Defining the environment in gene-environment research: lessons from social epidemiology

Jason D. Boardman
Jonathan Daw
Jeremy Freese

Published in *American Journal of Public Health*, 2013 August; 103(Suppl 1): S64–S72.

Defining the environment in gene-environment research: lessons from social epidemiology

Abstract

We review the current gene-environment interplay literature and show the importance of a social epidemiologic framework for contextualizing gene-environment interplay hypotheses. This sociological orientation to the environment includes a specific focus on a multilevel perspective in which environments are characterized as broad social contexts such as schools and neighborhoods. We encourage those already doing social epidemiology to consider the potential value of considering genetic moderators in their work. Likewise, we argue that researchers interested in gene-environment interplay should consider the merits of the social epidemiologic model of the environment. Toward this end, we emphasize the importance of future gene-environment research continuing to elaborate social aspects of the environment, the cumulative influence of social environments for an individual during their lifetime, change in environmental factors over time, and the genetic factors that make some particularly sensitive to broad environmental influences.

Introduction

Inquiry into the complex relationships between genetic and environmental influences on behavioral traits has increased substantially in the past decade (1, 2), a trend which is particularly pronounced in health research (3-6). A *PubMed* search yields 42 articles published in 2000 that contained the expression "gene-environment interaction" in the title, abstract, or keywords; this number increased to 704 by 2012. Although new and important findings have emerged from this large body of work, there are also strong criticisms of the existing gene-environment interaction (GxE) studies from researchers across the health, psychological, social sciences (2, 7-10). The most obvious criticism is the poor replication record for "established" gene-environment interaction results (11, 12). Others have pointed to problems with statistical power (8) and biological plausibility (7), potentially compromising many, if not most, candidate gene-by-environment (GxE) interaction studies.

To date, however, there remains very little discussion about a different shortcoming in existing GxE research. Specifically, there is no real consensus about nature and scope of the *environment* within GxE studies (13). As the *E* is one-half of the GxE framework it is critical to define the environment in a manner that maximizes the contributions from both social and biological sciences which enhances the potential for this work to improve public health. This need for cross-disciplinary discussions is echoed in the current efforts of the National Coalition for Health Profession Education in Genetics.[1] This group, with support from the Office of Behavioral and Social Science Research (OBSSR) with the National Institutes of Health (NIH), have developed a project entitled "Genetics and Social Science" with the explicit goal to "create an educational program that will improve social scientists' genetics literacy." This project points to a variety of collaboration opportunities within the area of gene-environment interplay and states in part that

[1] http://www.nchpeg.org/bssr/. Accessed 12/14/12.

"geneticists may be less familiar with measures used to quantify the observable external environments, and can benefit from the guidance of social and behavioral researchers." One goal of this paper is to address the latter comment by offering guidance for operationalizing and measuring the *social environment* in GxE studies. Consensus regarding the definition of the social environment will help to guide future work and locate GxE evidence in a more coherent framework, and address major limitations in conceptualizations of the environment in research on GE interplay.

We make three contributions toward this goal. First, we discuss the importance of existing social epidemiological and sociological theory for understanding the environment in a multilevel, multidomain, longitudinal framework that accounts for upstream processes influencing health outcomes. In particular, this approach draws a sharp distinction between individual and family attributes and the broader social contexts in and through which they arise. Second and relatedly, we emphasize the potentially important role that characteristics of intermediate levels of social organization such as neighborhoods, schools, and the workplace have to play in a more thoughtful account of the environment in gene-environment interplay research. Finally, we discuss different forms and models of gene-environment interplay with frequent reference to previous published research.

What is the environment? The social epidemiologic perspective

In one of the first papers describe a general framework for GxE associations in epidemiologic research, Ottman (14) defines the environment as follows: "The environmental risk factor can be an exposure, either physical (e.g., radiation, temperature), chemical (e.g., polycyclic aromatic hydrocarbons), or biological (e.g., a virus); a behavior pattern (e.g., late age at first pregnancy); or a "life event" (e.g., job loss, injury)" (pp. 764). Although this statement accurately summarizes how most GxE research approaches the environment, it is limited in at least two respects. First, each of the factors that are described may be thought of as *proximate* environmental

moderators of genetic associations. This same characterization of the environment is evident in the Gene Environment Association Studies (GENEVA) consortium which is led by NIH and NHGRI through the Gene, Environment, and Health Initiative (GEI). The list of published papers from this group includes "environments" such as obesity (15) and maternal smoking (16), which are far downstream from social environmental factors that structure exposure in the first place (17). In contrast, the *fundamental cause* perspective (17) argues that "individually-based risk factors must be contextualized, by examining what puts people at *risk of risks*, if we are to craft effective interventions and improve the nation's health." (pp. 80). Full understanding of the determinants of a health outcome requires understanding the social structure from which proximate risks and exposures have arisen.

Second, emphasis on individual environments does not account for group-level behavioral, normative, and cultural processes that shape individual health and behavior. To illustrate the importance of this issues within GxE research, consider a recent paper in the *American Journal of Epidemiology* (18) in which researchers examined the interaction between single nucleotide polymorphisms (SNPs) within 38 genes and specific health behaviors (e.g., smoking, drinking, exercise, and nutrition) on body mass index (BMI) among white and black adults. They provide evidence for gene-behavior interactions (GxB) by demonstrating that the association between each health behavior and BMI depends on the genotype of individuals. By labeling these GxBs as GxEs, this approach takes at best a very limited view of the nature of the environment. This same emphasis on behavioral factors as environmental moderators of genetic influences on obesity and BMI is also evident elsewhere (19). Understanding how genes moderate the consequences of behavior is an important component of a genetic epidemiologic understanding of health but, as others have made very clear (20), it is distinct from GxE research. Individuals do help shape environments through their behaviors, but it is nevertheless important to distinguish between the actions of people and the

circumstances in which these actions occur. The latter incorporates a much more comprehensive approach to the environment for gene-environment interplay research.

