotomous (expressed, for example, as completion rates of TB therapy). In most cases, an individual who is not totally compliant but who meets some minimum criteria for therapeutic efficacy will be coded as compliant. In a more limited number of cases, clinical efficacy requires total compliance. In either case, the measurement strategy relies on a clinical criterion.

However, some compliance measures are intrinsically scalar. For example, category change for HIV-positive persons has not been achieved. Interestingly, in this case, two different functions relating compliance to outcomes are discussed by Lucas (2001). For reduction of transmission risk, a preferred measure of compliance is reduction in viral load and the function is monotonic. However, when the measure is development of drug-resistant HIV, the function appears to be bell-shaped. Lucas does report pill-count adherence rates (which can be measured in any number of ways, including bottle caps with microchips), but then goes on to relate these to rates of primary-drug resistance mutations and finds a threshold at 60 percent adherence.

Greer and Greer (1991) measured physician compliance with practice guidelines through interviews with practitioners. Among many topics, these interviews explored whether any guidelines were used by the practitioner, how they were used, and reasons for use or non-use. From these data, the investigators concluded that compliance with guidelines (which were, in very large degree, optional at the time of the study) was minimal. The most important policy findings from this study concerned the reasons for non-compliance, a topic beyond the scope and purpose of this paper.

Descriptive approaches can be powerful and compelling, but become more complex as the unit of analysis increases in size. Thus, a descriptive analysis of policy enforcement may be feasible in a modestly-sized community. Expanding this analysis to an entire State would require either a very large study or the construction of a sample of communities. It should be noted that, although descriptive studies produce qualitative data, these data can and have been used in sophisticated quantitative policy analyses (Gruenewald, 1997).

## Linear

*Linear* measurement approaches are familiar to anyone who has taken an introductory course in social science methods. Adapting Figure 1.1 of Holder (1998), this approach can be diagrammed as follows:



The linear measurement approach tests the hypothesis that changes in enforcement compliance output changes in social or health problems through mechanisms that may or may not be specified. The similarity to the open system described above is obvious. Actions are inputs, the “black box” contains the throughputs, and desired results are outputs.

Like the open system model, the measurement model underlying the linear approach allows no feedback from outcomes to enforcement compliance and either ignores variables that may affect both enforcement/compliance and outcomes or controls them in the analysis. This model is generally an acceptable representation of reality when certain conditions are met. Primary among these is that the measures are direct and largely uncontaminated by extraneous system effects (Gruenewald et al. 1997).<sup>9</sup> Of course, the linear measurement approach requires that data collection and analysis conform to standard desiderata of survey research (sampling, replacement, analysis, etc). So, for example, intersection cameras and radar studies (Hingson et al. 1996) provide direct measures of compliance with traffic laws, ozone testing provides a direct assessment of air quality, and surveys provide a direct assessment of self-reported criminality (e.g., Grasmick et al., 1993) and citizen beliefs about police enforcement (Hingson et al., 1996).

In the alcohol area, random roadside blood-alcohol-content (BAC) surveys (Homel, 1988) provide a direct measure of compliance with BAC laws and can be expected to be relatively stable (within the limits of secular trends) absent specific interventions to alter them. Similarly, the use of underage decoys (Holder, 2000) provides a direct and largely uncontaminated measure of compliance with sales to underage individuals. Other examples of fruitful applications of the linear measurement approach to alcohol problems are provided by Ross (1982), Gruenewald et al. (1993), and Watts and Rabow (1983).<sup>10</sup>

The methodological issues associated with direct measurement of enforcement or compliance are the same as those encountered in any survey research. Primary among these is sampling. Some issues, such as sampling retail outlets (Holder, 2000) and radar speed studies (Hingson et al., 1996), have been addressed in several studies. However, textbook solutions are unlikely to be found for other, newer methods. More details on how the above cited studies addressed these issues will be presented in a revision of this paper.

