The Importance of Context in Understanding Behavior and Promoting Health

Edwin B. Fisher, Ph.D.

Published online: 16 February 2008
© The Society of Behavioral Medicine 2008

Abstract Behavior, the roles of behavior in health, health promotion, health, quality of life, and death are all context-dependent. This paper begins with a review of behavioral and ecological models, emphasizing their shared emphasis on context. It then turns to genetics and the importance of contexts in understanding genetic influences. Jumping from genes to geography, it examines how spatial analysis provides both a model and framework for considering contextual influence. Continuing with analytic models, it returns to genetics, and considers how it provide models for integrating our understanding of broad social and community influences. The paper extends this thinking through multilevel analysis and proposes “analytic multilevel designs” as a way of studying “context focused interventions” (as opposed to context independent interventions for which conventional experimental designs are often well-suited). It closes with reflections on ways in which we cultivate and extend our knowledge base and on the intellectual contexts of positivism and postmodernism that surround behavioral and ecological thinking.

Keywords Behavior · Behavioral and ecological models · Contexts · Evaluation · Genetic influences · Health promotion

Introduction

Whether our lives are directed by events around us or events within us, “not in our stars, but in ourselves,” is of concern in ethics, aesthetics, law, religion (the Old Testament of laws and the New Testament of “faith as a grain of mustard seed” within us), and, of course, behavioral science, biology, and health. Without taking sides on this enduring question, the present paper does assert the importance of contexts in health and health behavior and examines how we conceptualize context amidst the many other influences on our actions.

Roots of Behavioral Medicine in Behaviorism and Behavior Therapy

An important base for the development of behavioral medicine was the intellectual tradition of behaviorism in

1 Shakespeare, Julius Caesar, Act 1, Scene 2.
western psychology of the 20th century. The first meeting of the Society of Behavioral Medicine was in conjunction with a meeting of the Association for the Advancement of Behavior Therapy, developed in the 1960s to advance what was at that time a radical departure from conventional psychotherapeutic perspectives. Early leaders in behavioral medicine such as John Paul Brady and Stuart Agras were also leaders in behavior therapy. Neil Miller, a major figure in twentieth-century behaviorism and the application of behaviorism to complex behavior and psychotherapy [1] was also a major figure in the founding of behavioral medicine.

Richard Stuart, developer of some of the first behavioral interventions to counter the assumption that weight loss was not possible without resolving supposed underlying personality disorders, and G. Terrence Wilson, long a major contributor to research on weight loss and eating disorders, were both presidents of the Association for the Advancement of Behavior Therapy. Biofeedback [2], stress management interventions featuring relaxation and active coping with stressors [3–6], group and individual interventions for weight management or smoking cessation [7, 8], all had their roots in behavior therapy and its development of self control, and self-management interventions in the 1960s [9–11].

Past Experience and Behavior

A common misconception about behaviorism is that it views the actions of the individual as responses only to current stimuli. This fails to recognize the fundamental point that behavior is learned, that past experience guides present behavior. Thus behaviorism is fundamentally historical, fundamentally directed to how the aggregate of past experience brought to the current situation governs the impact of the present.

That behaviorism emphasizes the guidance of present behavior by past experience does not mean it views these relationships as simple. Much of behavioral research over the past 50 years has focused on using a rigorous behavioral perspective to understand complex behavior, for example, choice and self control [12, 13]. A good example of complexity in the influence of past experience on present behavior lies in individual differences in responses to ambiguous stimuli.3 For example, Edith Chen et al. [14] have examined adolescents’ responses to ambiguously threatening videos such as one in which a teacher raises anxiety about suspected cheating in a class and then, without explaining why, asks one of the students to stay after class. Serving as a marker of previous exposure to unpredictable, threatening and stressful experiences, low socioeconomic status was associated with greater likelihood of viewing ambiguous videos as threatening. Perceptions of threat in ambiguous stimuli were also associated with higher night-time heart rate [15] and served as a mediator between socioeconomic status and heart rate as well as other cardiovascular risk indicators such as blood pressure [16]. Differences in perceptions of ambiguous stimuli reflect differences in the aggregate of experiences associated with low socioeconomic status and previous exposure to unpredictable, threatening, and stressful experiences.

B. F. Skinner, Clark Hull, and other behaviorists emphasized as critical to the evolution of *homo sapiens* this propensity to generalize from one stimulus to another, from one response to another response [17, 18]. Although never told by their teacher that they were under suspicion after their class was scolded about a serious cheating incident, the adolescents in Chen’s studies, responded differentially according to their past experience. This illustrates *homo sapiens’* key adaptive capacity of generalization. Along with a complementary ability to discriminate among settings and behaviors as differential consequences dictate, generalization tunes our behavior exquisitely to situational nuance. Our behavior reflects our contexts.

Articulating a Broad View of Experience and Environment—Ecological Perspectives

Ecological perspectives have gained substantial popularity in public health and much literature seeking to address population approaches to health [19–21]. In ecological approaches, the behavior of the individual is viewed as guided by layers of influences including the family, proximal social influences such as social networks or neighborhoods, organizational influences such as worksite or community systems or healthcare systems, and larger social influences such as government, policy, or large economic structures. Different models may specify different layers of influence and different components of each, but they share two important emphases: (1) that the behavior of the individual reflects the influence of all the layers; and (2) that the layers interact in their influence so that, e.g., communities may influence families but families may also influence communities [20].

There is a fundamental congruence between behaviorism’s emphasis on individual experience, and ecological perspectives’ articulation of the social and organizational layers that are the architecture of that experience. As in Chen’s research examining exposure to multiple stressors and reactions to ambiguous threatening stimuli [14], both
behaviorism and ecological perspectives share an emphasis on how experience across a wide array of settings shapes behavior. Ecological perspectives add attention to how organizational and policy levels provide structure to the more immediate settings—the stimuli and reinforcers—that organize individual experience. Both also share the assumption that the same approaches to analyzing individual behavior can be applied to the behavior of larger social units such as groups or organizations and the influences of their environments.