This distinction conforms to the social epidemiologic emphasis on the upstream sources of risk exposure. Social epidemiology explicitly reframes traditional epidemiologic paradigms by emphasizing *the role played by an individual's location within a particular social structure as a fundamental determinant of vulnerability and exposure* (21). Accordingly, we conceptualize the social environment as an external, multilevel, and multidimensional feature that determines an individual's exposure to risks and access to resources and constrains or enables people to engage in healthy lifestyles at different stages of the lifecourse.

A unique contribution of the social epidemiological perspective in this context is it emphasis on the *embodiment* of social arrangements, or "how we literally incorporate, biologically, the material and social world in which we live, from conception to death" (22:672, 23). Sociologists' contribution to this idea is the explication of *pathways of embodiment* that constrain and enable individuals' capacities to live healthy lives, including social structures (24-26). These pathways are multilevel, multidomain, and multi-timescaled. Multilevel pathways incorporate contextual dynamics at supraindividual, often nested, levels of analysis (e.g., families, schools, neighborhoods, states, countries). Multidomain pathways span different spheres of people's lives (e.g., social, economic, physical, and institutional). Multi-timescaled pathways encompass both change within individuals over the life course and historical changes in populations. Importantly, Krieger (22) writes that embodiment provides a "biological expression of social relations" (672) and as such, the complex, dynamic, and transactional nature of the social environment becomes a critical input into basic biological processes.

One important aspect of this perspective is that environmental risk factors are not characterized as independent of one another. For instance, the joint distribution of collective efficacy, socioeconomic status, and crime rates (27) across neighborhoods in Chicago makes it

difficult to consider each of these factors as independent variables in traditional multivariate models. Just as the 'fundamental cause' perspective focuses on an individual location within the social order as relative factors rather than an objective indicator of 'exposure', the clustering of social characteristics within geographically defined neighborhoods and schools provides important evidence about the relative position of a particular social context along a continuum of privilege and disadvantage. Identifying the mechanisms through which this allocation system affects measured phenotypes is critical, but exclusive focus on downstream processes like stressful life events and behaviors loses sight of the possibility that ill-health and social risks will often be derived from the same source.

This understanding is very important because it helps to contextualize findings from genetic epidemiology studies in which genetic associations are shown to be different for members of different racial, ethnic, and socioeconomic groups. As discussed below, environmental factors may fundamentally alter the way in which genes are associated with health outcomes because in some residential areas, health may be driven exclusively by the physical and social features of the neighborhood and genes have virtually nothing to do with individual differences in health within these communities. For example, using data from the Chicago Health and Aging Project, researchers have shown that the association between the ApoE-ε4 genotype and change in cognitive function is the strongest in the most socially organized neighborhoods in the Chicago area (28). Consistent with the "social distinction" model we will describe below, these researchers argue that the comparably small influence of genotype is further muted by social factors that may profoundly influence cognitive decline in the most disorganized communities.

This understanding is also in line with the social construction perspective on racial and ethnic identity (29) that is shared by most social scientists. This includes focusing research on features of the social environment that are amenable to policy interventions and are precursors to

the observable behavior, rather than emphasizing racial phenotype as a cause. Absent reliable and valid measures of the environment and theory linking environmental factors to health behaviors, results from genetic association studies may, at times, provide misleading conclusions. In an influential example, Turkheimer and colleagues provide convincing evidence that the heritability of cognitive test scores is virtually zero for those who are raised in the most disadvantaged homes but increases dramatically as the level of socioeconomic resources increase (30). Others report similar results (31), and together this research indicates that genetic factors linked to cognitive performance may not be fully realized for those in the most disadvantaged communities.

The social epidemiological focus on pathways from social structure to health is critical because it better clarifies the factors that structure both differential exposure and mitigating resources. Nevertheless, this approach is limited by its inattention to gene-environment interplay. Consider health-related behaviors such as exercise, nutrition, substance use, and adherence to medical treatments. All of these are necessarily linked to the ecosocial precursors but, just as importantly, people from comparable ecosocial environments respond differently to similar environmental conditions. The links between social structure, the physical and social environment, health behaviors, and morbidities are well established, and yet it is increasingly clear that genotype may factor into this conceptual orientation at each stage of the process. In other words, as research document the chains by which broad social conditions have specific health consequences, genetic differences appear as potential moderators of each link in this chain.

In this respect, gene-environment interplay provides a great opportunity for elaboration of the social epidemiological perspective in public health. Advances in molecular technology have made it possible for researchers to incorporate genotypic information into this traditional social epidemiologic framework to ask new and important questions that involve genetic differences and yet nevertheless remain true to core principles of social epidemiological thinking. The notion of

embodiment as both an *indicator* of social location and a *cause* of future health trajectories becomes more, not less, relevant as we learn more about the human genome. As others have made clear, understanding both social and genetic risks at each developmental stage is critical to understanding specific pathways to divergent health outcomes throughout the lifecourse (32-34).

The ecosocial perspective emphasizes the role of places in which individuals reside, work, interact, and attend school in health; life course theory emphasizes that the environments which are most important change in predictable ways across the life course. During gestation, the uterine environment and determinants of maternal health are the most important environmental influences on health outcomes. During childhood and adolescence one's parents, neighborhood, school, and social networks are the most robustly influential. In adulthood, the workplace becomes an increasingly important environment, and one's formed family and home become increasingly important from young adulthood to old age. Each of these social environments serve to link individuals' place in the broader social structure to their lived lives, and influence the set of risks and resources available to them which influence their embodied health and health behaviors. Measures exist for several well established social environmental factors related to health including social integration (35), collective efficacy (27), social capital (36-38), psychosocial stressors (39), behavioral norms (40), and segregation (41).

We argue that genetic influences should be incorporated into this model, as they potentially influence all of these connections (42). Genetic differences influence how individuals end up in different types of environments (43). Genetic differences moderate how particular environments translate into environmental risks, resources, and health behaviors. Finally, genetic differences also likely moderate how these risks, resources, and behaviors all influence embodied health outcomes.