---

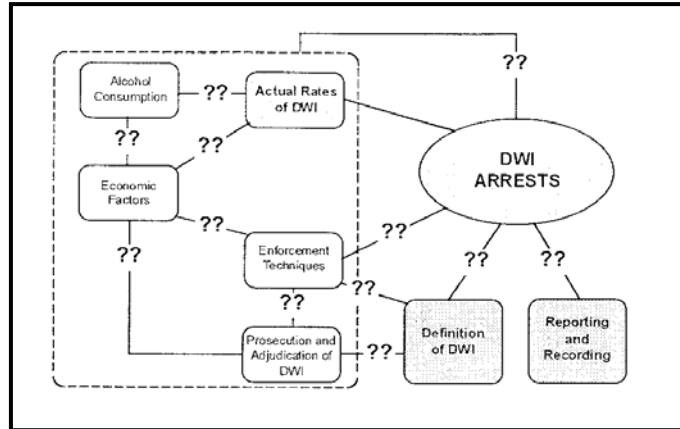
<sup>9</sup> To use a medical example, both venipuncture and sphygmomanometry can be used to measure blood pressure. However, the former approach is relatively direct while the latter approach is prone to contamination through extraneous system effects, such as variations in technique or placement of the cuff.

<sup>10</sup> It is tempting to suggest that the linear approach will find many more applications in measuring compliance than in measuring enforcement. This is because compliance is a circumscribed (if sometimes complex) behavior of individuals or organizations. As such, traditional psychometric approaches from the behavioral sciences (of which the linear approach is a prime example) provide a variety of well-tested measurement strategies and options. By contrast, enforcement is most often the output of some sort of system (e.g., the activities of police, prosecutors, and courts, legislated sanctions). Traditional psychometric approaches offer fewer measurement tools under these circumstances.

## Synthetic

Problems arise when the conditions discussed in the previous section are not met. Arrest rates (a traditional and standard index of crime and enforcement rates; see Maltz, 1984; Sellin & Wolfgang, 1964) provide the canonical example of these problems.

Gruenewald et al. (1997) show the futility of using arrests to index either compliance or enforcement (in this case, with driving-while-intoxicated [DWI] laws). Their Figure 2.1 (page 13) provides an example of the complex relationships that govern changes in arrest rates. In this model, enforcement is represented by four variables; enforcement techniques, prosecution and adjudication, definition of DWI, and reporting and recording.



Some of these variables are, in turn, directly related to economic factors (e.g., enforcement budgets), and indirectly related to alcohol consumption and actual rates of DWI. Finally, some or all of the relationships in the figure may be reciprocal. Under these conditions, the linear model provides a hopelessly inadequate representation of the relationships among enforcement, arrests, and rates of DWI. This example, which could easily be redrawn for many common indicators of enforcement and compliance, illustrates the issue of embeddedness raised earlier.

Holder (1998) and Gruenewald et al. (1997) suggest that the Gordian knot represented in the figure (and in similarly complex relationships among variables) can be untied (or at least loosened) by an approach in which measurement and modeling interact. That is, systems models are used to suggest data collection and analysis strategies which, in turn, inform refinement of the models. This measurement approach is an outgrowth of the systems approach. Gruenewald et al. (1997) summarize the *synthetic* process this way:

The goals of the modeling strategy are (1) to represent adequately the theoretical structure of community systems that come to bear on alcohol problems; (2) to incorporate into these theoretical-structural representations estimates of effects obtained from empirical research; and (3) to mimic the dynamics of community systems underlying alcohol problems (p. 25).

That is, measurement approaches are guided by systems models which are, in turn, tested and refined using measurement data.

In the simplest application of a systems model to enforcement, two or more variables are used to model the system that outputs enforcement. Continuing with the example of arrest data, a simple model can be built on the premise that arrests reflect both rates of criminality and enforcement pressure (i.e., the effort expended by enforcement personnel in apprehending specific types of offenders, or, more usually, all offenders). As noted by Gruenewald et al. (1997), a measure of full-time law enforcement officers provides a measure of the effect of enforcement pressure on measured arrest rates. This measure,

while crude, offers increased precision in measuring enforcement (or crime) and is generally readily available.

