Decomposing the causes of the SES-health gradient with biometrical modeling

by

S. Mason Garrison

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

December, 2016

Nashville, Tennessee

Approved:

Joseph Lee Rodgers, Ph.D.

Kristopher Preacher, Ph.D.

Carolyn J. Heinrich, Ph.D.

ACKNOWLEDGMENTS

I would first like to thank my advisor Joe Rodgers. His patience, enthusiasm, insight, and unwavering support were instrumental for the writing of this thesis. I could not imagine having a better mentor and advisor.

I would like to thank the rest of my thesis committee (Kris Preacher and Carolyn J. Heinrich) for their time, thoughts, and support.

I would also like to thank in alphabetical order the following individuals for their thought-provoking insights on this project.

- Ali Hadd,
- Mike Hunter,
- Patrick O'Keefe, and
- the Quantitative Methods Faculty.

This thesis is based on longitudinal data collected as part of the National Longitudinal Survey of Youth 1979, (NLSY79). These data are publicly available; a bibliography of articles using these surveys can be found at `https://nlsinfo.org/bibliography-start` This material is based upon work that has been supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. (DGE-1445197), and various means of institutional support from the following universities: Vanderbilt University.

Last compile was 11/08/2016 at 15:52:53

ABSTRACT

The consistent relationship between socioeconomic status (SES) and health has been widely covered in both the popular media and in scientific journals. These articles, both popular and scientific, draw strongly upon conventional explanations, that physical health inequalities are caused by material disadvantage directly (*e.g.*, access to medical care; Hummer, Rogers, & Eberstein, 1998) or indirectly (*e.g.*, chronic environmental stress; Baum, Garofalo, & Yali, 1999; McEwen & Stellar, 1993). Such explanations account for differences between those who have resources and those who do not, but they do not account for the finely stratified health differences that exist across the entire range of SES (see Gottfredson, 2004, for discussion of additional limitations and contradictions).

Recent theories to explain the SES-health gradient have grappled with limitations and contradictions of early models, and have implicated very different pathways to explain the gradient. For example, articles in differential epidemiology have argued that individual differences, such as personality and cognitive ability, are the 'fundamental cause' of the gradient, acting through genetic sources (Arden et al., 2016; Marioni et al., 2014). Alternatively, variants of the allostatic load theory (McEwen & Stellar, 1993), like the Risky Family model (Repetti, Taylor, & Seeman, 2002) implicate the early home-environment. Surprisingly, very little research has applied behavior genetic modeling to understanding the sources of the SES-health gradient.

The purpose of this paper is to narrow the scope of fundamental causes, by explicitly untangling the sources of variance associated with the gradient. Specifically, we decompose the gradient into genetic (a^2), shared-environmental (c^2), and non-shared environmental (e^2) pathways, using data from the National Longitudinal Survey of Youth 1979. Monozygotic-twin, dizygotic-twin, full-sibling, half-sibling, and cousin pairs from recently validated kinship links are used in the current study (n= 4018 pairs; J. L. Rodgers et al., 2016).

Our findings have decomposed the relationship between SES and health into its biometrical components. Our results differed by health construct. Mental health's relationship with SES is primarily explained by the shared-environment, which is consistent with the Risky Family model of the gradient. The relation of physical health relationship with SES is primarily explained by genetic effects.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The relationship between socioeconomic status (SES) and health has been widely covered in the popular media; within the last year,[1] the Atlantic, Time Magazine, and the New York Times have all featured the SES-health gradient at length. Both popular and scientific articles draw strongly upon conventional explanations that material disadvantage directly (*e.g.*, access to medical care; Hummer et al., 1998) or indirectly (*e.g.*, chronic environmental stress; Baum et al., 1999; McEwen & Stellar, 1993) causes physical health inequalities. Such explanations account for differences between those who have resources and who have not.

Those situational explanations, however, do not account for the finely stratified health differences that exist across the entire range of SES (see Gottfredson, 2004, for an extensive discussion). They do not distinguish between mental and physical health, and even assume that the pathways are the same. In some cases, mental health is treated as a mediator, rather than as an outcome (Adler et al., 1994). Further, these explanations do not explain why greater access to health care (Siddiqi & Hertzman, 2007; Steenland, Henley, & Thun, 2002), accounting for rates of morbidity and mortality (Steenland et al., 2002), and improved education (Conti, Heckman, & Urzua, 2010) steepen the gradient instead of flattening it.