To summarize, we argue that previous GxE research has adopted an improperly atomistic view of the social environment, often even treating behaviors as environmental characteristics. In

contrast, a social epidemiological perspective contextualizes individual actions and attributes within the broader organization of society into institutions and meaningful social groups, to which health risks and resources are systematically and jointly distributed. Taking the nature of this allocation system seriously in gene-environment interplay research entails a move away from mere risk factor epidemiology and toward a focus on environmental pathways to embodiment of social conditions from macro to micro. This joint distribution of health-relevant features of the social environment means that genetic influences on health may be far more important in some contexts than others, in some stages of the life course than others, and for some socially meaningful groups than others. Finally, it may frequently be the case that specific genetic loci serve to modify the effects of these environmental risks and resources on health outcomes, as is discussed presently.

Types of Gene-Environment Interplay

The social epidemiologic perspective provides a useful framework to delineate meaningful social environments for research on gene-environment interplay. Most broadly, this "interplay" encompasses a combination of gene-environment interactions (GxE) and gene-environment correlations (rGE). We consider the latter below. *Gene-environment interactions* describe situations in which the observed effect of environmental differences is contingent on genetic differences. Such interactions can be usefully subdivided into two distinct types. A *heritability by environment (HxE) interaction* is a population-based model that estimates the relative contribution of genetic influences to overall phenotypic variance across different environments (32, 44). As with the bulk of the GxE research, much of this work focuses on proximate environmental influences at the individual and family levels. For example, Silventoinen and colleagues (45) used samples of twins from Denmark and Finland to examine the heritability of body size, showing that genetic associations for body mass are lower for those who exercise more and those whose diets contain a larger portion of protein compared to those who do not exercise and eat less protein. Likewise, Gottlieb and colleagues (46)

use data from the Framingham Heart Study and demonstrate that the heritability of lung function (FEV1) increases from .05 in the entire population to .18 when they only consider current smokers. In this case, some genetic differences that would otherwise be inconsequential for lung function may influence lung function among those who smoke. This individual focus is equally clear in the Turkheimer and colleagues (30) paper described above, in which the heritability of IQ increases substantial as familial socioeconomic resources increase.

This same emphasis on proximate environmental determinants is also evident in studies that rely upon candidate gene-by-environment research designs. Because these studies focus on environmental moderation of the association between a specific allele and a health outcome, this type of GxE association can be referred to as an *allele by environment interaction* (AxE; the distinction between HxE and AxE is also referred to as the difference between "latent" and "measured" GxE (47)). The most widely cited AxE interaction, despite a fairly weak replication record (8, 48), is found in the work of Caspi and colleagues (49) who show that carriers of the short allele in a gene that codes for the serotonin transporter (*5HTTLPR*) are particularly sensitive to individual-specific stressful life events but that the carriers of two long alleles at this loci are fairly immune to the deleterious effects of regular exposure to strain and stress. In a similar manner, Mitchell and colleagues (50) report two genetic polymorphisms that are associated with a crossover in the relationship between socioeconomic status and postpartum depression: the genotypes that conferred more risk for poor mothers conferred *less* risk for wealthier ones. There are countless examples of AxE research in the psychological, social scientific, and health literatures, but the overwhelming share of these findings operationalize and measure environmental exposure as a proximate and individual-level characteristic (see Duncan and Keller (8) for a review).

This body of work is critical to public health research because it signals a need to consider specific environmental contingencies that may mask or illuminate genetic influences on health and

well-being. However, it is limited because the environmental factors are typically either behaviors (e.g., smoking) or family characteristics (e.g., socioeconomic status). In the past decade a body of research has emerged that focuses on exogenous and more broadly defined social environments such as neighborhoods (51), schools (44, 52), states of residence (53), and historical periods (3, 54, 55) as important environmental moderators of genetic effects on health and health behaviors. The focus on these broad social environments is important because it delineates a range of social contexts in which individuals are socialized about health-related behaviors that are pegged to key developmental periods. These environments also provide socially and geographically meaningful boundaries for policy makers to implement specific public health initiatives.

The limited examples of this work have provided important substantive and methodological contributions to the GxE research. For example, a recent paper shows that the magnitude of the association between one SNP (rs1801282) and metabolic syndrome varies depending on the availability of exercise facilities (56). In other words, changes to the structure and aspects of built environments can affect the association between specific genetic variants and specific health outcomes.

Gene-environment interactions can also be distinguished by the functional forms of the relationship between genotype, environment, and outcome. The four rows of Figure 1 distinguish four models implied by a GxE typology that is used by researchers (32, 57), differentiated by their HxE formulation (left column) or AxE forumulation (right columns).

[Figure 1 about here]

The first two rows of Figure 1 depict the *diathesis stress* and *differential susceptibility* models (58-61). Both propose that individuals with chronic exposure to socially risky environments are more likely to display poor health. The diathesis stress model suggests that the genetic differences that are associated with negative outcomes in risky environments will have either an attenuated or entirely

muted relationship in low-risk environments. This is best characterized by the work of Caspi and colleagues (62) described above. As shown in Figure 1, a diathesis-stress model implies increasing heritability in negative environments, and an allelic divergence as adversity increases.

A complement to the diathesis-stress model is one that calls attention to how gene-outcome relationships can be attenuated by *social control* (63). As an HxE example, Boardman finds that the heritability of regular smoking is significantly reduced in states that have the most restrictive policies regarding the sale of cigarettes and in states that have the highest taxes per pack on cigarettes (53). An AxE counterpart is shown in the work of Fletcher (64), who shows that the association between a single nucleotide polymorphism (SNP) in the CHRNA6 gene (rs2304297) and tobacco use described by others (65) is significantly reduced for those who live in states with the highest levels of tax on tobacco products.