In the alcohol area, a similar though more sophisticated approach was taken by Wagenaar and Wolfson (1994). These investigators modeled enforcement of underage drinking laws using:

1. Total population in state
2. Total number of law enforcement officers in state
3. Rate of officers per 100,000 population
4. Rates of serious crime arrests per 100,000 population
5. Rates of minor crime arrests per 100,000 population excluding liquor law arrests
6. Rates of narcotic arrests per 100,000 population
7. Rates of liquor law arrests among adults per 100,000 population
8. Rates of liquor law arrests among youth under age 21 per 100,000 population

The only analysis presented in their paper is a table of correlations among these variables. Modeling of enforcement using these predictors is mentioned but not discussed.

As noted previously, the major use of systems models in the social sciences has been in the area of macroeconomics. As discussed by Gruenewald et al. (1997), a theoretically driven or synthetic approach to model development has proven beneficial to both an understanding of the complex dynamics of economic systems and the limits of underlying theories of economic behavior. These approaches rely on a careful theoretic development of the processes assumed to underlie the outcome(s) of interest (e.g., performance of the U.S. economy), the indexing of these processes with available data from empirical analyses of the systems studied (e.g., retail activity, gross domestic product), and the simulation of the system to see whether the results of the model adequately mimic both qualitative and quantitative aspects of the systems studied.

To date, the most extensive applications of synthetic measurement to public health and public policy are *SimCom* (community alcohol problems) and *SimSmoke* (community tobacco problems) developed by Holder and his colleagues (Holder, 1998; Levy et al. 2001). Holder (1998) describes *SimCom* this way:

The *SimCom* model recreates the systems dynamics of a targeted community with regard to alcohol retail activity, alcohol consumption patterns, drinking and driving behavior, social norms, and regulatory controls. Published research findings, survey data, and results from secondary data analyses are used to define and mathematically specify relationships among variables within and across the model subsystems.

For example, DUI enforcement in *SimCom* is a function of staffing, the time and difficulty of making an arrest (reduced by special DUI patrols), and the priority placed on such enforcement. These variables are part of a larger-community Formal Regulation and Control subsystem which includes resources for enforcement as well as the laws and regulations themselves. Changes in laws (e.g., BAC levels or evidentiary standards) will affect the ease of procedures used for enforcement. The Formal Regulation and Control system, in turn, interacts with other subsystems, including those representing social

norms and retail alcohol sales. Based on actual or simulation data, *SimCom* outputs annual rates of DUI arrests and convictions, driver fatalities, crashes resulting in injury, and measures of mortality and morbidity. The concordance of these outputs with actual outcomes measures provides a test of the adequacy of the model (its variables and relationships among them).<sup>11</sup> A formal description of the modeling algorithms used in *SimCom* will be presented in a revision of this paper.

When a satisfactory level of concordance is achieved, models like *SimCom* and *SimSmoke* yield powerful policy planning and research tools for public health and public policy. They are arguably the strongest available approach to empirical study of health and social problems that arise as the complex interaction of individuals with economic, social, and cultural factors in the environment. In terms of enforcement (and other embedded constructs), these models provide empirically validated maps by which policy researchers can navigate between the many Scyllas and Charybdises that attend the measurement of variables in complex systems.

## CONCLUSION

The public health/public policy literature offers no magic bullets for measuring enforcement and compliance. However, a number of useful implications and directions can be identified. As the ultimate purpose of this review is to inform decisions related to APIS, we'll confine this discussion to alcohol policy and policy research.

### **Most Promising/Reliable/Valid Measures**

First, it seems clear that measurement of compliance is less complex than the measurement of enforcement. As noted, compliance is largely reflected in the behavior of individuals and organizations. Thus, linear (psychometric) measures of this behavior provide adequate indices when these measures are largely uncontaminated by extraneous system effects. We have identified random roadside breath tests, radar surveys, decoys, and self-report surveys as potentially useful compliance measures for alcohol policy research. Descriptive analysis also has been used to index compliance, although such studies are generally more expensive and labor intensive than linear measures.