Recent theories have grappled with limitations and contradictions of early models and have implicated very different pathways to explain the gradient. For example, articles in differential epidemiology have argued that individual differences, such as personality and cognitive ability, are the 'fundamental cause' of the gradient, acting through genetic sources (Arden et al., 2016; Marioni et al., 2014). Alternatively, variants of the allostatic load theory (McEwen & Stellar, 1993), like the Risky Family model (Repetti et al., 2002)

[1] as of the time of this writing, 2015-2016

implicate the early home-environment. These theory-driven pathways align with behavior genetic models, which decompose relationships into genetic and environmental sources. Surprisingly, very little research has applied behavior genetic modeling to understanding the sources of the SES-health gradient. To the best of the author's knowledge, only one paper exists that explicitly attempts to do so (Lichtenstein, Harris, Pedersen, & McClearn, 1993, discussed in a later section).

The purpose of this paper is not to advocate for a specific fundamental cause of the gradient. Instead, the purpose of this paper is to narrow the scope of fundamental causal explanations, by explicitly analyzing and untangling the sources of the gradient using biometrical modeling. Specifically, I will decompose the gradient into genetic (a^2) and shared-environmental (c^2), and non-shared environmental (e^2) sources, using data from the National Longitudinal Survey of Youth 1979.

1.1 The SES-health gradient and its causes

The positive relationship between socioeconomic status and health (Adler et al., 1994; Antonovsky, 1967; Collins, 1926) is remarkably consistent across place, time, and method of assessment (Oakes & Rossi, 2003; Singh-Manoux, Marmot, & Adler, 2005). Conventional explanations of the gradient fall into three broad domains: (1) social causation, (2) social selection, and (3) social confounds (Adler et al., 1994; Adler & Ostrove, 1999; Adler & Stewart, 2010).

1. Social Causation theories argue that an individual's social standing affects their health (*e.g.*, through allostatic load, Baum et al., 1999; McEwen & Seeman, 1999; McEwen & Stellar, 1993, however see Matthews, Gallo, and Taylor, 2010).

2. Social Selection theories contend that health affects people's abilities to climb the social ladder (*e.g.*, social drift, J. W. Fox, 1990, see Rodgers, B. & Mann 1993 for methodology critiques).

3. Social Confound theories assert that a third variable causes health and wealth to covary, instead of either causally influencing the other (*e.g.*, cognitive ability, Gottfredson, 2004; Gottfredson & Deary, 2004).

Many theories, such as intergenerational transmission (Haas, 2006) and reciprocal influences (Smith, 1999) involve multiple domains. Accordingly, multiple pathways could be in play – either simultaneously or sequentially. Or, multiple theories, and thus different pathways could simultaneously explain specific health-wealth relationships. I note that many theorists (notably, Adler et al., 1994; Adler & Ostrove, 1999; Adler & Stewart, 2010) assert that Social Causation and Selection theories act exclusively through the environment, whereas Social Confound theories act exclusively through genes. However, these conceptualizations were developed outside of a behavior genetic framework, and thus lack some of the nuance and insight from current behavior genetics research. I discuss mapping current theory onto behavior genetic modeling in a later section.

1.2 Social Causation

Social Causation theories vary on two dimensions: baseline (abundance vs deprivation),[2] and pathway (direct vs indirect) (Wagstaff & Doorslaer, 2000). On the baseline dimension, the baseline outcome (*i.e.*, health) is improved by abundance (*e.g.*, more wealth) or it is harmed by deprivation (*e.g.*, poverty). Does being poor lead to worst health or does being rich lead to better health? The pathway dimension illustrates how far apart the cause and effect can be, and how many steps it takes to get from cause to effect (*i.e.*, few versus multiple. Direct pathways connect cause and effect in a few steps (typically, one or two; *e.g.*, improved wealth allows greater access to medical care, which improves health; Hummer et al., 1998), whereas indirect pathways connect cause to effect in multiple steps (*e.g.*, poverty leads to living in areas with more stressors, whereby the increasing stress levels

[2]Wagstaff and Doorslaer (2000) note that the theoretical distinctions between these dimensions have not be fully articulated in the literature

reduce health).

1.2.1 Direct Models

The Absolute Income Hypothesis (W. Evans, Wolfe, & Adler, 2012; Keynes, 1936) is a direct abundance model, where increased wealth allows individuals to buy better health (through, *e.g.*, improved access to healthcare). The deprivation equivalent model is the Absolute Deprivation Hypothesis or Poverty Hypothesis, where decreased wealth costs individuals health because they are unable to buy health goods (*e.g.*, health insurance). The distinction between these models is subtle and has not been clearly articulated in the literature; the abundance models assume that wealth positively influences baseline health, whereas deprivation models assume that lack of wealth negatively influences baseline health.

