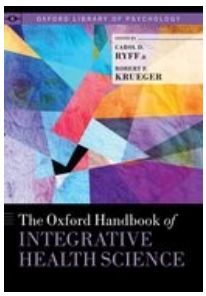

The Oxford Handbook of Integrative Health Science

Carol D. Ryff (ed.), Robert F. Krueger (ed.)

https://doi.org/10.1093/oxfordhb/9780190676384.001.0001

Published online: 09 October 2018 **Published in print:** 29 November 2018 **Online ISBN:** 9780190676407 **Print ISBN:** 9780190676384

CHAPTER

3 The Roles of Twin Studies and Modern Genomic Technologies in Integrative Health Science

Robert F. Krueger, Susan C. South

https://doi.org/10.1093/oxfordhb/9780190676384.013.28 Pages 35–46

Published: 09 October 2018

Abstract

This chapter focuses on genetically informative research design and strategy in integrative health science (IHS). A feature of IHS is studying individual differences in health outcomes together and in a multidisciplinary manner. The chapter focuses on the advantages of using genetically informative research strategies for IHS. Genetically informative strategies are tools to enrich inferences within the IHS paradigm. They help parse the meaning of observed associations between exposures and outcomes. Two strategies are considered for the Midlife in the United States study : (1) Gene × Environment interactions and (2) correlations between education and allostatic load. A strategy likely to be employed in IHS research involves using segments of RNA to understand mechanisms underlying health and illness, focusing on the conserved transcriptional response to adversity (CTRA). The conclusion is that IHS and genetically informative research strategies are natural allies in understanding origins of health and illness in the population at large.

Keywords: integrative health science, MIDUS study, twins, conscientiousness, alcohol use, RNA, health, illness, CTRA, genetically informative research

Subject: Health Psychology, Psychology

Series: Oxford Library of Psychology

Collection: Oxford Handbooks Online

Introduction

Genetically informative research is often conceptualized as a distinct field within the health sciences. There may be a number of reasons for this situation. One reason likely pertains to the methods of genetic research. Often, these methods appear highly specialized to investigators not directly involved in genetic research. Another reason may pertain to the aims of genetic research. Genetic research often involves aims that are specifically genetic in nature, such as hunting for polymorphisms (different forms of specific genes) that are relevant to health outcomes.

Our aim in this chapter is to explore a somewhat different perspective on genetically informative research. Specifically, we explore the perspective that genetically informative research provides strategies that are highly useful in "normal integrative health science (IHS) inquiry." What we mean by this is that methods and ideas that have emerged in specialized genetic contexts have broad applicability in furthering health research, beyond their specialized applications in genetics. With this general perspective in mind, we have two specific goals in this chapter. First, we hope to convey some of the methodological ideas that have emerged in behavior genetic research and explain why we feel these ideas have broad applicability in health research. Second, we try to peer just over the horizon at potential developments in genomics; we aim to describe how these developments might be helpful in typical IHS inquiry. For example, as genetic polymorphisms relevant to health and behavior are identified, they can be studied in largely any context, versus being topics of study only in a specialized genetics literature. Of course, for this strategy to work well,
p. 36 ↳ the effects of these polymorphisms need to be well understood, their assessment needs to be made reliably, investigators with diverse backgrounds need to know how to work with genetic data, and so on. In spite of these clear challenges, we sense that the field is on the cusp of potentially exciting developments along these lines. As such, our aim is to convey potential near-future scenarios to readers not working in genetics explicitly.

In this chapter, we focus in particular on the Midlife in the United States (MIDUS) project. We do this for at least two reasons. First, MIDUS is relatively unique among large-scale epidemiological studies of health in including a national sample of twins within its broader sample of participants. Many studies of human health outcomes—for entirely understandable reasons having to do with allocation of finite resources—focus specifically on twins or specifically on unrelated research participants (but rarely both). MIDUS is unique in attempting to pursue both participant strategies simultaneously, a strategy that has allowed for some novel contributions to the literature, in our estimation. Second, MIDUS is well positioned to be at the forefront of the kind of normal IHS inquiry approach to genetics we aim to articulate here. Rather than beginning anew with novel research participants, ongoing MIDUS research in which we and others are involved aims to augment existing MIDUS data with new data at the forefront of contemporary genomic technologies. We aim to give the reader some exposure to these directions, with the idea in mind of inviting readers to consider working directly with MIDUS data that may be pertinent to their own specific interests.

The Relevance of Twin Research to Integrative Health Science

This volume is focused specifically on IHS. We see IHS as an effort to study individual differences in health outcomes in concert, as opposed to the more traditional strategy of dividing efforts among putatively separable health outcomes, in separate literatures. In our view, the study of human twins is a key aspect of approaching and studying health outcomes in a more integrated way. Traditionally, twin studies have been thought of as relevant specifically to documenting *heritability*. Heritability is an important concept and refers to the proportion of phenotypic (observed) variation in a phenotype (a specific observable characteristic of an organism) attributable to genetic variation. Documenting heritability was a critically important accomplishment of twentieth-century research on twins. Prior to that time, many key individual differences were thought to be entirely environmental in origin, based on dominant theories about human development. An extensive body of literature, however, shows that diverse human individual differences are partly attributable to genetic differences (Polderman et al., 2015).