Commonly used measures of enforcement in the public health/public policy literature fall into two broad categories. Perhaps most common are the descriptive measures that rely on synoptic analysis of enforcement activity. In general, these analyses yield categorical indices (e.g., enforced/not enforced, strong/weak) that allow comparisons among jurisdictions or between two time periods. However, qualitative indices can also be integrated into sophisticated time series designs, as discussed by Gruenewald (1997).

A more limited number of analyses rely on simple models of enforcement to derive indices. Wagenaar and Wolfson (1994) provide an example. These models take two or more variables (e.g., arrests and rate of officers per 100,000 population) and combine them into a single indicator (a simple ratio in the example just given). As noted, these

---

<sup>11</sup> Such a comparison for San Diego County, CA, for the period 1970–1990 yielded very good concordance between estimates generated by *SimCom* and actual rates of traffic fatalities (Holder, 1998).

approaches use available data (e.g., arrest rates, police person power) and are thus relatively efficient when compared to descriptive approaches. These analyses do not address the embeddedness of enforcement in complex and sometimes reciprocal relationships with other variables, a concern that is mitigated in cross-sectional studies.

Finally, Holder and his colleagues have undertaken the development and validation of complex systems models that include (rather than control or ignore) the embeddedness of enforcement in the large community alcohol system. These models follow a synthetic development process in which model-building and measurement are complementary processes that refine both the model and the data it uses. Holder (1998) presents evidence that his *SimCom* model can mimic reality over a 20-year period for important alcohol-related outcomes (e.g., traffic crashes). It thus may provide an algorithm for indexing enforcement that has both theoretical and empirical support.

### **Potential Utility to APIS**

We assume three desiderata for an enforcement/compliance component for APIS. First, measurement of enforcement and compliance should be applicable to a wide variety of policies. Second, available data should be used whenever possible. Third, the enforcement/compliance component of APIS should conform to the high scientific standards applied elsewhere in the project. Based on these criteria, the following implications emerge:

- Although the descriptive approach has considerable appeal as a research tool, it seems to be of limited utility to APIS because extant descriptive data cannot generally be applied in new studies. In addition, descriptive methods may be highly dependent on the specific purposes of any given study and the geographic region in which the study takes place. Some methods (e.g., those used by Saltz for observation of servers) might be of general interest. APIS could describe these methods and provide access to protocols.
- The linear measurement approach has broad applicability when appropriate conditions are met (direct measurement uncontaminated by extraneous system effects). However, like descriptive studies, extant data that meet the needs of APIS will be rare. It is straightforward to index youth compliance with underage drinking laws from a number of national studies (e.g., Monitoring the Future [MTF], the Household Survey). From the MTF, for example, one might have a measure of underage drinking rates that could be disaggregated by geographic region, grade, gender, college plans, parents' education and race (a sample MTF table for tobacco use appears below). The Household Survey would allow similar analyses by State. Beyond this, however, no national surveys we are aware of gather the kinds of data that have been used in individual studies (e.g., roadside breath tests, decoys). Thus, like the descriptive approach, the linear approach may offer limited value to APIS. Again, APIS could describe these methods and provide access to protocols.