It is interesting to note that both behaviorism and ecological perspectives share a sensitivity to blaming the victim in analyses of individual behavior or behavior problems. In the war of nostrums of former “First Ladies” (and one presidential candidate), both behaviorism and ecological perspectives would rail against “Just say no” and embrace “It takes a village”! Both would look to socioeconomic, cultural, family, and individual experiences among other background factors in seeking to explain important health behaviors such as smoking, obesity, type A behavior, impulsiveness, etc.

Mistaken Opposition in Our Ranks—Congruence of Behaviorism and Ecological Perspectives

Given the broad compatibility between our behavioral roots and our ecological views, it is troubling to this unreconstructed behaviorist to see that many who cheerfully embrace ecological perspectives persist in 1960s-era rejection of behaviorism as mechanistic, treating individuals as objects. Science and Human Behavior [18] was B.F. Skinner’s effort to articulate an extension of fundamental principles of behaviorism to understanding large patterns of behavior such as self control, religion, and the behavior of groups. For example, in talking about how environmental influences shape the god/father metaphor in religious practice, Skinner argued that the emphasis on god/father as source of protection and forgiveness reflected reinforcement contingencies operating on religious organizations. The financial support and loyalty of adherents to a religion reinforces its development of effective metaphors for gaining influence over those adherents. The metaphor of god/father persists because it works, because this behavior of the organized religion is reinforced by the support of the adherents it draws. Without using the term “ecological,” Skinner clearly wrote in its spirit, showing how events at different ecological levels might reinforce each other.

Skinner’s provocatively titled Beyond Freedom and Dignity [22] addressed how views of “freedom” and “dignity” as intrinsic human characteristics provide the grounds for policies that hold the individual responsible for her or his behavior, an attributive tendency much akin to victim blaming. At the same time, the myth of intrinsic individual freedom distracts our institutions, government, and culture from more effective approaches to promoting healthy and adaptive behavior. From the perspective of developing interventions to advance human progress and quality of life, the key issue is not treating individuals as though they are free and autonomous, but providing interventions and engineering environments that will promote the skills and behaviors that end up making people feel free and possessed of dignity.

One of my mentors at Stony Brook, Leonard Krasner, is an example of the historical breadth of behavioral approaches and their extension to ecological perspectives. The “token economies” that he and others developed for chronic mental patients in the 1960s reinforced useful activities with tokens that were exchangeable for desired food, private rooms, more attractive clothes than those routinely issued, etc. [23] Krasner also published key papers showing that verbal behavior such as using plural nouns or first person pronouns could be shaped by the attention and responsiveness of the interlocutor, invading with the behavioral wedge the freedom and dignity of self and self presentation [24]. But Krasner went on to anticipate ecological models, professing to his students of the late 1960s and thereafter that behavior modification entailed “environmental design” as he extended his own attention to approaches such as open classrooms and his own writing to books such as Environmental Design and Human Behavior [25]. It included chapters on topics such as “Community Mental Health and Environmental Design” and “Environmental Design in Alternative Societies: The People’s Republic of China.” Even more broadly, his Behavior Influence and Personality: the Social Matrix of Human Action [26], written with his long-time collaborator, Leonard Ullmann, attempted to use behavioral and social influence perspectives to integrate our understanding of personality, psychopathology, behavior change, and psychotherapy. Krasner and Ullmann argued that models of psychopathology, themselves, reflect the social influences of professional and science establishments and contingencies that surround them. The course and breadth of his work, from verbal conditioning of plural pronouns and token reinforcers for making one’s bed in a mental hospital ward, to cultural influence in the People’s Republic, illustrates the many links and compatibilities between behavioral and ecological stances.

There is a fundamental compatibility between behavioral perspectives, so important in the origins of behavioral medicine, and ecological perspectives that have achieved prominence in public health. Both share an emphasis on contexts, be they termed “stimuli”, “setting events”, “experience”, “learning history”, “ecological layers”, or “social, cultural, community, economic, and policy influences”. 
Contexts and Disparities

Contexts play a large role in health disparities. Examples abound. In the course of my lifetime, smoking in the United States has evolved from a privilege of the well-to-do to a problem among those who are poorly educated, poorly paid, and/or burdened by a variety of personal and psychological problems such as depression, schizophrenia, or divorce [27]. Still in the United States, African Americans, Latinos/as, and American Indians are about twice as likely to have diabetes as the rest of the population. Internationally, infectious diseases, especially HIV/AIDS, are much more prevalent in poor nations and, within all nations, among poor people. The impact of HIV/AIDS in many countries in Africa threatens the core of their social fabric. Kelly Brownel has documented the socioeconomic structuring of obesity both within the U.S. and globally and the contributions to the obesity epidemic of the production, marketing, and drawing profit from the sale of food [28].

Through all of these, we in behavioral medicine see the hand of context and environment. An important illustration of this lies in analyses of ethnic differences in health burden. In many studies, the pattern is the same, ethnic differences are explained away or markedly reduced by economic and educational factors when these are included in statistical models. For example, in Edith Chen’s work, described earlier, a significant contribution of ethnicity to perception of threat becomes nonsignificant when a child’s socioeconomic status is included in the model.

Cross-national analyses support the view that disparities in health reflect variability in socioeconomic characteristics of countries [29]. Michael Marmot’s analysis of this global variability in health extends, however, beyond socioeconomic contexts per se. For example, the populations of the USA, Greece, Costa Rica, and Cuba have life expectancies ranging from 76.5 years (Cuba) to 78.1 years (Greece). However, their GNPs in US dollars range much more widely, from <$10,000 per person (Cuba and Costa Rica) to $34,000 (US). Marmot interprets such data as indicating that, along with income poverty and material conditions, social determinants must also play roles in the development of health risks and the paths of infectious disease transmission. Reporting as chair of the Commission on Social Determinants of Health of the WHO, Marmot notes that key social determinants include stress, early life circumstances, social exclusion, work, unemployment, social support, and addiction [30]. If obesity and other risks such as smoking and hypercholesterolemia and hypertension are the causes of noncommunicable diseases, then social determinants are among the “causes of the causes,” attention to which is likely to reduce population disease burden.