It might be thought that the relevance of twins to research on human individual differences is diminished in the contemporary era, given that the evidence for heritable contributions to human phenotypes is extensive. To understand why we think this is not the case, it may help to provide some simple equations to describe the quantitative concepts underlying heritability estimation. Stated mathematically, phenotypic variation P is understood to be the sum of genetic G and environmental variation E:

$$P = G + E \tag{3.1}$$

Following from this equation and our given definition of heritability *H*, we can specify heritability as the proportion of P (phenotypic variation) attributable to *G* (genetic variation):

$$H = G/P \tag{3.2}$$

For us, thinking about these equations leads to a couple of questions that help frame much of our research efforts. First, is G necessarily a constant for the entire population? The equation is set up in that way, but this is an assumption, not a necessity. A common misperception of the heritability statistic by lay individuals is that the estimate is individualized; that is, if heritability is .50, then one half of the "cause" of an outcome (e.g., a disorder) is due to genetics. But, our equations clarify that heritability is a means of parsing variance in an outcome within a population.

Recently, we have begun to ask if certain segments of the population may differ in the relative influence of G and E on P. If this possibility could be modeled empirically, it might lead to potential insights about how variables not yet introduced into this system might affect the system. For example, the system presented contains only one phenotype and does not take into account potential *moderators* external to the system as conceptualized in Equations 3.1 and 3.2 (e.g., socioeconomic status [SES]) that might affect P, G, and E and therefore H.

p. 37 Second, what is the meaning of E? E is often thought about as random influences left over net of G, where estimating G is the primary purpose of genetic inquiry. However, E is also *variation in P that is net of G*. This is not a trivial realization, in our view. One reason for its nontriviality relates to the logic of causal reasoning in observational research. Specifically, better handles on causality can be obtained through *counterfactual reasoning*.

The essence of counterfactual reasoning involves conceptualizing what might or could have happened, had some key variables been different, net of other relevant variables. For example, if something is amenable to experimental manipulation, random assignment of people to conditions allows for counterfactual reasoning. If we assign people randomly to conditions (e.g., receive active drug vs. receive placebo), we are trying to equate all potential confounding variables, such that all that varies is the receipt of drug versus placebo. If random assignment was successful and the drug condition is associated with some desirable outcome (but not the placebo condition), we conclude that the drug had a causal impact on the desirable outcome. We conclude (via counterfactual reasoning) that, if the placebo participants had received the drug, they also would have experienced the desirable outcome (and vice versa, i.e., had the drug condition participants *not* received the drug, they would *not* have experienced the desirable outcome).

With this thinking in mind, we emphasize how being able to hold G constant is not a trivial achievement in trying to understand the meaning of correlations between exposures and outcomes. This is particularly true when variables of critical public health relevance are not amenable to direct manipulation (e.g., many of the core constructs in social scientific health inquiry, such as SES). For example, comparing monozygotic (MZ) twins who differ on realized adulthood SES within pairs provides a comparison of persons who are matched on underlying genotype (and typically grew up in the same household). If health outcomes are worse in lower SES twins, within pairs where twins vary in SES, this evidence enhances our confidence in the idea that the SES–health association is at least partly mediated via a directly causal environmental pathway that is not confounded with genetic sources of variation.

Of course, the "natural twin experiment" is not perfect. For example, twins within pairs are not entirely randomly assigned to environmental exposures. Nevertheless, confidence in the idea that there exists some degree of environmentally mediated causality in the link between exposures and outcomes is enhanced by evidence that, in twins that differ in exposure, there are corresponding differences in outcomes. For example, one of the reasons smoking is thought to be causally related to diminished health outcomes is that, in twin pairs discordant for smoking, the smoking twin tends to have worse health outcomes (e.g., Kaprio & Koskenvuo, 1990). This design is helpful in enhancing our certainty about direct and environmentally mediated causality because, by studying twin pairs, we are controlling for a host of potential confounding third variables (most notably, underlying genotypes). This control is made possible because twins are matched on underlying genotypes, more so than unrelated research participants (and perfectly matched within MZ pairs).

Moreover, at the very least, such twin evidence helps in thinking about the plausibility of other potential models. For example, perhaps people with certain genotypes are more likely to smoke and to experience adverse health outcomes, where both smoking behavior and adverse health outcomes are the causal result of the same underlying genotypes. This model involves a type of *active gene–environment correlation.* In this scenario, specific genotypes contribute to specific persons actively exposing themselves to smoking environments, and those genotypes also directly cause diminished health (and not via smoking exposure). Twin studies can be used to compare this kind of model (mediation of exposure–outcome correlations primarily through G) with a more direct environmental causality model (mediation of exposure–outcome correlations primarily through E). Indeed, these models are not mutually exclusive, and both paths may contribute to observed correlations between exposures and outcomes.