**TABLE 1**  
**Long-Term Trends in Prevalence of Use of Cigarettes for Eighth, Tenth, and Twelfth Graders**

	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	'00-'01 change	
<b>Lifetime</b>																													
8th Grade																		44.0	45.2	45.3	46.1	46.4	49.2	47.3	45.7	44.1	40.5	36.6	-3.9ss
10th Grade																		55.1	53.5	56.3	56.9	57.6	61.2	60.2	57.7	57.6	55.1	52.8	-2.3s
12th Grade	73.6	75.4	75.7	75.3	74.0	71.0	71.0	70.1	70.6	69.7	68.8	67.6	67.2	66.4	65.7	64.4	63.1	61.8	61.9	62.0	64.2	63.5	65.4	65.3	64.6	62.5	61.0	-1.5	
<b>Thirty-Day</b>																													
8th Grade																		14.3	15.5	16.7	18.6	19.1	21.0	19.4	19.1	17.5	14.6	12.2	-2.5sss
10th Grade																		20.8	21.5	24.7	25.4	27.9	30.4	29.8	27.6	25.7	23.9	21.3	-2.6ss
12th Grade	36.7	38.8	38.4	36.7	34.4	30.5	29.4	30.0	30.3	29.3	30.1	29.6	29.4	28.7	28.6	29.4	28.3	27.8	29.9	31.2	33.5	34.0	36.5	35.1	34.6	31.4	29.5	-2.0	
<b>Daily</b>																													
8th Grade																		7.2	7.0	8.3	8.8	9.3	10.4	9.0	8.8	8.1	7.4	5.5	-1.9sss
10th Grade																		12.6	12.3	14.2	14.6	16.3	18.3	18.0	15.8	15.9	14.0	12.2	-1.8s
12th Grade	26.9	28.8	28.8	27.5	25.4	21.3	20.3	21.1	21.2	18.7	19.5	18.7	18.7	18.1	18.9	19.1	18.5	17.2	19.0	19.4	21.6	22.2	24.6	22.4	23.1	20.6	19.0	-1.6	
<b>1/2 pack- per day</b>																													
8th Grade																		3.1	2.9	3.5	3.6	3.4	4.3	3.5	3.6	3.3	2.8	2.3	-0.4
10th Grade																		6.5	6.0	7.0	7.6	8.3	9.4	8.6	7.9	7.6	6.2	5.5	-0.6
12th Grade	17.9	19.2	19.4	18.8	16.5	14.3	13.5	14.2	13.8	12.3	12.5	11.4	11.4	10.6	11.2	11.3	10.7	10.0	10.9	11.2	12.4	13.0	14.3	12.6	13.2	11.3	10.3	-1.0	
<b>Approx. Ns (in thousands)</b>																													
8th Grade																		17.5	18.6	18.3	17.3	17.5	17.8	18.6	18.1	16.7	16.7	16.2	
10th Grade																		14.8	14.8	15.3	15.8	17.0	15.6	15.5	15.0	13.6	14.3	14.0	
12th Grade	9.4	15.4	17.1	17.8	15.5	15.9	17.5	17.7	16.3	15.9	16.0	15.2	16.3	16.3	16.7	15.2	15.0	15.8	16.3	15.4	15.4	14.3	15.4	15.2	13.6	12.8	12.8		

NOTES: Level of significance of difference between the two most recent classes: s = .05, ss = .01, sss = .001.  
Any apparent inconsistency between the change estimate and the prevalence of use estimates for the two most recent classes is due to rounding error.  
SOURCE: The Monitoring the Future Study, The University of Michigan.

- The synthetic approach and the models that underlie it offer perhaps the greatest promise to APIS. Both relatively simple models such as that used by Wagenaar and Wolfson (1994) and the complex models used by Holder (1998) offer the possibility that researchers can be provided with empirically validated algorithms for indexing enforcement and possibly compliance. In the main, these models rely on secondary data, at least some of which could also be made available. Because the investigators cited in this paper do not provide detailed descriptions of the algorithms used, the feasibility of applying their models to APIS awaits further exploration.