The differential susceptibility hypothesis, on the other hand, implies that the same genotypes that are associated with negative outcomes in adverse environments may be associated with positive environments in less adverse ones. The study by Mitchell and colleagues discussed earlier (50) serves as one illustration. This is shown with the u-shaped HxE association and the crossover AxE association in Figure 1. As another example, Simons and colleagues (61) show that individuals with a higher number of plasticity alleles (the 7R allele in DRD4 and the S allele in 5HTTLPR) were the most aggressive in the most adverse social environments and least aggressive in the least adverse social environments. Their paper is an important extension to the GxE research because their research employed an inherently multilevel perspective emphasizing social resources from the respondent's neighborhood, school, and family levels of social support.

At the same time, the approach to the environment in this study does not contain any information describing the behavioral expectations, a description of the sanctions for violated norms, or a description of the mechanisms in place to enforce these norms. This difference is shown

in the work of Daw and colleagues (66) who examine the link between school-level smoking behaviors and the likelihood that individuals will begin smoking themselves. They show that increasing copies of the short allele in the 5HTTLPR gene increased the likelihood that individuals will adopt the smoking norms of their school. The association was even stronger for the drinking phenotype, and the differential susceptibility model seems to best characterize the link between school-level drinking patterns and individual risks of drinking. Specifically, in the schools that have the lowest drinking rates, those with the short allele drank the least, but the same allele was associated with the highest alcohol consumption in schools that had higher than average drinking levels. This is important because without this type of specification, one would not see an association between genotype and phenotype. This has been discussed recently in the debates regarding the power of candidate GxE associations (67) but it is also important because it suggests that normative factors that limit or enable specific behaviors should be considered as potentially important moderators of genetic effects like chronic stressors (49) or early childhood abuse (62).

An additional point this example highlights that is of particular concern to health researchers is the need to consider the environments across the full continuum rather than simply exposure. Consider a study in which differential susceptibility loci were the key element placing individuals at risk of smoking cigarettes. If this study were done in communities in which very few people smoke, one would conclude that the allele associated with lower smoking in the population as a whole was in fact the risk allele. If the same study were done in typical environments, one would not see any association, but if performed in the most risky environments, then the allele associated with smoking over all environments would emerge as risky, but extrapolation of the conclusions would misstate the direction of the effect for low-risk environments. Absent a complete representation of the individuals across the full range of environments, researchers can only tell one part of the story.

Characterizing the environment across the full continuum is also important because it allows one to examine the *social push* and *social distinction* GxE models. The social push model differentiates between typical and extreme social contexts and hypothesizes that genetic factors will be the most important within typical environments whereas social influences dominate within extreme environments. In these extreme environments, social factors so strongly influence the phenotype that ordinary genotypic differences have little room to differentiate individuals from one another. However, environments that have fewer social factors that limit individual differences allow for "biology to shine through" (68). The social distinction model is very similar to the social push model but it anticipates that the highest social risk environments will have the lowest heritability and lowest measured genetic associations. In this respect, the social risk hypothesis is essentially the statistical obverse of the diathesis-stress model, in that adverse environments mute the influence of genetic differences.

The social push and social distinction environments are not necessarily causal GxE models in the biologic sense of genes actually functioning differently in different environments. To illustrate the issue, researchers have shown that genetic factors related to smoking were virtually non-existent in the early 1960s but then became increasingly important for smoking initiation following the Surgeon General's report on the dangers of smoking (3). The researchers argue that those for whom smoking was driven by social factors were far less likely to initiate smoking, as well as more likely to successfully quite smoking, after the 1964 report, compared to those for whom smoking was largely due to genetic factors related to nicotine metabolism. In other words, this important scientific announcement had significantly less influence on the future smoking patterns for individuals with specific genetic risk profiles, because it affected the social costs and benefits of smoking, rather than any moderation of the role of genetic differences in nicotine metabolism itself. To the extent that reduction of overall smoking rates may have occurred largely among those for whom smoking was

intrinsically less rewarding, public health campaigns against smoking may have changed the actual allelic composition of the population of smokers while reducing the number of smokers overall (54).

Evidence for the social push and distinction models can be found in the public health and problem behavior literatures such as the work on ApoE described above (28). A similar result can be seen in the work of Tuvblad and colleagues (69) who examine antisocial behavior in 1,133 Swedish twin pairs (ages 16-17). The study uses a broad indicator of the social, economic, and behavioral context of the neighborhoods and finds that the heritability of antisocial behavior is significantly higher for those who reside in the most socioeconomically advantaged neighborhoods. As a last example, Boardman and colleagues (44) take advantage of the school-based design of the National Longitudinal Study of Adolescent Health to show that social understandings of body size substantially moderated the estimated influence of genetic differences on BMI. They examined the average BMI for those who said that they were "normal weight" to calculate a school-level norm about body size. In line with the social push models, they show that the heritability of BMI is the highest in schools with body size norms in the average range but lowest in schools in which the norm is very low or very high.

As noted earlier, gene-environment interplay encompasses not only gene-environment interaction but also *gene-environment correlation* (rGE), in which genotypes are associated with causally relevant aspects of the environments to which an individual is exposed (70). rGE potentially creates the appearance of a direct gene-health relationship where none exists. *Passive* gene-environment correlations result from genetic influences not directly on the individual but, rather, on biologically related individuals involved in the environment in which the individual develops, especially parents. Price and Jaffee (71) describe work in which parents with lower verbal ability raise children in environments that have more disorganization in the home, and that this disorganization has a causal

effect on the child's verbal ability. This has the side consequence of creating a spurious association between children's genes and verbal ability.