Moreover, much of what matters for health outcomes in the real world is not amenable to lasting or ethical manipulation in humans. Smoking again provides a good example; we cannot ethically randomly assign people to be heavy smokers, versus avoiding tobacco exposure. To illustrate this point further, consider early life adversities (ELAs). ELAs are known to be predictive of adult health outcomes, albeit this body of evidence is not always consistent or easily summarized (Chapter 5, this volume). Still, and obviously, ELAs cannot be manipulated directly or ethically. Hence, the observed correlation between ELAs and health outcomes is consistent with a variety of potential causal models. That correlation could mean that ↳ ELAs

directly cause adverse health outcomes via direct environmentally mediated pathways. It might also mean that genetic polymorphisms predictive of adverse family environments are manifest in both problematic parental behavior and physical health outcomes. For example, the parameters of neurophysiological systems are partly under genetic influence, and those systems may manifest in both adverse parenting behaviors and adverse health outcomes, in both parents and offspring. Tentative evidence for such a process comes from work using twins that demonstrates a nonzero correlation (r = .32) between genetic influences on childhood maltreatment and genetic influences on adult physical health (South, Schafer, & Ferraro, 2015).

As we hope this example makes clear, disentangling these sorts of potential explanations is highly challenging. Yet, it is also very important in public health. For example, if there is evidence for a direct environmentally mediated causal influence of ELAs on health outcomes, the public health implications are clear: Resources should be invested in programs designed to reduce ELA exposure. Even if part of the correlation is still attributable to shared genetic influences on both ELAs and health outcomes, the presence of an additional directly causal environmental pathway strongly suggests the possibility of disrupting that pathway through environmental interventions. Conversely, the absence of that pathway suggests environmental intervention may be less effective than direct physiological intervention. For example, psychopharmacology directed at modifying relevant neurophysiological systems may be an effective intervention strategy in this scenario. Again, both pathways may be present and are not necessarily mutually exclusive. Nevertheless, understanding how correlations between risk factors and outcomes are mediated is essential in effectively deploying limited public health resources. We turn now to consider specific perspectives on the meaning of both G and E in greater detail, drawing on two specific examples from our work in MIDUS.

Gene × Environment Interaction (Biometrical Moderation): The Case of Conscientiousness as a Moderator of Problem Alcohol Use

In this chapter, our aim is to illustrate conceptual and empirical advantages of genetically informed research for IHS. Hence, rather than provide an extensive review of pertinent literature, we have selected specific empirical examples from our own work with MIDUS as illustrative projects. In addition, as we have reviewed the broader literature more thoroughly elsewhere recently, interested readers are encouraged to consult this review (South, Hamdi, & Krueger, 2017).

The first strategy we illustrate focuses on the use of twin data to model potential Gene × Environment interactions, the situation in which etiological influences (genetics) on an outcome may differ depending on certain situations (environments). This strategy is also known as *biometrical moderation.* The terminology clarifies that what is being moderated is not only genetic influences, but also environmental influences on the phenotype. The basic idea in this strategy is that G and E in Equation 3.1 are not constants for the entire population and instead differ systematically as a function of another *moderating* variable M, which is usually thought of as broadly "environmental" but can range from parenting to peer groups to personality traits.

This type of model is illustrated in Figure 3.1, which is reprinted from an article in which we explored the possibility that genetic and environmental contributions to alcohol problems might differ as a function of the personality trait of conscientiousness. We know from extensive research that lower conscientiousness is correlated with alcohol use problems (e.g., Kotov, Gamez, Schmidt, & Watson, 2010), but that work does not explicate how these two phenotypes are linked. Biometrical moderation using twin data can provide the answer to one possibility, namely, that one's level of conscientiousness may act as a "psychological environment" that alters the magnitude of genetic contributions to alcohol problems. Thus, the hypothesis we sought to test was that the magnitude of underlying genetic and environmental contributions to alcohol

problems were not necessarily constant for the entire population but might differ as a function of the domain of personality most relevant to individual differences in the tendency to constrain versus express impulses (conscientiousness).

In Figure 3.1, the G and E terms from Equation 3.1 are further specified. G is given as A to refer to ***A***dditive genetic influences; that is, genetic influences are here conceptualized as many specific polymorphisms that together "add up" to contribute to observed phenotypic variation. E is broken down into two types of environments: those that make people similar because they grew up in the same families (shared or ***C***ommon environmental influences) and those that make people different in spite of the fact that they grew
p. 39 up in the ↳ same families (nonshared or unique ***E***nvironmental influences).