## References

- Alfred, L.E., and Graham, A.K. *Introduction to Urban Dynamics*. Cambridge, MA: Wright-Allen Press, 1976.
- Almond, G. and Coleman, J. (eds.) *The Politics of the Developing Areas*. Princeton: Princeton University Press, 1960.
- American Academy of Pediatrics (RE0059). Condom use by adolescents. *Pediatrics* 107(6):1463–1469, 2001.
- Cohen, P., and Kennedy, F. Interorganizational Network Analysis: A Method for the Assessment of Coordination of Planning and Service Delivery in High Blood Pressure Programs. Prepared for National Conference on Blood Pressure Control, Washington, DC, April 1983.
- Edwards, G., Anderson, P., Babor, T.F., Casswell, S., Ferrence, R., Giesbrecht, N., Godfrey, C., Holder, H.D., Lemmens, P., Makela, K., Midanik, L.T., Norstrom, T., Osterberg, E., Romelsjo, A., Room, R., Simpura, J., and Skog, O.J. *Alcohol Policy and the Public Good*. New York: Oxford University Press, 1994.
- Forrester, J.W. *Industrial Dynamics*. Cambridge, MA: MIT Press, 1961.
- Forrester, J.W. *Principles of Systems*. Cambridge, MA: Wright-Allen Press, 1968.
- Gould, S.J. *The Structure of Evolutionary Theory*. Cambridge, MA: The Belknap Press of Harvard University Press, 2002.
- Grasmick, H., Bursik, R., and Arneklev, B. Reduction in drunk driving as a response to increased threats of shame, embarrassment, and legal sanctions. *Criminology* 31(1):41–67, 1993.
- Greer, A. The two cultures of biomedicine: Can there be consensus? *JAMA* 258(1):2739–2740, 1987.
- Greer A.L., and Greer, S: Medical technology decisions in hospitals: Changes since 1980. Final report to the Agency for Health Care Policy and Research, Grant HS06054. Milwaukee: Urban Research Center, University of Wisconsin-Milwaukee, 1991.
- Gruenewald, P.J., Ponicki, W.R., and Holder, H.D. The relationship of outlet densities to alcohol consumption: A time series cross-sectional analysis. *Alcoholism: Clinical and Experimental Research* 17(1): 38–47, 1993.
- Gruenewald, P.J. Analysis approaches to community evaluation. *Evaluation Review* 21:209–230, 1997.
- Gruenewald, P., Treno, A., Taff, G., and Klitzner, M. *Measuring Community Indicators: A Systems Approach to Drug and Alcohol Problems*. Thousand Oaks, CA: Sage Publications, 1997.
- Hanratty, M., Lightsey, R., and Larson, A. Open Systems and the Systems Engineering Process. DoD Discretionary Document, January 1999.  
<http://web2.deskbook.osd.mil/reflib/DDOD/028EV/028EVDOC.HTM>
- Hingson, R., McGovern, T., Howland, J., Heeren, T., Winter, M., and Zakocs, R. Reducing alcohol-impaired driving in Massachusetts: The Saving Lives Program. *AJPH* 86(6):91–797, 1996.