Alternatively, genetically-influenced individual traits can influence the environments that an individual may experience. Thus, genetic factors are an indirect cause of whatever other traits these environments may influence. The key distinction often drawn here are between traits influencing their selection of environments (*active* rGE) and environments responding differently to individuals based on observable traits (*evocative* rGE). As an example of the latter, if differences in skin color lead to differential treatment and experiences of discrimination, then pathways from discrimination to health outcomes could induce a correlation between genetic causes of skin color variation and health (72). In this way evocative rGE closely corresponds to the sociological notion of ascription (73-75), insofar as the latter is based on genetic foundations. Active gene-environment correlations encompass genetic influences on the environments that individuals seek out. For instance, Cleveland and colleagues (76) find evidence for genetic influence on whether one has friends who smoke and drink. If these friendships, in turn, influence whether adolescents smoke and drink themselves, then friendship selection mediates a relationship between genes and these health behaviors.

Gene-environment correlation is very important for the GxE research described above because a key assumption of GxE research is that the environmental exposure is assumed to be independent of genotype. Others have shown that violations of this assumption can have important implications for the interpretation of the GxE estimates (77). The most effective strategy to deal with the possibility of rGE in GxE studies is to consider environmental factors that are exogenous to genetic characteristics of individuals (78). This further highlights the importance of the ecosocial perspective because the emphasis on large environmental contexts such as schools, neighborhoods, or counties reduces the likelihood that genetic and environmental factors are correlated.

Discussion

Although researchers have given much attention to gene-environment interplay, this work has thus far focused on a fairly narrow characterization of the environment. As social epidemiology and sociological research has shown, the social environment is more than a set of independent risk factors and protective influences. Instead, society and its major institutions and contexts, such as schools, neighborhoods, businesses, and the built environment are jointly distributed in a manner which disproportionately channels health-promoting resources to the wealthy and powerful at the expense of the poor and powerless. Thus, good schools and safe neighborhoods (79), opportunities for good careers (80), and access to nutritious food, health care, and conditions amenable to exercise (81) are disproportionately available to higher SES families. Equally important, the distribution of resources and risks obviously has substantial consequences for health inequality (82), and genetic epidemiology has heretofore paid limited attention to these lessons from social epidemiology. To be sure, sociological researchers have expressed valid concern regarding the potential consequences of blind enthusiasm for the marriage between genetic and social explanations for behaviors (83). But as others have pointed out (25), sociological explanations become far more relevant when the genetic influences on social forces are made clear. Advancing understanding of these processes should therefore be a high priority for both sociology and public health.

However, much work remains to be done in this area of research. Perhaps the most important limitations are: a limited conceptualization of the nature and scope of the environment and its interaction with the genome; limited sample sizes available to study this topic in a biologically-informative manner; the weak replication record for some of the most widely cited GxE associations (8, 48); and the lack of analytical strategies which offer causally satisfying interpretations. In this paper we have sought to address the first limitation, and the second is increasingly being addressed by efforts to genotype long-standing, large-sample, population-

representative social science datasets such as the Health and Retirement Study, the National Longitudinal Study of Adolescent Health, and the Wisconsin Longitudinal Study. The incorporation of genetic samples in to moderately sized and representative data sources may help to clarify the salience of the GxE perspective and it will certainly help stabilize the GxE parameter estimates that show a great deal of variation across different and at times fairly small studies (11). However, the sample sizes of these studies still fall well short of nearly 100,000 observations that some have argued are needed to identify true GxE associations (84). Indeed, the presence of statistically significant GxE associations within the literature has led some to assert that the bulk of these associations are likely to be false positives and appear in scholarly journals because of publication bias (8).

Concerning the last limitation, most research on GE interplay in public health and elsewhere is primarily correlative, providing evidence on interactive associations but not necessarily causal ones. Because genetic effects on the phenotype may be confounded by population stratification (85) and rGEs (77, 86), and the environmental effects may be confounded as well, these methodological limitations threaten to undermine the interpretability of this important research. This last point is important and it is worth repeating. Population stratification is one form of gene-environment correlation that is a strong challenge to our claim of exogenous environmental exposure. That is, residential segregation by race and ethnicity remains a fundamental feature of social life in the United States (87). Small differences among socially defined racial and ethnic differences in allele frequencies for genes that are related to specific health behaviors is the primary concern of population stratification but these same small differences may be correlated with neighborhood characteristics that we are describing as exogenous. As such, we encourage researchers to employ one of many standard statistical approaches to adjust the possibility that environmental exposure and genotype are independent above and beyond population differences across the genome. These

methods include 'ancestrally-informative markers' (88), principal components (89, 90), sibling fixed effects or family-based studies to reduce this influence of this form of rGE (66, 91, 92).

Conclusion

Our discussion offers three primary lessons for gene-environment interplay research within public health. These lessons are derived from the demonstration that most health behaviors of interest to public health researchers have a heritable component but that the relative influence of genes is often contingent upon environmental factors. First, we advocate taking the multilevel, multidomain, and longitudinal nature of the environment seriously in gene-environment interplay research. We believe that the social epidemiological framework offers the best approach to doing so, as its focus on the upstream processes of social organization which lead to the joint allocation of health risks, resources, and norms within society offers a fuller understanding of the environment than that shown in most research on this topic. This approach emphasizes that behaviors are not environments, that individual and familial environmental influences are best understood in their broader social contexts, and that proximate risks and rewards in the pathway between social structure and health are often systematically and jointly distributed.

Second, we emphasize the role of intermediate levels of social organization, such as neighborhoods, schools, workplaces, and social networks, as important features of the social environment for understanding gene-environment interplay and health. These units of organization provide important linkages between the broader social structure and individual lives, and have the benefit of providing plausibly exogenous sources of environmental variation for models of GE interplay. Which of these units of social organization are most consequential varies systematically through the life course. In addition, the specific ways that these intermediate levels influence individuals' lives are highly variegated, but assessing their comparative importance can provide important clues toward identifying their key etiologic attributes.

Third, we highlight different basic forms of gene-environment interactions and gene-environment correlations, with examples that have been observed. The differences between these forms have consequences in predicting the population consequences of environmental changes, and distinguishing among them requires information on the full range of phenotypes. Articulating the models also provides an opportunity to emphasize the difference between biologic and statistical interaction, as changing social conditions can influence the observed population association between a gene and an outcome without at all moderating the biologic effect of genes.