In Figure 3.1, conscientiousness is both a phenotype and a moderator. That is, the model conceptualizes individual differences in conscientiousness as a phenotype, resulting from both genetic and environmental influences (as operationalized via A, C, and E paths leading directly to conscientiousness). These genetic and environmental influences on conscientiousness also lead directly to alcohol problems, and there are also genetic and environmental influences on alcohol problems specifically (beyond the influences in common with conscientiousness). For all of these A, C, and E paths leading to alcohol problems, genetic and environmental influences are not constant but are given as a function of both constant influences and influences that vary as a function of the level of the moderator (conscientiousness in this case, i.e., conscientiousness is the variable labeled M in the equations in Figure 3.1). This arrangement is operationalized by the betas in the paths leading to alcohol problems in this model. Consider, for example the path leading from A_c (genetic influences in common between both conscientiousness and alcohol problems). This path includes a constant genetic effect that does not differ as a function of the moderator (a_c) and an effect that *does* differ as a function of the moderator ($\beta_{Xac}M$). The beta in this equation determines how much weight to place on the level of the moderator (M; conscientiousness) in affecting the magnitude of genetic contributions to both conscientiousness and alcohol problems (and there are also unique and moderated genetic and environmental effects specific to alcohol problems, i.e., A_u, C_u, and E_u in Figure 3.1).

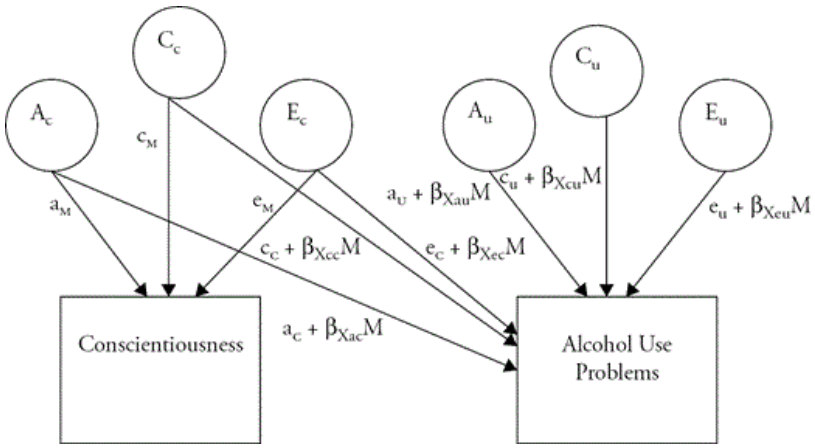

Figure 3.1 Biometrical model of the moderation of genetic and environmental variance in alcohol use problems by conscientiousness. A = Additive genetic effects; C = Common or shared environmental effects; E = Nonshared environmental effects. The subscript c indicates effects in common between conscientiousness and alcohol use problems. The subscript u indicates effects unique to alcohol use problems. The variable M represents the value of the moderator (conscientiousness in this case). The Greek letter beta (β) indicates the magnitude of a specific moderation effect. For example, β_{Xau} represents the impact of the moderator on the unique additive genetic variance in alcohol use problems as a function of the value of the moderator *M*.

Reprinted with permission from South, S. C., & Krueger, R. F. (2014). Genetic strategies for probing conscientiousness and its relationship to aging. *Developmental Psychology, 50,* 1362–1376.

Figure 3.2 shows the results reported in South and Krueger (2014) from fitting this model to data obtained from twin participants in the MIDUS project (please see the original publication for pertinent sample and methodological details). As reported, the moderation model that includes additional parameters to allow A, C, and E influences to differ fit better than the model without moderation (i.e., a model where all the betas in Figure 3.1 are set to zero). Figure 3.2 shows moderation using total variances (the left-hand side of Figure 3.2), and also using proportions of variance (the right-hand side of Figure 3.2). As Figure 3.2 illustrates, the key finding was that the total amount of nonshared environmental variance in alcohol problems (variance contributing to making people unique, in spite of growing up together) was *lower at higher* levels of conscientiousness. Because genetic variance (A) was relatively constant across levels of conscientiousness, and shared environmental variance (C) was essentially nil, the *proportion* of genetic variance (heritability) in alcohol problems was *higher* at higher levels of conscientiousness. Essentially, at higher levels of conscientiousness, alcohol problems were relatively more heritable (A was a greater proportion of total overall variation). Put somewhat differently, ↳ at higher levels of conscientiousness, genetic effects were
p. 40 expressed with greater clarity because there was less total variance in alcohol problems, attributable to less nonshared (unique) environmental perturbations.

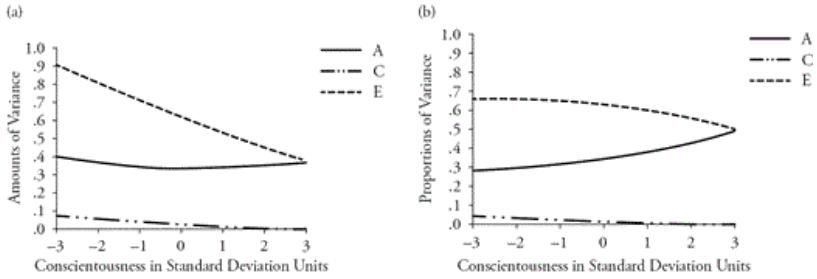

Figure 3.2 Model-derived decomposition of the genetic and environmental variance in alcohol use problems as a function of conscientiousness. A = Additive genetic effects; C = Common or shared environmental effects; E = Nonshared environmental effects.