- Holder, H. Drinking, alcohol availability and injuries: A systems model of complex relationships. In: N. Giesbrecht, R. Gonzalez, M. Grant, E. Osterberg, R. Room, I. Rootman, L. Towle, Eds., *Drinking and Casualties: Accidents, Poisonings and Violence in an International Perspective*, New York, NY: Routledge, 1989, pp. 133–148.
- Holder, H. *Alcohol and the Community: A Systems Approach to Prevention*. Cambridge, UK: Cambridge University Press, 1998.
- Holder, H. Community prevention of alcohol problems. *Addictive Behavior* 25(6):843–859, 2000.
- Homel, R. *Policing and punishing the drinking driver: A study of general and specific deterrence*. New York, NY: Springer-Verlag, 1988.
- Huang, C.C. The impact of child support enforcement on non-marital and marital births: Does it differ by racial and age groups. *Social Service Review* 76:275–301, 2002.
- Klein, L.R. *Comparative Performance of U.S. Econometric Models*. New York: Oxford University Press, 1991.
- Klitzner, M., Bamberger, E., and Gruenewald, P. The assessment of parent-led prevention programs: A national descriptive survey. *Journal of Drug Education* 20(2), 1990.
- Klitzner, M. Prevention through reinventing environments. In: Hansen, W.B., Giles, S.M., and Fearnow-Kenny, M.D., eds. *Improving Prevention Effectiveness*. Greensboro, NC: Tanglewood Research, 2000.
- Levi-Strauss, C. *Structural Anthropology*. Tr. by C. Jacobson and B. Schoepf. New York: Basic Books, 1963 (original publication date, French language edition, 1958).
- Levin, G., Hirsch, G., and Roberts, E. Narcotics and the community: A system simulation. *AJPH* 62: 861–873, 1972.
- Levin, G., Roberts, E.B. and Hirsch, G.B. *The Persistent Poppy: A Computer-Aided Search for Heroin Policy*. Cambridge, MA: Ballinger Publishing Company, 1975.
- Levy, D., Friend, K., and Polishchuck, T. Effect of clean indoor air laws on smokers: The clean air module of the SimSmoke computer simulation model. *Tobacco Control* 10: 345–351, 2001.
- Lewin, K. Frontiers in group dynamics. *Human Relations* 1:2–38, 1947.
- Lucas, G.M. Directly observed therapy for the treatment of HIV: Promises and pitfalls. *The Hopkins HIV Report Archives*, July 2001.
- MacRae, D., Jr. *Policy Indicators: Links Between Social Science and Public Debate*. Chapel Hill: University of North Carolina Press, 1985.
- Maltz, M.D. *Recidivism*. New York: Academic, 1984.
- Mechanic, D. *Medical Sociology, Second Edition*. New York: The Free Press, 1978.
- Mitroff, I.I., and Sagasti, F. Epistemology as general systems theory: An approach to the design of complex decision-making experiments. *Philosophy of Social Science* 3:117–134, 1973.
- Morrissey, E.R. Methodologic issues in the study of special populations. In: Marc A. Schuckit, Ed., *Alcohol Patterns and Problems*, New Brunswick, NJ: Rutgers University Press, 1985. pp. 79–111.

- Moskowitz, J. The primary prevention of alcohol problems: A critical review of the literature. *Journal of Studies on Alcohol* 1:54–88, 1989.
- Pressman, J., and Wildavsky, A. *Implementation, Second Edition*. Berkeley, CA: University of California Press, 1973.
- Rigotti, N.A., McKool, K.M., and Schiffman, S. Predictors of smoking cessation after coronary artery bypass graft surgery: Results of a randomized trial with 5-year follow-up. *Annals of Internal Medicine* 120:287–293, 1994.
- Ross, H.L. *Deterring the Drinking Driver: Legal Policy and Social Control*. Lexington, MA: D.C. Heath, 1982.
- Ross, L. Social control through deterrence: Drinking and driving laws. *Annual Review of Sociology*, 10:21-35, 1984.
- Saltz, R., and Stangherra, P. A community-wide responsible beverage service program in three communities: Early findings. *Addiction* 92 (supplement 2) S237–S249, 1997.
- Saronson, S. *The Psychological Sense of Community: Prospects for a Community Psychology*. San Francisco, CA: Josey-Bass, Inc., 1974.
- Sellin, T. The basis of a crime index. *Journal of Criminal Law and Criminology* 346, 1991.
- Sellin, T., and Wolfgang, M.E. *The Measurement of Delinquency*. New York: John Wiley, 1964.
- Smith, E.A., and Winterhalder, B. *Evolutionary Ecology and Human Behavior*. New York: Aldine de Gruyter, 1992.
- Stoneall, L. *Country Life, City Life: Five Theories of Community*. New York: Praeger Scientific, 1983.
- Tuma, N.B., and Hannan, M.T. *Social Dynamics: Models and Methods*. New York: Academic, 1984.
- Wagenaar, A.C., and Wolfson, M. Deterring sales and provision of alcohol to minors: A study of enforcement in 295 counties in four states. *Public Health Reports* 100(4): 419–427, 1994.
- Watts, R.K., and Rabow, J. Alcohol availability and alcohol-related problems in 213 California cities. *Alcoholism: Clinical and Experimental Research* 7(1): 47–58, 1983.