In keeping with ecological perspectives [31], the understanding of psychosocial and stress factors rests also on other contextual factors such as economic status, social and cultural factors, spatially distributed variables, or organizational and political factors. From an ecological perspective’s emphasis on influences among ecological layers, economic factors, for example, may have direct influences on a variety of health behaviors and health indicators but may also interact with other cultural, political, and organizational factors in health impacts mediated by psychosocial distress.

Genetics and Complexities Among Influences on Behavior

Many think of genes as causes that obviate other influences on behavior. Old controversies as to whether one or another disease, e.g., schizophrenia, is either genetic or learned presumed that the one trumps the other. The reality is that genetic, other biological, behavioral, and environmental variables interact in complex ways to lead to behaviors and health states [32]. A case in point lies in the relationships among serotonin, psychological factors such as hostility and stress, and cardiovascular risk and disease. As contributed to and summarized by Redford Williams et al. [33, 34], there is considerable evidence that low levels of serotonin activity in the central nervous system (CNS serotonin) are related to hostility, stress, depression, and related negative emotions that are, in turn, related to heightened risks of cardiovascular disease (CVD). Furthermore, long and short alleles of the serotonin transporter gene promoter polymorphism appear to affect CNS serotonin activity in ways that can impact CVD risk. Persons with one or two long alleles, for example, exhibit altered levels of CNS serotonin function, and greater blood pressure responses to stress. They also appear to be at greater risk of developing coronary heart disease.

Environment also guides the function of the serotonin transporter gene. Among Rhesus monkeys reared by their parents, there is no difference in CNS serotonin levels between those with long and short alleles. However, among those reared among peers, the short allele is associated with reduced CNS serotonin [35].

Figure 1 summarizes observations around CNS serotonin and CVD risk. Genes clearly play an important role, but their impact may also be governed by environmental factors such as rearing conditions. Furthermore, the model suggests how socioeconomic and social factors may influence these processes. For example, overstressed parents or neighborhood crime may be analogous in humans to the peer vs parental rearing that moderates gene expression in monkeys. As also described in Fig. 1, there are several contextual factors that influence the pathway from genotype to CVD risk. The prevalence of long-allele genotypes varies by country of origin, from less than 30% in China and Japan to over 70%
among populations originating in Africa [36]. Also, substantial literature indicates influences of socio-economic status on the negative emotions that appear to raise CVD risk [37].

Michael Meany has carried out an important series of studies showing that maternal rearing conditions (e.g., frequency of rat dams licking their pups) influence expression of genes related to stress response in adults. He has taken this work as far as being able to identify “epimutations”, specific changes in methylation of cytosines on genes that mediate the relationship between rearing and adult stress response [38]. Pursuing this line of thinking, Meany has argued (comments at International Congress of Behavioral Medicine, Bangkok, December 2006) that within a cell, the expression of the cell’s genes is dependent on the environment of the cell that is the phenotype of those genes. Thus, in an important way, phenotype precedes genotypic expression. At the level of the cell, then, one can see a very close interdependence between gene and environment that sets a model for gene X environment interactions at the levels of whole animals and populations.

The impact of context is underscored further by a perplexing aspect of the research on long and short alleles of the serotonin transporter gene. As noted above, the work of Williams et al. [33] indicates that the long alleles are the “bad actors”; those with one or two long alleles have significantly greater blood pressure responses to stress and greater CVD risk. However, in a longitudinal study of depression among young adults, the number of short alleles (either one or two) was related to greater likelihoods of depression and suicidality [39]. In other studies of Williams et al. [40], whether the bad actor is the long or short allele appears to vary by sex and geographic population origin. This variability in the contributions of the long and short alleles may reflect gene–gene and gene–population interactions, both extending the central point here, that genetic expression is context dependent.

What Meany points out at the level of the cell is parallel to what others have called “reciprocal determinism” [41] in the relationships between human behavior and its environmental surround. Just as the cell phenotype acts as an environment that influences the expression of the cell’s genetic material and the further emergence of the cell’s phenotype, so our environment governs our actions which create important parts of the environment that will govern our next actions. Continuing up the ladder of complexity, one can see the same kind of pattern in the influence of:

- the group on the individual and the individual on the group
- the organization on the division and the division on the organization
- policies on organizations and organizations on policies.

This pattern of reciprocal influence of surround on agent and of agent on surround appears an important dynamic across living systems. It poses an important counterpoint to more primitive models such as those which get lost in debate over whether genes or environment are important, models in which there can only be one cause and a single thing can be only a cause or an effect but not both.

It is worth noting that, statistically, interactions are even handed. Whether we graph the influence of A as dependent on the presence/absence of B or the influence of B as dependent on the presence/absence of A, it is the same interaction. Going back to the work of Meany on maternal rearing and gene expression, one could say that the expression of genes in adult stress response is dependent on maternal rearing, or one could say that the impact of maternal rearing on adult stress response is dependent on particular genes. Both refer to the same data. As another example, Pima Indians in the US show “the highest prevalence of type 2 diabetes mellitus … of any population in the world” [42]. Yet, Pimas living in Mexico have relatively low levels of diabetes. Ample evidence links genetics to diabetes within the Pima population [42]. Thus, the relationships among genes, environment, and diabetes among the Pimas can be stated in either of two ways:

- Genetic factors associated with membership in the Pima population have a strong influence on prevalence of diabetes among a population exposed to the obesigenic environment of US diet and food distribution
- The obesigenic environment of the US has a strong influence on prevalence of diabetes among a population genetically predisposed to high rates of diabetes

Further consideration of gene–environment interactions leads to recognition of additional ways in which context governs events. Assume that the SBM allele leads to interest in behavior and environment (IBE) and, in the absence of this allele, there is no IBE. Assume also that being raised by parents who own Birkenstocks (PB) leads to IBE and, in the absence of PB, there is no IBE. Let us
make one further simplifying assumption. Ignoring the literature on gene–gene interaction, let’s assume that the SBM allele and PB are independent of each other.