Reprinted with permission from South, S. C., & Krueger, R. F. (2014). Genetic strategies for probing conscientiousness and its relationship to aging. *Developmental Psychology, 50*, 1362–1376.

Stepping back from this specific example, what does it illustrate about the broader utility of the biometrical moderation strategy in IHS? In our estimation, it shows the enhanced interpretive leverage made possible by studying samples of twins. Without twin observations, the model in Figure 3.1 would reduce to a correlation between conscientiousness and alcohol problems. Knowing these variables are correlated provides little in the way of potential reasons *why* conscientiousness and alcohol problems are correlated and, thereby, little in the way of uncovering processes that may underlie this kind of correlation. The findings portrayed in Figure 3.2, by contrast, suggest that high levels of conscientiousness are associated with less unique environmental perturbations in levels of alcohol problems. Genetic risk for alcohol problems is expressed with greater clarity at high levels of conscientiousness. This kind of finding suggests that genetic factors are a key reason people develop alcohol problems in spite of relatively higher conscientiousness. There are clearly individual differences in alcohol problems, even at high levels of conscientiousness, and one key reason for these individual differences is individual differences in genetic risk for alcohol problems.

Biometric moderation models using twin data may be particularly useful tools in the study of behavioral health because so many known mental and physical disorders have a nonzero heritability estimate, yet the specific loci behind these estimates are still being sought. For most disorders, we are still quite far from being able to use genotype data to predict a health outcome with any certainty (Kong et al., 2015). But with biometric moderation, we can come ever closer to a personalized estimate of genetic and environmental risk for a particular health outcome, based on an individual's standing on the moderator variable of interest. In sum, the potential of IHS to inform our understanding of health issues such as alcohol problems is notably enhanced by studying twins.

Co-Twin Control: The Case of Education and Allostatic Load

Focusing specifically on effects within families provides another example of the utility of twins in IHS. As the example of conscientiousness and alcohol problems illustrates, G (genetic effects) are not typically 100% of P (phenotypic variation). Moreover, often, a substantial part of P is attributable to E (nonshared environments) and not C (shared environments). For example, little of the variation in alcohol problems was attributable to C (see Figure 3.2). By contrast, much of the variance in alcohol problems was attributable to E (particularly at lower levels of conscientiousness). How might these types of observations be leveraged to better understand the processes underlying observed correlations?

In addition to the biometrical moderation approach described, another way of thinking about and modeling twin data involves *co-twin control approaches*. These approaches focus on E (nonshared environmental
p. 41 variation) and the ability ↳ to use twin observations to control for A (additive genetic effects) and C (shared environmental effects). As noted, twins are matched on genetic polymorphisms, perfectly in the case of identical or MZ twins (at the level of the underlying nucleotide sequence, i.e., the series of "letters" that make up the underlying human DNA code). In addition, if twins were raised in the same household, they are matched on objective features of that situation (e.g., the objective qualities of the house they grew up in together, such as the neighborhood in which that house was located).

Importantly, however, even identical twins are not phenotypically identical in adulthood. This observation is critically important in understanding the co-twin control approach. Essentially, twins can be used as "controls" for each other, within families. This approach controls for genetic variation and for shared environmental variation, leaving "purified" nonshared environmental variation. The idea here is very similar to other approaches in epidemiology, such as propensity score matching, or any other approach that aims to introduce numerous covariates in an attempt to help ensure that a correlation is not attributable to confounding by any of those covariates. However, co-twin control has the advantage of controlling for not only the covariates that one might think to include in an analysis, but also all unobserved sources of variation that are embedded in A and C effects.

To illustrate the value of this approach, we next describe a specific co-twin control study pursued in MIDUS. In our estimation, one of the most fascinating findings in IHS is the correlation between lower SES and poorer health (see, e.g., Chapter 31, this volume). As many chapters in this volume convey, this correlation is surprisingly robust and can be observed across numerous measures of health and across numerous potentially moderating variables. Nevertheless, the underlying reasons why this correlation is observed are challenging to uncover, at the level of an etiologic or causal understanding. For example, might the correlation be driven at least in part by direct environmental causation, such that enhanced SES (e.g., increased education) directly causes improvements in health? Or, by contrast, is the correlation attributable to background family factors (potentially both genetic and in the shared environment) that lead to an association between SES and health?

A number of co-twin control studies have been conducted to address these possibilities and were reviewed recently by Hamdi, South, and Krueger (2016). These studies reached somewhat mixed conclusions regarding the etiology of the association between SES and health, albeit with a number of them suggesting the link is not directly environmentally mediated. Also notable is the extent to which the existing twin literature has focused on survey, registry, and hospital record data in characterizing the health phenotype. More direct observations of health status (e.g., biomarkers) were generally absent from the extant twin literature. As a result, we (Hamdi et al., 2016) undertook a co-twin control study of years of education and allostatic load (McEwen, 2000, 2017) in MIDUS twins (64 fraternal and 81 identical pairs, along with 89 twin persons whose twin brother or sister did not participate). Allostatic load was operationalized as the sum of seven physiological indices. Each of the seven indices was based on between two and five biomarkers (or directly observed health indices such as body mass index [BMI]). Together, these indices covered sympathetic, parasympathetic, hypothalamic–pituitary–adrenal (HPA) axis, inflammation, cardiovascular, glucose, and lipid systems.