Variants on these direct models include the Relative Income Hypothesis (Duesenberry, 1967), which suggests that relative increases in wealth allow individuals to buy better health. Similarly, the Relative Status Hypothesis suggests that *relative* increases in status (*e.g.*, wealth, prestige) allow individuals to purchase health. The Relative Status Hypothesis is the broadest of the direct abundance models as it allows SES gains in rank as well as wealth. I direct readers to Lynch et al. (2004) for discussion on the mixed empirical support for direct models, and note that US samples show the most support. These direct models have distinct implications for interventions (see, W. Evans et al., 2012, for discussion).

1.2.2 Indirect Models

Indirect models of Social Causation argue that rather than SES directly influencing health, SES places individuals in environments that in turn influence health. In the majority of models, this indirect influence of health is linked to stress (Adler, 2013), conceptualized both as life events requiring adaptation and as a state where perceived demands exceed their coping abilities. Under allostatic load theory (Baum et al., 1999; McEwen & Stellar, 1993), these environmental stresses of poverty cause "wear and tear on the body" that accumulate

over an individual's lifetime. Although longitudinal evidence for this effect in adulthood is mixed (Matthews et al., 2010), evidence from childhood is not. Specifically, the Risky Families model identifies early childhood environments created by violence, overt aggression, or neglect as sources of poor health across the lifespan (Repetti et al., 2002). Children of poverty are at greater risk for those household stressors (*e.g.*, Dodge, Pettit, & Bates, 1994; Yoshikawa, Aber, & Beardslee, 2012). Moreover, there is evidence that the effects of poverty on later health are mediated by childhood exposure to early environmental stresses (G. W. Evans & English, 2002).

1.3 Social Selection

Social Selection theories reverse the causal arrow – instead of SES causing health outcomes, health causes changes in SES. The theories in this area can be partitioned into two mechanisms of action: drift and stunting (Haas, 2006). In drift-based theories, individuals with poorer health slowly drift into lower and lower socioeconomic classes. This downward shift can be the result of discrimination (Ameri et al., 2015), reduced employment (Marwaha & Johnson, 2004), and decreased personal wealth (Chirikos & Nestel, 1985). For example, early longitudinal research on schizophrenia observed between-generation drift and within-generation drift (Goldberg & Morrison, 1963). At their first hospital admission, patients with schizophrenia held similar status jobs to patients with diagnoses of other than schizophrenia (predominantly "anxiety states"). At follow-up (between one and four years later), patients with schizophrenia had a noticeable within-person decline in job status, whereas patients with other diagnoses had no such decline. In drift models, individuals move down the social ladder.

The other mechanism is stunting, where poor health during critical periods of development negatively impact an individual's ability to meet key economic/educational milestones, such as learning to read or graduating high school. For example, findings from the Bucharest Early Intervention Project suggested that deprivation during the first two years of

life permanently impaired cognitive development (Nelson et al., 2007). In the longer term, stunting prevents individuals from accumulating status, and thereby inhibits their social mobility.

Lately, these theories have become less popular in the psychology and other behavioral science literature. For example, recent reviews by Adler and Stewart (2010) have explicitly dismissed social selection, whereas other reviews have implicitly dismissed them through omission (*e.g.*, Adler, 2013; Braveman & Gottlieb, 2014) or brief coverage (*e.g.*, W. Evans et al., 2012). Although social selection is commonly dismissed as the cause of the gradient, clearly there are subsets of individuals for whom these theories apply. Moreover, the selection effects might be of more importance depending on context – perhaps within the contemporary developing world and in the yesteryears of the developed world.

1.4 Social Confound

Unlike Social Causation and Social Selection theories, Social Confound theories argue that a third variable causes health and wealth to covary, instead of either causally influencing the other. Predominantly, the third variables identified are individual differences in personality and cognitive ability (Deary, 2010; Deary, Weiss, & Batty, 2010). Although many individual differences are linked with SES and health (Chapman, Roberts, & Duberstein, 2011; Deary et al., 2010), I will focus the discussion on cognitive ability[3] and conscientiousness[4] because their relationships are the most investigated and best understood for both SES and health.

[3] "[Cognitive ability] is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience. It is not merely book-learning, a narrow academic skill, or test-taking [ability]. Rather, it reflects a broader and deeper capability for comprehending our surroundings 'catching on', 'making sense' of things, or 'figuring out' what to do" (Gottfredson, 1997).