It is our hope that the research will continue to provide new insights for public health research from the simultaneous consideration of genetic and social factors. We hope that this framework and language will help to organize the otherwise atomized results from the large body of GxE research. We stressed the need to consider social components of the environment that provide cues about specific health behaviors in specific social contexts and specific times in the life course; environmental risks involve shared understandings about the meaning of risks which are critically related to norm formation and enforcement across different contexts (93). Treating risk as a characteristic of an individual may be a very useful model for the medical sciences but it does very little to advance our understanding of public health because we lose sight of the social origins of individual beliefs and behaviors. In this manner, it is our hope that social scientists recognize that processes of gene-environment interplay are an important subsequent of the class of generic social processes whereby features of the environment and the individual recursively influence health.

REFERENCES

1. Bearman P. Exploring Genetics and Social Structure: Introduction. American Journal of Sociology 2008;114:I-VI.
2. Freese J, Shostak S. Genetics and Social Inquiry. In: Annual Review of Sociology. Palo Alto: Annual Reviews; 2009. p. 107-128.
3. Boardman JD, Blalock CL, Pampel FC. Trends in the Genetic Influences on Smoking. Journal of Health and Social Behavior 2010;51(1):108-123.
4. Schnittker J. Happiness and Success: Genes, Families, and the Psychological Effects of Socioeconomic Position and Social Support. American Journal of Sociology 2008;114:S233-S259.
5. Guo G, Roettger EM, Cai T. The Integration of Genetic Propensities into Social Control Models of Delinquency and Violence among Male Youths. American Sociological Review 2008;73:543-568.
6. Seabrook JA, Avison WR. Genotype-environment interaction and sociology: Contributions and complexities. Social Science & Medicine 2010;70(9):1277-1284.
7. Charney E, English W. Candidate Genes and Political Behavior. American Political Science Review 2012;106(1):1-34.
8. Duncan LE, Keller MC. A Critical Review of the First Ten Years of Candidate Gene-by-Environment Interaction Research in Psychiatry. American Journal of Psychiatry 2011;168(10):1041-1049.
9. Eaves LJ. Genotype x environment interaction in psychopathology: Fact or artifact? Twin Research and Human Genetics 2006;9(1):1-8.
10. North KE, Martin LJ. The Importance of Gene-Environment Interaction Implications for Social Scientists. Sociological Methods & Research 2008;37(2):164-200.
11. Risch N, Herrell R, Lehner T, Liang K-Y, Eaves L, Hoh J, et al. Interaction Between the Serotonin Transporter Gene (5-HTTLPR), Stressful Life Events, and Risk of Depression: A Meta-analysis. Journal of the American Medical Association 2009;301(23):2462-2471.
12. Munafo MR, Durrant C, Lewis G, Flint J. Gene x Environment Interactions at the Serotonin Transporter Locus. Biological Psychiatry 2009;65(3):211-219.
13. Perrin AJ, Lee H. The undertheorized environment: Sociological theory and the ontology of behavioral genetics. Sociological Perspectives 2007;50(2):303-322.
14. Ottman R. Gene-environment interaction: Definitions and study designs. Preventive Medicine 1996;25(6):764-770.
15. Cornelis MC, Tchetgen EJT, Liang LM, Qi L, Chatterjee N, Hu FB, et al. Gene-Environment Interactions in Genome-Wide Association Studies: A Comparative Study of Tests Applied to Empirical Studies of Type 2 Diabetes. American Journal of Epidemiology 2012;175(3):191-202.
16. Beaty TH, Ruczinski I, Murray JC, Marazita ML, Munger RG, Hetmanski JB, et al. Evidence for Gene-Environment Interaction in a Genome Wide Study of Nonsyndromic Cleft Palate. Genetic Epidemiology 2011;35(6):469-478.
17. Link BG, Phelan J. Social Conditions as Fundamental Causes of Disease. Journal of Health and Social Behavior 1995:80-94.
18. Edwards TL, Edwards DRV, Villegas R, Cohen SS, Buchowski MS, Fowke JH, et al. HTR1B, ADIPOR1, PPARGC1A, and CYP19A1 and Obesity in a Cohort of Caucasians and African Americans: An Evaluation of Gene-Environment Interactions and Candidate Genes. American Journal of Epidemiology 2012;175(1):11-21.