If about half of homo sapiens have the SBM allele and about half are exposed to PB, then about one quarter (0.5 × 0.5 = .25) will show IBE. Crudely, both the SBM allele and PB will contribute about 50% of the variance in IBE. Suppose, however, that the presence of the SBM allele were 100% in a group. In that case, the expression of IBE would become identical to the prevalence of PB and PB would account for 100% of the variance in IBE. Reciprocally, if PB were 100% prevalent, the expression of IBE would become identical to the prevalence of the SBM allele. If instead of 100%, the prevalence of PB fell to zero, then the presence/absence of the SBM allele would have no effect and IBE expression would be zero. So what’s conditional on what? It seems to depend on the base rates of the factors being considered. The proportion of variance attributable to environment depends, in part, on the base rates of critical genes. Reciprocally, the proportion attributable to genes depends in part on the prevalence of critical environmental factors. The gene–environment interaction is, itself, context dependent. The wide variation in the prevalence of the long allele of the serotonin transporter gene, from <30% among those of Chinese descent to >70% among populations originating in Africa [36] makes clear the issue is not unimportant.

**Spatial Analysis and the Legacy of Lewin**

Generations of students in behavioral and social sciences have been introduced to the work of Kurt Lewin, best know for $B=f(I×E)$—behavior is an interactive function of the individual and the environment. For years, we all read this, wrote wonderful paragraphs in our exams on the interplay between the individual and the environment, and then went merrily on our way pursuing research examining individual and, occasionally, family factors as they influence behavior, with, perhaps, socioeconomic or demographic variables “thrown in”, but usually only to be “controlled away”.

The advent of spatial analysis in recent years has excited the inclination to examine how the environment influences behavior. As Tip O’Neill noted that “All politics is local”, we might also ponder whether all behavior is spatial. Spatial analysis allows us to analyze powerfully the effects of geographically arrayed influences such as those associated with census tract location or proximity to services and resources. For example, before limitations on some cigarette advertising over the past decade, Douglas Luke et al. (43) showed sharp stratification of location of tobacco billboards in St. Louis, Missouri, much more likely in poor and African American neighborhoods than in White and middle class neighborhoods (see Fig. 2).

Similarly, Elizabeth Baker et al. [44] examined the distribution of supermarkets and fast food restaurants, also in St. Louis. Supermarkets were audited and sorted into tertiles according to their offering fresh fruits and vegetables and lean, low-fat, and fat-free meat, poultry and dairy products. Of 21 supermarkets in census tracts with greater than 75% African American population, none were in the highest tertile, whereas, in census tracts with less than 10% below the poverty level and more than 75% white population, 17 of 30 (57%) were in the top tertile.

Do neighborhood resources make a difference? Morland et al. [45] examined the relationships among obesity and supermarkets and convenience stores in neighborhoods. After adjusting for gender, race, age, income, education, and physical activity, it turns out the presence of supermarkets in a census tract is associated with a lower prevalence of obesity (prevalence ratio = 0.83 relative to census tracts with no supermarkets) while the prevalence of convenience stores was associated with a higher prevalence of obesity (prevalence ratio = 1.16 relative to neighborhoods with no convenience stores). Those in census tracts with only convenience stores were 1.45 times as likely to be obese as those in tracts with only supermarkets.

The availability through spatial analysis of data reflecting census tract characteristics and location of individuals’ relative to community resources such as supermarkets, mass transit, walking trails, indoor malls with walking programs, or recreation centers raises the possibility of combining such data with those from individuals. For example, we might wonder whether individual levels of hostility or depression are related to neighborhood levels of crime or economic indicators. In analyzing interventions, we might wonder, for example, whether an intervention promoting healthy eating is more effective for those living in neighborhoods with accessible sources of healthy foods as opposed to those whose neighborhoods contain only convenience food outlets.

**Multilevel Analysis**

Multilevel analyses allow integration of indicators from a variety of domains such as spatial, neighborhood, organization, or family as well as the individual level. Consider the work of my colleagues at the University of North Carolina at Chapel Hill, Karl Bauman, Susan Ennett, Vangie Foshee, and Ying-Chih Chuang 4 regarding teens’ use of alcohol and tobacco. They have shown complex paths among:

- neighborhood-level variables: neighborhood socioeconomic status

---

4 Now at Graduate Institute of Public Health, Taipei Medical University, Taiwan.
• family-level variables: parent drug use and parental monitoring of teen drug use
• peer influences: peer drug use
• individual level behaviors: use of alcohol and tobacco

A simplified version of results from one of their studies is in Fig. 3. In Path A, low socioeconomic status of neighborhood was related to higher levels of parental monitoring which was, in turn, related to lower levels of alcohol use among adolescents. In Path B, however, low-SES neighborhoods had higher levels of peer drinking which was associated with greater adolescent drinking. So, the same neighborhood characteristic, low SES, was associated with one path through parental monitoring that was linked to lower drinking and with another path through peer drinking that was associated with greater drinking. In contrast to low neighborhood SES, high neighborhood SES was associated, in Path C, with higher levels of parent drinking which was associated with greater adolescent drinking [46].

Another example of this way of thinking lies in work examining type of social support within the context of social networks and these, in turn, in the context of...
neighborhood climate or social capital, carried out with my colleagues Jeanne Gabriele, Mark Walker, and Joan Heins. Adults from St. Louis (n=301, 76.7% female, 52.2% African American, 76% educated beyond high school) completed measures of:

(a) Nondirective support (cooperating without “taking over,” accepting choices and feelings, e.g., “wow, what a bummer”) and Directive support (taking control, prescribing correct choices and feelings, e.g., “you’ve just got to look at the half of the glass that’s full”)

(b) social network and integration [47]

(c) positive and negative neighborhood climate (e.g., “If you fell on the sidewalk or street in your neighborhood, would people help you?” “Do you see people in angry arguments in your neighborhood?”), and

(d) cynical mistrust.