We observed a notable and significant zero-order negative phenotypic correlation between this index of allostatic load and years of education in this sample (r = -.22). That is, at the level of twin observed in this research, more years of education corresponded with lower allostatic load (thereby replicating the often-obtained finding that higher SES corresponds with better health). We then decomposed this correlation into both between-family and within-family effects (net of sex and age, which were used as covariates in this analysis). Between-family effects correspond essentially to the sum of unmoderated A and C effects in Figure 3.1, whereas within-family effects correspond to the unmoderated E effect in Figure 3.1. We found

that the within-family effect was essentially zero, such that the entire phenotypic effect we observed was mediated through between-family individual differences. That is, there was no evidence in this study of a direct, causal, and environmentally mediated link between years of education and allostatic load. Instead, at least in this project, effects at the family level mediated the observed relationship between years of education and allostatic load. These between-family effects could be either at the level of shared environments or genetics (albeit twin correlations for both education and allostatic load in this project were
p. 42 notably higher in identical vs. fraternal twins, suggesting the possibility the ↳ connection between these phenotypes is genetic in nature).

We again step back at this point and contemplate the broader relevance of this specific project for IHS. In our view, this project illustrates the considerably enhanced interpretive leverage made possible by studying twins. The correlation between lower SES and worse health is observed frequently and often is robust even in the presence of a variety of potential covariates. Co-twin control approaches, however, control not just for measured and selected covariates. Instead, these approaches control for unobserved family-level confounds, whether genetic or shared family environments. In exerting these controls, the hypothesis of direct environmental causality (often implicit in discussions of the impact of SES on health) can be evaluated.

Somewhat provocatively, our findings for education and allostatic load suggest that direct environmental causality is unlikely to be the explanation of the link (consistent with much—albeit not all—other twin literature reviewed by Hamdi et al., 2016). Rather than direct environmental causality (mediation via the nonshared environment), our findings are instead compatible with genetic or shared environmental (family-level) factors being responsible for the association between education and allostatic load. For example, genetic polymorphisms that are associated with diminished health may also manifest in diminished socioeconomic status. Indeed, emerging molecular genetic evidence pertinent to understanding the link between education and longevity is compatible with this perspective (Marioni et al., 2016).

Emerging Directions at the Intersection of Contemporary Genomic Technologies and Integrative Health Science

We hope the examples described illustrate how twin research contributes vitally to IHS. Interestingly, some have speculated that the advent of modern genomic technologies will render twin studies obsolete. However, understanding the inferences made possible by twin studies helps explain why that is not a likely scenario. Again, our point is that working with twins allows much more compelling inferences, beyond heritability estimation and of the sorts we described. Indeed, marrying contemporary genomic technologies with twin study designs provides enhanced interpretive leverage, beyond what is possible with either approach alone.

Next, we describe a contemporary genomic approach that is likely to have utility in the next phase of projects in IHS, such as the MIDUS project. We focus specifically on using individual differences in processes whereby DNA is transcribed into proteins (via RNA) to understand biological processes underlying health and illness. In the context of this strategy, we again emphasize how twins are relevant in improving inferences about the nature of associations between risk factors and outcomes.

Using Gene Expression Data to Understand Mechanisms of Health and Illness: The Conserved Transcriptional Response to Adversity (CTRA)

Health and illness can be conceptualized in a variety of complementary ways, from subjective reports of overall health and well-being, to biomarkers of functioning in specific physiological systems. Technologies relevant to characterizing physiological functioning and biological mechanisms relevant to those functions continue to grow by leaps and bounds. Relatedly, technologies used in characterizing genetic mechanisms are often thought of as relevant to identifying underlying segments of DNA that might be predictive of risk for disease. Indeed, the genome-wide association study (GWAS) approach, in which single nucleotide polymorphisms (SNPs) across the genome are scanned for potential associations with health-relevant phenotypes, has been generally successful at identifying thousands of variants relevant to health and illness. This is particularly true as sample sizes have increased into the hundreds of thousands of participants and methods for synthesizing the results of multiple studies at the summary data level have become more tractable (Pasaniuc & Price, 2017).