[4] "Conscientiousness is defined as the propensity to follow socially prescribed norms for impulse control, to be goal directed, to plan, and to be able to delay gratification" (Roberts, Jackson, Fayard, Edmonds, & Meints, 2009).

1.4.1 Individual Differences and SES

Conscientiousness and cognitive ability are consistently associated with composite measures of SES (Chapman, Fiscella, Kawachi, & Duberstein, 2010; Hernstein & Murray, 1994) and its components (education, occupation, income; Ng, Eby, Sorensen, & Feldman, 2005; Strenze, 2007). Given that I am interested in these associations as confounds rather than mediators in causation and selection models, I focus exclusively on how cognitive ability and conscientiousness influence SES.[5] As is the case with many outcomes, cognitive ability and conscientiousness indirectly influence the components of SES through decision-making and various behaviors. These behaviors and decisions are numerous, so I briefly review them here.

Education and academic performance. Individuals with higher cognitive ability perform better in classes and attain higher levels of education (Deary, Strand, Smith, & Fernandes, 2007; Strenze, 2007), even within the highest percentiles of ability (Kell, Lubinski, & Benbow, 2013). Similarly, self-ratings of conscientiousness predict performance in medical school (Lievens, Ones, & Dilchert, 2009) and teacher-ratings of conscientiousness predicted higher levels of education attainment (Lleras, 2008). I direct readers to Deary and Johnson (2010) and Kuncel, Ones, and Sackett (2010) for treatment of the relationship between individual differences and education.

Occupation and specialization. Cognitive ability predicts later occupational attainment, where higher ability individuals tend to obtain more prestigious occupations (Strenze, 2007). Performance on the GRE varies by intended graduate major (Education Testing Services, 2016); Physical Science graduate majors had the highest overall scores (309), although Science Engineering graduate majors scored highest in Quantitative Reasoning (159), and Arts and Humanities majors scored highest in Verbal Reasoning (157). More-

[5] Discussions and research concerning the causes of these associations, especially for cognitive ability's association with SES are controversial. In many instances (*e.g.*, Jensen, 1969b), they devolve into heated public debates (Alfert, 1969a, 1969b; Davies, 1969; Jensen, 1969a, 1969c) Even the direction of the causal arrow is controversial. I note that these controversies are predominantly social/political in nature, rather than scientific.

over, the differences in domains of verbal and quantitative ability at age 13 are predictive of specializing in the humanities or the sciences 30 years later (Kell et al., 2013). Less work has been done on the relationship between conscientiousness and later occupation. Conscientious individuals tend to seek out investigative occupations (*e.g.*, scientists, statisticians, doctors; Judge, Higgins, Thoresen, & Barrick, 1999), and this tendency appears to indirectly influence later occupation through vocational interests (Woods & Hampson, 2010). I direct readers to Schmidt and Hunter (2004) and Major, Johnson, and Deary (2014) for treatment on the relationship between individual differences and occupation.

Income and employment. Cognitive ability and conscientiousness are consistently identified as the most effective measures in personnel selection (*e.g.*, Behling, 1998). Higher levels of conscientiousness are positively associated with job attendance (Judge, Martocchio, & Thoresen, 1997), performance (Dudley, Orvis, Lebiecki, & Cortina, 2006), retention (Eskreis-Winkler, Shulman, Beal, & Duckworth, 2014), promotion, and salary (Ng et al., 2005). Similarly, cognitive ability predicts job performance (Ree & Earles, 1992; Schmidt & Hunter, 2004), salary (Ng et al., 2005; Strenze, 2007), and training performance (Ree & Earles, 1990), among others.

1.4.2 Differential Epidemiology

Conscientiousness and cognitive ability consistently predict health (Gottfredson & Deary, 2004; Hampson, Goldberg, Vogt, & Dubanoski, 2007), and do so across the lifecourse, including the ultimate measure of health – longevity (Batty, Deary, & Gottfredson, 2007; Jackson, Connolly, Garrison, Leveille, & Connolly, 2015; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). Cognitive ability's effect on longevity does not vary in strength across most causes of death (Christensen, Mortensen, Christensen, & Osler, 2016). Although health can impact cognitive ability and personality (*e.g.* traumatic brain injury Catroppa, Anderson, Morse, Haritou, & Rosenfeld, 2008; Mandleberg & Brooks, 1975; Rochat et al., 2010), I focus exclusively on how these individual differences influence health. Primarily,