As can be seen in Fig. 4, structural equation modeling showed a path from positive neighborhood climate through social network and Nondirective support to lower levels of cynical mistrust. A second path links poorer neighborhood climate to higher levels of Directive support and, in turn, greater cynical mistrust. At the same time, each of neighborhood climate and social network retain independent links to cynical mistrust. Rather than the several social variables—climate, network, and type of support—collapsing in an amorphous statistical mush, these results illustrate both distinct influences of each and complex layering among them.

### Genetics as a Model for Analyzing Contexts and Other Influences on Behavior

In genomics, causal relationships are inferred through cluster analysis and related statistical techniques that compare differences in probabilities of hundreds or even thousands [48] of genes among those with varied phenotypes. As an example, Fig. 5 shows gene arrays characterizing women with poor or good “signatures” for likelihood of subsequent metastases following incident breast cancer [49]. In such analyses, no one gene is the cause or indicator of the phenotype. Instead, the relationship between phenotype and all the genetic markers in the analysis is probabilistic, not all or none.

This approach to characterizing genetic influences is descriptive but persuasive as to the likely causal relationship between profiles and outcomes. To what extent does it provide a model for making judgments about causal influences in a multilevel approach to complex behavior, such as might be arrayed by genetic, personal, social, organizational, and geographic influences?

Michael Gibbons [50] has suggested the metaphor of genomics as a base for populomics. From the perspective of the individual, we can envision complex webs of influence including genetic and other individual characteristics as well as, outside the individual, the ecologic layering of family, neighborhood, community, worksite, government and policy, all arrayed in a spatial analysis. These multi-level complexes could be examined as they explain, for example, likelihood of smoking and its relationship with...
rates of cardiovascular disease and cancer, or BMI and its relationship with diabetes, obesity, and other related diseases.

**Implications for Evaluation of Context Focused Interventions**

Experimental designs *isolate* influences from their contexts and so enable inferences that the experimental manipulation, itself, is responsible for observed changes or outcomes. Much intervention research is intended to show that interventions are context independent or robust amidst varying contexts. For example, a surgical procedure is, ideally, independent of the behavioral and social contexts that surround it. However, this feature of experimental designs poses challenges for evaluating “context focused” interventions that are intended to interact with or adapt to their contexts.

A key point of difficulty in using experimental designs to evaluate context focused interventions is the imperative that experimental variables be standardized. To all who have studied experimental design, this seems an unassailable, bedrock assumption. However, let us consider it further. How does one standardize context focused interventions that, in the “real world,” include substantial adjustment or tailoring to individual, provider, setting, or community factors? This is especially a problem for community organization approaches to health promotion that stress development of interventions according to the perspectives and with the involvement and active direction of community members [51]. That an intervention might be “whatever the community decides to do” evokes horror among those of us schooled to insist on standardization.

The COMMIT trial of community interventions in promoting nonsmoking was designed to test a community approach in a rigorous, experimental manner. Each of eleven pairs of communities were randomized to receive the COMMIT community intervention or usual care. The results of COMMIT were considered disappointing. Light to moderate smokers were more likely to quit in the COMMIT intervention communities. However, the major hypothesized outcome was impact on heavy smokers and this was not significantly greater in the intervention communities [52, 53].

An editorial that accompanied the publication of the COMMIT results in the *American Journal of Public Health* discussed “the tribulations of trials” [54]. Noting the secular trends in smoking and the overall demonstrable impacts of a complex of informational, social, community, and economic factors driving reductions of smoking, it reframed the failure of COMMIT not as a failure of community interventions *per se* but as a failure of our methodologies to capture what broader, epidemiologic research indicates are clearly powerful influences of community, social, mass media, and related influences on health behavior.

One way of considering the disappointing results of COMMIT is in terms of its approach to standardizing the intervention. A core protocol specified elements to be implemented in each of the intervention communities. This specification was extensive, totaling 167 mandated activities in each community, for example, 31 for each worksite and 20 for public media [55]. Standardization was underscored by identification of intervention approaches that were not allowed, enforced in several communities that were prevented from implementing approaches they had favored.

The standardized community intervention in COMMIT contrasts with the tradition of community organization approaches such as those of Jack Rothman and Saul Alinsky popularized in the 1950s and 1960s and embraced by many in public health such as Meredith Minkler, Larry Green, and my colleagues at the University of North Carolina at Chapel Hill, Eugenia Eng, Alice Ammerman, Marci Campbell, John Hatch, and Laura Linnan. These community organization approaches focus on engaging communities in setting priorities for their programs, planning interventions, identifying collaborators, or setting the content of interventions. COMMIT did not test such an approach.

Suppose that instead of the *content* of the community organization program, COMMIT had standardized the process of community organization taken in the 11 experimental communities. This would have required identifying key features of how community based programs addressed a number of challenges e.g., types of coalitions, whether or not to have a coalition, approaches to audience involvement, roles and authority of community members, rules for setting program objectives and identifying program strategies, etc. In this way, a well-specified community organization approach (albeit with varying smoking cessation program content) might have been compared to usual care and monitoring. If successful, the lesson learned would be that the community organization process—not any specific smoking cessation tactic employed—is useful in promoting smoking cessation.

The report of the main results of COMMIT concluded, “Based on sound principles of experimental design, COMMIT allowed a rigorous evaluation of its community-based intervention” [52] (p. 190) COMMIT was not wrong. It did test one type of community program, one that was centrally planned and locally implemented, but it surely did not test community approaches in general or, in particular, those that emphasize context focused community organization and engagement processes as the base for specific interventions [56].

---

5 If you ever get there, please send postcard to Ed Fisher, Box 7440, University of North Carolina, Chapel Hill, NC 27599.
At its base, standardizing or operationalizing variables requires judgment as to the key dimensions by which concepts, principles or key intervention functions should be specified. When the American Lung Association’s® group smoking cessation program, Freedom from Smoking®, was first being disseminated, an eager participant in one training session mentioned in passing that, in their setting, the program would need to be implemented in eight sessions. One of my colleagues on the training staff cautioned, “You can’t do that. It’s only been validated as a seven-session class.”