Nevertheless, GWAS strategies are only one way of leveraging genomic technologies in IHS. Other relatively newer methods for genome-wide transcriptional profiling and related bioinformatic interpretive strategies can be used to study how the expression profiles of genes central to the body's immune-inflammatory response are affected by social contexts. This type of approach relies not on sequencing DNA, but rather on studying RNA transcripts. This approach focuses on *gene expression*, which can be thought of as the fundamental biological process whereby environmental events (e.g., injury) lead to the expression of genes relevant to health and illness.

p. 43 Environmental events extend beyond physical insults. Indeed, a body of evidence shows how ↳ adverse psychosocial conditions such as low SES and social isolation affect the expression of several specific sets of genes relevant to health outcomes. These studies have identified a conserved transcriptional response to adversity" (CTRA) marked by upregulation of pro-inflammatory genes and downregulation of genes involved in Type I interferon responses and production of specific antibody isotypes (Cole, 2014). Importantly, this type of research has yet to be conducted in nationally representative samples and can benefit substantially from expansion to a broader range of potentially relevant psychosocial risk factors. As a result, we are pursuing this approach in ongoing MIDUS research. The basic framework for this approach is seen in Figure 3.3. Most research in IHS has (often by necessity) been at the level of observations above the oval on Figure 3.3, associating, for example, aspects of the environment with health outcomes. The novelty of this approach in MIDUS lies in getting "under the skin," with the aim of more directly studying the gene expression processes that mediate psychosocial–outcome associations (processes portrayed within the oval on Figure 3.3, in particular transcript abundance).

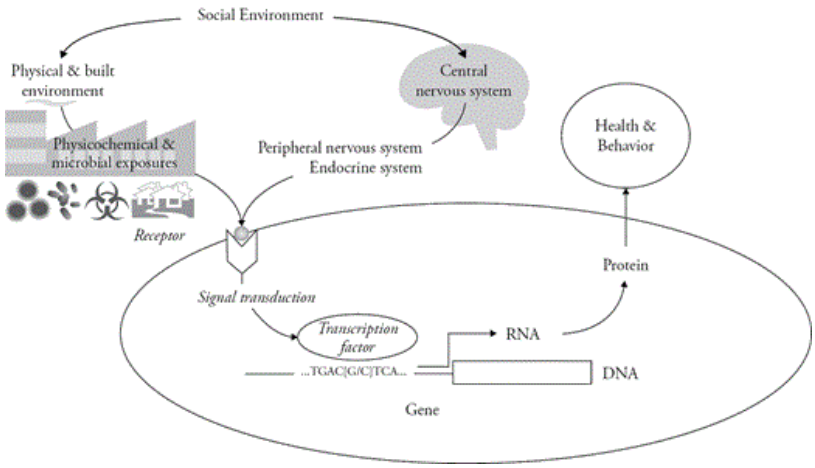

Figure 3.3 A model of mechanisms in the biological pathways linking the social environment with health and behavior.

Reprinted with permission from Cole, S. W. (2014). Human social genomics. *PLoS Genetics, 10*, e1004601.

MIDUS provides numerous advantages as a context for studying the CTRA. For example, much of the relevant literature has focused on psychosocial disadvantage. MIDUS allows the ability to study both advantage and disadvantage simultaneously because (as this volume makes clear) MIDUS contains numerous assessments over time, of both risk and resilience variables. Hence, we are working to understand how psychosocial *advantages* might ameliorate the impact of adverse life experiences. For example, in earlier work using data from married twins in MIDUS, we showed that genetic influences on self-reported physical health were highest among those with both very satisfying and very unsatisfying marriages (South & Krueger, 2013), suggesting that the same genes that convey liability to poor health in the wrong conditions may also provide a buffer against poor health when one is surround by an advantaged environment. More recent work in this vein showed that individuals who lost a spouse were buffered against subsequent depressive symptoms if they were higher in a polygenic risk score associated with subjective well-being (Domingue, Liu, Okbay, & Belsky. 2017). In sum, the overarching aim of our current MIDUS efforts involve integrating gene expression profiling and related bioinformatic assessment of CTRA transcriptome profiles into the rich knowledge of psychosocial aging afforded by MIDUS.

Additionally, some participants in this project are members of identical twin pairs. Their presence provides the ability to see if associations between the CTRA and psychosocial risks, which are documented using
p. 44 unrelated research participants, are environmentally mediated. Similar to the ↳ literature on education and allostatic load, it is frequently assumed that psychosocial risks cause the CTRA in a direct and environmentally mediated manner. However, it is also possible that the CTRA and psychosocial risks are correlated for genetic reasons. For example, persons at relatively higher genetic risk for adverse health outcomes might select into environments that further enhance that risk, a kind of active gene–environment correlation process. By studying pairs of identical twins, support for these possibilities can be evaluated, similar to our aforementioned study of education and allostatic load.

Future Directions

We conclude this chapter by briefly sketching select future directions in genetics and IHS. Our own work in this area will focus on using MIDUS to further contextualize the CTRA, as described in this chapter. Ultimately, however, the forthcoming MIDUS RNA data might also be used for discovery science. Beyond the focus on the specific genes that are relevant to the CTRA, other transcripts may show meaningful variance in the MIDUS sample. Hence, there is the possibility of discovering novel associations between specific transcripts and other specific MIDUS variables. The size and scope of the MIDUS endeavor makes this sort of discovery science an actual possibility because the projected sample is considerably larger than most samples in the human social genomic literature. Hence, there is the possibility of documenting novel associations, controlling the false discovery rate (FDR).

Possibilities also exist at the level of DNA and family designs. Although identical twins are identical at the level of the underlying nucleotide (DNA) sequence, other sibling pairs differ (e.g., fraternal twins and other siblings). This provides a powerful and compelling approach to controlling for family background factors in studying DNA–phenotype associations (see, e.g., Ott, Kamatani, & Lathrop, 2011).