19. Wareham NJ, Young EH, Loos RJF. Epidemiological Study Designs to Investigate Gene-Behavior Interactions in the Context of Human Obesity. Obesity 2008;16:S66-S71.
20. Bouchard C, Agurs-Collins T. Studying Gene-Behavior Interactions: Summary of Recommendations. Obesity 2008;16:S95-S96.
21. Berkman LF, Kawachi I. Social Epidemiology. Oxford: Oxford University Press; 2000.
22. Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. International Journal of Epidemiology 2001;30(4):668-677.
23. Krieger N. Epidemiology and the Web of Causation - Has Anyone Seen the Spider. Social Science & Medicine 1994;39(7):887-903.
24. Williams GH. The determinants of health: structure, context and agency. Sociology of Health & Illness 2003;25:131-154.
25. Schnittker J. Education and the changing shape of the income gradient in health. Journal of Health and Social Behavior 2004;45(3):286-305.
26. Link BG. Epidemiological Sociology and the Social Shaping of Population Health. Journal of Health and Social Behavior 2008;49(4):367-384.
27. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: A multilevel study of collective efficacy. Science 1997;277(5328):918-924.
28. Boardman JD, Barnes LL, Wilson RS, Evans DA, de Leon CFM. Social disorder, APOE-E4 genotype, and change in cognitive function among older adults living in Chicago. Social Science & Medicine 2012;74(10):1584-1590.
29. Omi M, Winant H. Racial Formation in the United States: From the 1960s to the 1990s: Routledge; 1994.
30. Turkheimer E, Haley A, Waldron M, D'Onofrio B, Gottesman II. Socioeconomic Status Modifies Heritability of IQ in Young Children. Psychological Science 2003;14(6):623-628.
31. Rowe DC, Jacobson KC, Van den Oord E. Genetic and environmental influences on vocabulary IQ: Parental education level as moderator. Child Development 1999;70(5):1151-1162.
32. Shanahan MJ, Boardman JD. Genetics and Behavior in the Life Course: A Promising Frontier. In: Elder GH, Jr., Giele JZ, editors. The Craft of Life Course Research. New York: The Guilford Press; 2009.
33. Burt A. Some Key Issues in the Study of Gene-Environment Interplay: Activation, Deactivation, and the Role of Development. Research in Human Development 2011;8(3-4):192-210.
34. Wanke KL, Spittel ML. Advancing Research in Gene-Environment Interplay: Can Developmental Science Lead the Way? Research in Human Development 2011;8(3-4):165-172.
35. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Social Science & Medicine 2000;51(6):843-857.
36. Portes A. Social Capital: Its origins and applications in modern sociology. Annual Review of Sociology 1998;24:1-24.
37. Coleman JS. Social Capital in the Creation of Human-Capital. American Journal of Sociology 1988;94:S95-S120.
38. Kawachi I, Kennedy BP, Lochner K, ProthrowStith D. Social capital, income inequality, and mortality. American Journal of Public Health 1997;87(9):1491-1498.
39. Boardman JD, Finch BK, Ellison CG, Williams DR, Jackson JS. Neighborhood disadvantage, stress, and drug use among adults. Journal of Health and Social Behavior 2001;42(2):151-165.

40. Boardman JD, Saint Onge JM, Rogers RG, Denney JT. Race differentials in obesity: The impact of place. Journal of Health and Social Behavior 2005;46(3):229-243.
41. Massey DS, Denton NA. The Dimensions of Residential Segregation. Social Forces 1988;67(2):281-315.
42. Dick DM. Gene-Environment Interaction in Psychological Traits and Disorders. In: NolenHoeksema S, Cannon TD, Widiger T, editors. Annual Review of Clinical Psychology. Palo Alto: Annual Reviews; 2011. p. 383-409.
43. Dick DM, Agrawal A, Schuckit MA, Bierut L, Hinrichs A, Fox L, et al. Marital status, alcohol dependence, and GABRA2: Evidence for gene-environment correlation and interaction. Journal of Studies on Alcohol 2006;67(2):185-194.
44. Boardman JD, Roettger ME, Domingue BW, McQueen MB, Haberstick BC, Harris KM. Gene-environment interactions related to body mass: school policies and social context as environmental moderators. Journal of Theoretical Politics 2012;24(3):370-388.
45. Silventoinen K, Hasselbalch AL, Lallukka T, Bogl L, Pietilainen KH, Heitmann BL, et al. Modification effects of physical activity and protein intake on heritability of body size and composition. American Journal of Clinical Nutrition 2009;90(4):1096-1103.
46. Gottlieb DJ, Wilk JB, Harmon M, Evans JC, Joost O, Levy D, et al. Heritability of longitudinal change in lung function - The Framingham Study. American Journal of Respiratory and Critical Care Medicine 2001;164(9):1655-1659.
47. Dick DM. An Interdisciplinary Approach to Studying Gene-Environment Interactions: From Twin Studies to Gene Identification and Back. Research in Human Development 2011;8(3-4):211-226.
48. Karg K, Burmeister M, Shedden K, Sen S. The Serotonin Transporter Promoter Variant (5-HTTLPR), Stress, and Depression Meta-Analysis Revisited. Archives of General Psychiatry 2011;68(5):444-454.
49. Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, et al. Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. Science 2003;301(5631):386-389.
50. Mitchell C, Notterman D, Brooks-Gunn J, Hobcraft J, Garfinkel I, Jaeger K, et al. Role of Mother's Genes and Environment in Postpartum Depression. Proceedings of the National Academy of Sciences of the United States of America 2011;108:8189-8193.
51. Cleveland HH. Disadvantaged neighborhoods and adolescent aggression: Behavioral genetic evidence of contextual effects. Journal of Research on Adolescence 2003;13(2):211-238.
52. Boardman JD, Saint Onge JM, Haberstick BC, Timberlake DS, Hewitt JK. Do schools moderate the genetic determinants of smoking? Behavior Genetics 2008;38(3):234-246.
53. Boardman JD. State-Level Moderation of Genetic Tendencies to Smoke. American Journal of Public Health 2009;99(3):480-486.
54. Boardman JD, Blalock CL, Pampel FC, Hatemi PK, Heath AC, Eaves LJ. Population Composition, Public Policy, and the Genetics of Smoking. Demography 2011;48(4):1517-1533.
55. Adkins DE, Guo G. Societal Development and Shifting Influence of the Genome on Status Attainment. Research in Social Stratification and Mobility 2008;26:235-256.
56. Wang SH, Chen WJ, Chuang LM, Hsiao PC, Liu PH, Hsiao CK. Inference of Cross-Level Interaction between Genes and Contextual Factors in a Matched Case-Control Metabolic Syndrome Study: A Bayesian Approach. PLoS One 2013;8(2).