Wise people may differ as to whether reorganizing a curriculum as an eight- rather than seven-session class would have constituted a violation of the validated protocol. Nevertheless we must recognize boundary conditions of operationalizations. How broadly should they be drawn? We probably would all agree that one study site printing a manual on pale green paper and another printing the same manual on pale blue paper would not constitute a meaningful violation of standardization. However, we would all probably respond with alarm to a weight loss intervention that was operationalized in one site as a series of lectures about the food pyramid and, in another, as dream interpretation of emotional factors related to eating. Identifying the key dimensions of operationalization requires careful judgment and cannot be decided simply by following rules of research design.

One approach to evaluating context focused interventions is to differentiate their key functions from the specific ways in which those functions may be addressed in a particular setting. In organizational behavior, this reflects “equifinality,” recognition that different practices can achieve similar effects [57]. In the Diabetes Initiative of the Robert Wood Johnson Foundation with which I have had the privilege of being associated, we faced the challenge of encouraging self management programs in 14 different primary care and community settings around the U.S. [58] To do this, we identified key “Resources and Supports for Self Management”: individualized assessment, collaborative goal setting, opportunities to learn skills, ongoing follow-up and support, community resources such as for physical activity or buying healthy food, and continuity of quality clinical care [59]. We then encouraged programs to be wide ranging and creative in how they addressed these resources and supports. Evaluation can then focus on relationships between outcomes and the extent to which key functions (e.g., the Resources and Supports for Self Management) are addressed or implemented in each setting.

**An Alternative—Analytic Multilevel Designs**

An approach that may have advantages for (a) evaluating context-focused, complex interventions, and allowing considerable variability in their implementation, while (b) analyzing their interactions with context variables is through analyses that draw inferential power from their numbers, measurement of key variable, and sophisticated analytic techniques such as multilevel analysis, structural equation modeling, and spatial analysis.

In its eleven pairs of experimental and control communities, COMMIT collected longitudinal data from over 20,000 individuals. Imagine that street addresses of all of these were available to support spatial analysis and inclusion of block level census data in evaluations along with all the individual data routinely gathered in such studies. Imagine that communities were provided resources to implement programs and given considerable latitude in how they pursued this but were required to maintain careful records of activities implemented and, where appropriate, their location. Imagine that those monitored in the study were provided modest incentives for quarterly or half yearly reports of their exposure to intervention elements and for providing a sample of blood to be stored for subsequent analysis of genotypes found related to smoking or difficulty quitting. Such a design would support an enormously powerful evaluation integrating: (a) socioeconomic data at the individual and neighborhood level; (b) neighborhood and community characteristics such as presence of neighborhood based resources or size of media markets, (c) characteristics of organizations implementing intervention, (d) level of intervention implementation and type of intervention implemented; (e) individual characteristics; (f) extent to which individuals reported engaging in or being exposed to interventions; (g) history of smoking over the course of the program, and, perhaps (h) candidate genes related to smoking and/or difficulty quitting. With the large numbers involved, quantification of intervention and exposure and their relationships with outcomes could be carefully controlled for a whole host of behavioral, sociodemographic, community, and even genetic variables in longitudinal study rendering strong predictive relationships of very formidable inferential weight. Figure 6 portrays an imagined simplified model of some of these relationships.

Mark Walker and my colleagues in St. Louis and I carried out an analysis something like that in Fig. 6 in evaluating the impacts of a community-based program to promote asthma management among low income, African American children and parents. As depicted in Fig. 7, a descriptive analysis using structural equation modeling identified apparent benefits of key intervention elements, asthma management classes and community health workers. Furthermore, it showed that these were significantly predictive of “hard” measures of reductions in hospitalizations and emergency care from hospital records. These were significant after controlling for baseline level of hospital and emergency care, standard confounders in the field such as child’s age and sex, and parent’s social isolation [60].
There are pros and cons of analytic multilevel designs relative to the more conventional experimental design of COMMIT and similar trials. On the one hand, the experimental design of COMMIT offers the level of certainty an experimental design provides regarding the impact of that manipulation or set of manipulations to which individuals or units of analysis are randomized. On the other hand, the analytic multilevel model offers considerable flexibility in varieties of interventions (requiring then careful quantification of implementation and exposure) and allocation of resources to supporting broad implementation, evaluation, and analysis. In terms of making conclusions about interventions, the one offers whatever confidence experimental designs provide. The other offers perhaps greater external validity and opportunity to study interactions between interventions and their contexts yet still providing data of great precision with powerful control of potential confounders and sufficient inferential power to support decisions regarding interventions. Clearly, the validity of analytic multilevel designs depends on the extent to which we can quantify (a) variation among interventions and (b) important mediators and moderators or confounders of our independent and dependent variables.

Consider areas like smoking cessation in which a great deal of knowledge may guide identification of key variables, from advertising to candidate genes. Of what practical importance is the possibility that (a) after quantification of the impacts of most or all known confounders, (b) after specification of the quantitative relationships among policy, community, organizational, audience, and setting factors influencing the implementation of an intervention, and (c) after highly refined quantification of the relationship between key outcomes and extent of exposure to and engagement in interventions, it may not be absolutely certain that the intervention rather than some unanticipated and/or unmeasured variable accounts for observed changes in outcomes? Recalling that an experimental finding has a p value associated with it and some possibility that it might have occurred by chance alone, and that experimental designs often constrain operationalization of complex interventions, how much greater is the doubt in the validity of the findings of such an analytic multilevel design than the doubt that the experimental finding significant at $p<0.001$ may represent the one time in a thousand that the effect attributed to the experimental manipulation in fact occurred by chance alone? What if $p<0.01, 0.05$? It may be worth noting that most marketing decisions regarding consumer products and sales are made with much less powerful data than those of the analytic multilevel designs envisioned here. These decisions may serve as examples for important decisions in behavioral medicine and public health, such as which versions of already validated programs to offer which sections of a population.
If good science includes use of methods finely tuned to the questions at hand, planning of intervention research should begin by asking whether information is most needed in which of topics such as (a) the individual impacts of the intervention, independent of its context, (b) the impacts of a complex set of interventions elements and the range and variety of their implementation, (c) the interactions among intervention elements and their policy, community, organizational, or social contexts, or (d) the extent to which variability in intervention elements and implementation may be associated with the extent to which key audiences are reached, engaged, and benefit from intervention. The choice of design should be based on answers to these kinds of questions rather than based on a priori judgments about “gold standards” and quality of design. Our field would gain vigor and excitement and yield a broader range of useful information if, in addition to traditional randomized experimental designs, it incorporated a variety of designs such as preference designs [61] or practical clinical trials [62] as well as the analytic multilevel designs advocated here without a priori scoring of their strength.