Indeed, the "classical" twin design likely has continued utility in other areas touched on in the current volume and specifically within the context of the MIDUS research endeavor. We point to Turiano and colleagues (Chapter 22, this volume), who focused on the association between personality and health behaviors. As they note in their section on future directions, the twin design provides novel perspectives on the association between personality and health-diminishing behaviors, such as substance use and abuse. Some connections are likely genetic, and others are likely environmental. For example, some of the variation in substance use per se is likely directly attributable to the same genetic variation that affects relevant personality dispositions, such as disinhibition. Other variation, however, is likely more environmental. Consider, for example, the impact of availability of specific substances at the microenvironmental level and broader factors such as taxation at the macro level. Through the use of twin and other family designs, it is possible to begin to parse individual differences, such as personality and substance use behaviors, into these distinguishable proportions and to begin to trace the specific pathways connecting pieces of the broader IHS puzzle.

In sum, we hope we have illustrated how MIDUS and other IHS endeavors benefit notably from studying twins, in particular from working to merge family and twin designs with cutting-edge genomic technologies. Rather than outliving their usefulness in an age of genomic discovery, our view is that genomic science and twin–family modeling approaches are naturally aligned. Ultimately, these approaches provide tools that we see as critically relevant in moving from documenting correlations to being able to articulate more mechanistic and perhaps even causal accounts of epidemiological findings in IHS.

References

Cole, S. W. (2014). Human social genomics. *PLoS Genetics*, *10*, e1004601.
Google Scholar WorldCat

Domingue, B. W., Liu, H., Okbay, A., & Belsky, D. W. (2017). Genetic heterogeneity in depressive symptoms following the death of a spouse: Polygenic score analysis of the US Health and Retirement Study. *American Journal of Psychiatry*, *174*, 963–970. E Pub ahead of print.
Google Scholar WorldCat

Hamdi, N., South, S., & Krueger, R. F. (2016). Does education lower allostatic load? A co-twin control study. *Brain, Behavior, and Immunity*, *56*, 221–229.
Google Scholar WorldCat

Kaprio, J., & Koskenvuo, M. (1990). Cigarette smoking as a cause of lung cancer and coronary heart disease. A study of smoking-discordant twin pairs. *Acta Geneticae Medicae et Gemellologiae: Twin Research*, *39*, 25–34.
Google Scholar WorldCat

Kong, S. W., Lee, I.-H., Leschiner, I., Krier, J., Kraft, P., Rehm, H. L., . . . the MedSeq Project. (2015). Summarizing polygenic risks for complex diseases in a clinical whole genome report. *Genetics in Medicine*, *17*, 536–544.
Google Scholar WorldCat

Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking "Big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin*, *136*, 768–821.
Google Scholar WorldCat

Marioni, R. E., Ritchie, S. J., Joshi, P. K., Hagenaars, S. P., Okbay, A., Fischer, K., . . . Deary, I. J. (2016). Genetic variants linked to
p. 45 education predict longevity. *Proceedings of the National Academy of Sciences of the United States of* ↳ *America*, *113*(47), 13366–13371. http://doi.org/10.1073/pnas.1605334113
WorldCat

McEwen, B. S. (2000). Allostasis and allostatic load: Implications for neuropsychopharmacology. *Neuropsychopharmacology*, *22*, 108–124.
Google Scholar WorldCat

McEwen, B. S. (2017). Allostasis and the epigenetics of brain and body health over the life course: The brain on stress. *JAMA Psychiatry*, *74*(6), 551–552.
Google Scholar WorldCat

Ott, J., Kamatani, Y., & Lathrop, M. (2011). Family-based designs for genome-wide association studies. *Nature Reviews Genetics*, *12*, 465–474.
Google Scholar WorldCat

Pasaniuc, B., & Price, A. L. (2017). Dissecting the genetics of complex traits using summary association statistics. *Nature Reviews Genetics*, *18*, 117–127.
Google Scholar WorldCat

Polderman, T. J., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*, 702–709.
Google Scholar WorldCat

South, S. C., Hamdi, N. R., & Krueger, R. F. (2017). Biometric modeling of gene-environment interplay: The intersection of theory and method and applications for social inequality. *Journal of Personality*, *85*, 22–37.
Google Scholar WorldCat

South, S. C., & Krueger, R. F. (2013). Marital satisfaction and physical health: Evidence for an orchid effect. *Psychological Science*, *24*, 373–378.
Google Scholar WorldCat

South, S. C., & Krueger, R. F. (2014). Genetic strategies for probing conscientiousness and its relationship to aging. *Developmental Psychology*, *50*, 1362–1376.
Google Scholar WorldCat

South, S. C., Schafer, M. H., & Ferraro, K. F. (2015). Genetic and environmental overlap between childhood maltreatment and
p. 46 adult physical health. *Twin Research and Human Genetics*, *18*, 533–544. ↳
Google Scholar WorldCat