57. Shanahan MJ, Hofer SM. Social context in gene-environment interactions: Retrospect and prospect. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 2005;60:65-76.
58. Belsky J, Pluess M. Beyond Diathesis Stress: Differential Susceptibility to Environmental Influences. Psychological Bulletin 2009;135(6):885-908.
59. Belsky J, Pluess M. The Nature (and Nurture?) of Plasticity in Early Human Development. Perspectives on Psychological Science 2009;4(4):345-351.
60. Ellis BJ, Boyce WT, Belsky J, Bakermans-Kranenburg MJ, van Ijzendoorn MH. Differential susceptibility to the environment: An evolutionary-neurodevelopmental theory. Development and Psychopathology 2011;23(1):7-28.
61. Simons RL, Lei MK, Beach SRH, Brody GH, Philibert RA, Gibbons FX. Social Environment, Genes, and Aggression: Evidence Supporting the Differential Susceptibility Perspective. American Sociological Review 2011;76(6):883-912.
62. Caspi A, McClay J, Moffitt TE, Mill J, Martin J, Craig IW, et al. Role of genotype in the cycle of violence in maltreated children. Science 2002;297(5582):851-854.
63. Koopmans JR, Slutske WS, van Baal GCM, Boomsma DI. The influence of religion on alcohol use initiation: Evidence for genotype x environment interaction. Behavior Genetics 1999;29(6):445-453.
64. Fletcher JM. Why Have Tobacco Control Policies Stalled? Using Genetic Moderation to Examine Policy Impacts. PLOS One 2012;7(12):e50576.
65. Saccone SF, Hinrichs AL, Saccone NL, Chase GA, Konvicka K, Madden PAF, et al. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Human Molecular Genetics 2007;16(1):36-49.
66. Daw J, Shanahan M, Harris KM, Smolen A, Haberstick B, Boardman JD. Genetic Sensitivity to Peer Behaviors: 5HTTLPR, Smoking, and Alcohol Consumption. Journal of Health and Social Behavior 2013;54(1):92-108.
67. Pluess M, Belsky J. Conceptual Issues in Psychiatric Gene-Environment Interaction Research. American Journal of Psychiatry 2012;169(2):222-223.
68. Raine A. Biosocial studies of antisocial and violent behavior in children and adults: A review. Journal of Abnormal Child Psychology 2002;30(4):311-326.
69. Tuvblad C, Grann M, Lichtenstein P. Heritability for adolescent antisocial behavior differs with socioeconomic status: gene-environment interaction. Journal of Child Psychology and Psychiatry 2006;47(7):734-743.
70. Plomin R, Defries JC, Loehlin JC. Genotype-Environment Interaction and Correlation in Analysis of Human-Behavior. Psychological Bulletin 1977;84(2):309-322.
71. Price TS, Jaffee SR. Effects of the Family Environment: Gene-Environment Interaction and Passive Gene-Environment Correlation. Developmental Psychology 2008;44:305-315.
72. Branigan AR, Freese J, Patir A, McDade TW, Liu K, Kiefe C. Skin Color, Sex, and Educational Attainment in the Post-Civil-Rights Era. In: Meeting of the Research Committee on Social Stratification and Mobility (RC28) of the International Sociological Association. Colchester, UK; 2011.
73. Stovel K, Savage M, Bearman P. Ascription into achievement: Models of career systems at Lloyds Bank, 1890-1970. American Journal of Sociology 1996;102(2):358-399.
74. Reskin BF, McBrier DB. Why not ascription? Organizations' employment of male and female managers. American Sociological Review 2000;65(2):210-233.

75. Nielsen F. Achievement and ascription in educational attainment: Genetic and environmental influences on adolescent schooling. Social Forces 2006;85(1):193-216.
76. Cleveland HH, Wiebe RP, Rowe DC. Sources of exposure to smoking and drinking friends among adolescents: A nehavioral-genetic evaluation. Journal of Genetic Psychology 2005;166(2):153-169.
77. Jaffee SR, Price TS. Gene-environment correlations: a review of the evidence and implications for prevention of mental illness. Molecular Psychiatry 2007;12(5):432-442.
78. Fletcher JM. Enhancing the Gene-Environment Interaction Framework: Evidence from differential responses to September 11th. In: Annual Meeting of the Population Association of America. Washington, D.C.; 2011.
79. Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing "neighborhood effects": Social processes and new directions in research. Annual Review of Sociology 2002;28:443-478.
80. Theorell T. Working Conditions and Health. In: Berkman LF, Kawachi I, editors. Social Epidemiology. New York: Oxford University Press; 2000. p. 95-117.
81. Reidpath D, Burns C, Garrard J, Mahoney M, Townsend M. An ecological study of the relationship between socioeconomic status and obesogenic environments. Health and Place 2002;8:141-145.
82. Williams DR, Collins C. US Socioeconomic and Racial-Differences in Health - Patterns and Explanations. Annual Review of Sociology 1995;21:349-386.
83. Duster T. Comparative perspectives and competing explanations: Taking on the newly configured reductionist challenge to sociology. American Sociological Review 2006;71(1):1-15.
84. Hartz SM, Short SE, Saccone NL, Culverhouse R, Chen LS, Schwantes-An TH, et al. Increased Genetic Vulnerability to Smoking at CHRNA5 in Early-Onset Smokers. Archives of General Psychiatry 2012;69(8):854-861.
85. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet 2003;361(9357):598.
86. Rutter M, Moffitt TE, Caspi A. Gene-environment interplay and psychopathology: multiple varieties but real effects. Journal of Child Psychology and Psychiatry 2006;47(3-4):226-261.
87. Massey DS, Rothwell J, Domina T. The Changing Bases of Segregation in the United States. Annals of the American Academy of Political and Social Science 2009;626:74-90.
88. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. American Journal of Human Genetics 2000;67(1):170-181.
89. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. 2006:in press.
90. Daw J, Guo G. The influence of three genes on whether adolescents use contraception, USA 1994—2002. Population Studies 2011;65(3):253-271.
91. Fletcher JM, Lehrer SF. Genetic lotteries within families. Journal of Health Economics 2011;30(4):647-659.
92. Wang KS, Liu XF, Aragam N, Jian XQ, Mullersman JE, Liu YL, et al. Family-based association analysis of alcohol dependence in the COGA sample and replication in the Australian twin-family study. Journal of Neural Transmission 2011;118(9):1293-1299.
93. Blumer H. Symbolic interactionism: Perspective and method. Englewood Cliffs, NJ: Prentice-Hall; 1969.

Figure 1. Models of Gene-Environment Interactions