To summarize this discussion of method, it points again to the importance of context. Without specifying the context of a study—the scholarly questions to be asked, the settings and populations, the type of intervention, etc.—it is pointless to talk about a “good design.”

Another concern about the application of randomized clinical trials to many community interventions lies in their extension from tests of isolated components of care, such as a drug, to studies of disparities in care. With many health disparities, disproportionate morbidity, burden, and mortality emerge from failure to diseminate and promote core elements of care that have already been shown to be effective (e.g., healthy diet, nonsmoking, cancer screening, or regular diabetes or asthma care). In such cases, assignment of a group to usual care that has already been shown to result in maldistribution of proven core elements of care raises serious ethical questions. Compare this to an RCT of, say, cancer chemotherapy in which (a) usual care of high quality is compared to (b) that usual care plus a medication of unconfirmed quality. The trial participant is offered the best care available along with the advantage of very active monitoring from the clinical team plus the possibility of being randomized to receive an additional medication of possible but unproven benefit. In contrast, consider the member of a disadvantaged group randomized to either (a) usual care that has already been documented to result in disproportionate morbidity or mortality in his or her group, or (b) a benign but perhaps unproven approach to promoting greater access and utilization of already proven interventions. The ethical parallel between the two—the RCT evaluating a medication of unproven value amidst adequate delivery of usual care, and the RCT comparing benign promotion of already validated, recommended care with a usual system of inadequate care—is far from close. We need to develop ways to study innovative approaches to promoting well validated interventions without assigning individuals or groups to conditions we already know will leave them at disadvantage.

Data Interpretation and Inference

In considering the quality of methods, I often find exceedingly helpful an observation of Arnold Lazarus from now over 40 years ago. At the time, the clinical evaluation literature in behavior therapy was composed almost exclusively of case studies. Many were carefully documented to provide strong bases for inferences regarding useful techniques but, nevertheless, remained case studies. I recall a small conference in 1965 at which a questioner challenged Lazarus on this evidence base. Lazarus replied that, although we tend to equate “scientific” with methodologies, he thought the essence of science is the criticality with which we talk about our data rather than the methods by which we gather them. From this perspective, then, a high level of modesty and criticality is necessary in talking about the inferences one can make from a single case. But the conversation is not eliminated because it is based on a case study.

It cuts both ways. We have seen often the lack of criticality in discussion of data gathered through methods that were considered solid amidst the standards of their day. Consider the “IQ controversy” or, currently, application of behavioral genetic findings that are naïve vis a vis the complexity of gene–environment and gene–gene interactions and of the phenotypes being “explained” (e.g., intelligence as a single thing). 6

Lazarus’ observation about the essence of science points to another important consideration, the irreducible role of scholarly judgment that is intrinsic in any extension of research findings to a real world setting. No matter the number of trials supporting an intervention or the number of settings in which the intervention may have been shown effective, there is always a level of scholarly or professional judgment necessary to make the decision that a given case or setting is an instance of those to which the intervention should apply—and that there are not other important factors present in the case or setting that may compromise that application [63]. That is, application of research always entails human judgment. The extent of that dependence on human judgment may vary and the extent to which research may provide guidance to judgment may vary, but the idea that our evidence can be so tight as to eliminate human judgment is quixotic. If we do not recognize this role of

judgment and exercise it with modesty and criticality, we are prone on the one hand to fail to extend our knowledge to those it may benefit, or, on the other, to extend it in wooden or foolhardy ways to those who need it applied wisely.

The current Guide to Community Preventive Services includes the finding that sufficient evidence supports diabetes self management education if offered through “community gathering places” such as “community centers, libraries, private facilities (e.g., cardiovascular risk reduction centers), and faith institutions” [64] (p. 201) but not if offered through worksites. Only one study was found reporting self management education in worksites and it “had design limitations” [64] (p. 207). Consider these findings from the perspective of someone developing a diabetes program and whose advisory board believes in evidence-based practice. A proposal that the program include self management classes implemented in a local worksite whose management has expressed enthusiasm for the idea and willingness to support it energetically might well be greeted with criticism that there is no evidence to support the program in worksites. Assuming a suitable environment of confidentiality and management support, how many of us would really believe a diabetes self management class offered to volunteers in a worksite is less likely to be effective than the same program offered in a church or a private CVD risk-reduction clinic? The presence of sufficient evidence for one category and the lack for the other reflects (a) the judgment to place in the category of community gathering places such diverse settings as libraries, cardiovascular risk reduction centers, and faith institutions (without assessing the evidence separately for each of these) while leaving worksites in their own category, and (b) the fact that there was only one study found for worksites. It does not necessarily reflect the true efficacy of diabetes self management education in each of the several settings evaluated. This points to at least three considerations for using such compilations of evidence: (1) consider the real world impact of the finding of “insufficient evidence”; (2) consider how the volume of research in different areas shapes conclusions and whether that volume, itself, truly represents the best available knowledge to guide practice; and (3) use good judgment in framing questions such as the decision to distinguish between worksite and community settings in the first place.

Concern for parsimony teaches us to assume generality of findings until data force us to do otherwise. Without statistical or other evidence documenting a significant difference between interventions A and B, asserting a difference between them in the form that there is evidence for treatment A but insufficient evidence for treatment B should be trimmed by the sharp edge of Occam’s razor. Returning to the example above, if there are no data comparing worksites to other community settings that show significant differences between them, there can be no evidence based distinction made regarding their efficacy. One cannot prove the null hypothesis. This is especially a problem if stated limits become the basis for assertions of “no evidence” in areas which have merely received insufficient research attention. Indeed, the practice of asserting a conclusion of “insufficient evidence” in areas that have received insufficient attention runs the risk of a damaging spiral in which topics for which it is hard to carry out conventional research designs (e.g., dissemination of practices to underserved groups from whom it is difficult to obtain consent for enrolment in randomized trials) will tend to receive little research attention meeting current standards of “acceptable evidence,” leading to conclusions that “there is nothing [evidence-based] we can do,” all leading to failure to find the things we can do.

As an undergraduate, I had the privilege of studying with Gail Kennedy, a professor of philosophy who introduced me to Freud and, as a student himself of John Dewey, also introduced me to American Pragmatism. I remember the very kind and generous Professor Kennedy offering the pointed observation “Getting the right answers is easy. It’s asking the right questions that’s difficult”. There is an almost infinite set of questions that might be asked of the data on which evidence based guidelines are based. The choice of questions, e.g., the decision to separate the evidence for diabetes education in community and worksite settings in the first place, needs to reflect our best wisdom about the field.

**Contexts, Dissemination, and Maintenance of Behavior Change**

Consideration of the application of evidence raises issues around dissemination of our knowledge. Surely, contexts are critical to dissemination. Whether behavioral medicine is supported and promoted and made available to those who may benefit will depend on decisions made by individuals and organizations with no particular allegiance to behavioral medicine. It will be influenced by a variety of factors, even the subtleties with which apparently innocent phrases like “personalized medicine” can be turned into strategies for projecting a narrow view of healthcare and progress in healthcare as resting primarily in esoteric regions of biology [65].

As Ernst Wynder a number of years ago and Larry Green, Russ Glasgow, and others have urged more recently, we need dissemination research that starts with an assumption of efficacious intervention and examines ways in which it can be implemented, promoted, disseminated, and broadened in its benefit. Along these lines, Steven Schroeder has recently analyzed failures of the U.S. health system and identified the potential contributions of
behavior change and improved access to care in efforts to “become number one in health” [66]. It has become clear that interventions which may have been shown in controlled trials to have efficacy may nevertheless fail to demonstrate benefits when implemented more widely [67]. Fortunately, recent calls for greater emphasis on external validity [68] have been supported by policies of research journals [69].

Glasgow et al. [70] have articulated the RE-AIM framework for organizing research into dissemination of interventions. This provides an orderly series of questions to address whether interventions are reaching those for whom they are intended, whether they are effective within the group, and whether they are maintained and integrated into routine practice. The kind of analytic multilevel designs suggested above would have considerable utility in studying the implementation, engagement, and impacts of interventions across large populations with the power of data sets resting more on the numbers of observations and range of settings considered than on the depth of characterization of any one individual and setting.

In many areas of behavioral medicine, efficacy has been established. There is a state of the art for helping people quit smoking, helping them lose weight, helping them manage stress and reduce associated cardiovascular risks, increasing physical activity, or promoting chronic disease management. It is debatable whether these areas require better interventions or better ways of promoting and disseminating the interventions we already have. Rather than efficacy in the purposefully limited context of the experimental design, more timely questions surround reach, success in engaging high priority populations, or robustness and replicability of impacts.

As the previous paragraph listed the interventions for which one can claim efficacy, the critical reader may have responded, “Yes, we have shown we can help people quit smoking, lose weight, and so forth but we haven’t shown we can help them stay off or keep it off.” This points to another area, highly dependent on contexts that links efficacy and dissemination. That is sustaining the behavior changes accomplished. We all have great respect for intervention studies that include follow-up of 2 or 3 years. Consider now that the average individual with type 2 diabetes will live 20, 30, or 40+ years with the disease. How do we make the extension from studying maintenance of change over two or three years to developing systematic ways of supporting individuals needing to maintain changes for decades? In the Diabetes Initiative of the Robert Wood Johnson Foundation, we have come to realize that the most important characteristic of type 2 diabetes and self management of it is that it is “for the rest of your life” [71]. It sounds simple but we have been struck by how often this consideration reframes our thinking about self management programs. As an example consider the goals in working with a 45-year-old adult whose diabetes is in poor control. Is the goal getting that control improved in the next 3 months? Or, is the goal establishing an approach to living with diabetes that will help the individual attain the best possible control over the next three or four decades? Does the choice of goal have implications for the approach to helping the individual? The lifespan is an important context of behavioral medicine and one we are just beginning to grasp.

Conclusion

Twentieth century positivism helped identify causal relationships among behavior, feelings, health and death, relationships previously thought beyond the reach of objective analysis. Implicit in such causal analysis is an emphasis on context—the causal analysis forces attention outside the effect to something else that influences it. Causal analysis and sui generis are fundamentally incompatible. In this sense, contemporary ecological perspectives and twentieth-century positivism are in accord.

Closely aligned with positivism was the modernist impulse that clear thinking can find the cause, analyze the situation, and prescribe the solution, as in planned communities that now look hopelessly rigid and without spirit. This modernist certainty has given way to a postmodern recognition of the turbulence of causes and contingencies and resulting uncertainty concerning any simple statement that A led to B or that we even know B. Consider for example the rejection of simple narrative and certainty in Barton Fink, Fargo, and other movies of Ethan and Joel Coen. Integrating our positivist roots with a postmodern recognition of the contingent nature of all things—whether the long allele is good or bad depends on its context—forces us to a broader approach to knowing that embraces and studies context, rather than the quixotic view that the only real knowledge requires that we control it.

References