they do so in two ways. First, cognitive ability and conscientiousness indirectly influence health through decision-making and various behaviors. Conscientiousness is consistently associated[6] with health-related behaviors (such as risky driving, physical activity, violence, tobacco use, *etc*., Bogg & Roberts, 2004), which partially mediates its relationship with health (Lodi-Smith et al., 2010). Cognitive ability is also associated with health-related behaviors (Gottfredson, 2004, but see Garrison and Rodgers, 2016), and is linked to health literacy (Beier & Ackerman, 2003). Indeed, cognitive ability accounts for much of the relationship between education and health behaviors (Cutler & Lleras-Muney, 2010). Second, cognitive ability is an indicator of overall "system integrity" (Deary, 2012; Lubinski, 2009). System integrity is a "general latent trait of a well-functioning body" (Gale, Batty, Cooper, & Deary, 2009), and reflects how efficiently the human body handles complex systems. "Well-wired" individuals would be more intelligent and healthier on average than "poorly-wired" individuals. This construct may be related to allostatic load – perhaps "well-wired" individuals are more able to endure environmental stresses.

1.5 Linking Gradient Theories to Environmentality and Heritability

In the previous sections, I have reviewed the three broad theories on causes of the SES-Health gradient (*i.e*., Social Causation, Social Selection, and Social Confounds). Thus far, linkages of theory to biometrical modeling have been sparse by design. I have noted that many theorists (notably, Adler et al., 1994; Adler & Ostrove, 1999; Adler & Stewart, 2010) assert that Social Causation and Selection theories act exclusively through the environment, whereas Social Confound theories act exclusively through genes.

However, these conceptualizations were developed outside of a behavior genetic framework, and thus lack some of the nuance and insight from current behavior genetics research. In reality, some of the specific theories map nicely onto biometrically-informed conceptu-

[6]This association is positive for health protective behaviors like physical activity, and negative for risky behaviors like violence

alizations of gene and environmental influences (*e.g.*, the Risky Families Model onto the shared-environment, Repetti et al., 2002), moreso than other theories. Thus, this section will review the biometrically-informed conceptualizations of gene and environmental influences. I will not link every theory to its biometrical source, but I will illustrate a select few.

1.5.1 Environmentality

In behavior genetics, the proportion of variance in a trait that is not due to genes is environmentality (the analog for heritability). These environmental influences can be divided into shared- and non-shared environmental experiences. The shared-environment, in a biometrical sense, consists of experiences shared by siblings, whereas the non-shared-environment consists of experiences not shared by siblings. These experiences can also be conceptualized as between- and within-family experiences. The shared-environment is exclusively made from between-family experiences. These between-family experiences result in family members becoming more similar to one another than genetic similarity alone would predict. Classic examples of shared-environmental experiences include parental discord, divorce, and socioeconomic status. Shared experiences do not exclusively arrive by means of parents. They can also include sibling shared-experiences, or even twin-specific (*e.g.*, shared placenta) experiences (Neale & Maes, 2004).

The non-shared environment has two sources of non-shared experiences: objective and effective (Goldsmith, 1993; Turkheimer & Waldron, 2000). Objective sources are within-family influences, where explicit differences within a family result in non-shared experiences. These experiences include:

- activities (such as trips to the museum, J. L. Rodgers, Rowe, & May, 1994),
- education environments (like the classroom, Vernon, Jang, Harris, & McCarthy, 1997),

- parenting behaviors (including discipline, McGuire, Dunn, & Plomin, 1995),
- peer groups,
- sibling interactions, or family structure (*e.g.*, birth order, Rowe, Rodgers, & Meseck-Bushey, 1992).

In contrast, "effective" experiences are between-family experiences, where a shared-event leads to unique experiences. These common experiences include parental discord, divorce, and socioeconomic status.[7] Early behavior genetic work misattributed these between-family sources of non-shared experiences as exclusively shared-environmental. Superficially, two children within a household will both experience poverty.[8] However, those two hypothetical children experience that same moment of poverty at different stages of development, which in turn has the potential to differentially impact them. The majority of non-shared environmental variance comes from these between-family experiences (Turkheimer & Waldron, 2000).

1.5.2 Heritability

Heritability reflects the extent to which genetic differences contribute to observed individual differences (see Visscher, Hill, & Wray, 2008, for a broad review). Environmentality – its analog – has been discussed above and reflects the extent that environmental differences contribute to observed individual differences. Heritability estimates can vary by age (Bergen, Gardner, & Kendler, 2007), vary by cohort (Silventoinen, Kaprio, Lahelma, & Koskenvuo, 2000), be moderated by the environment (*e.g.*, Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003), and do not reflect extreme environments (only those typical within a population). Most, if not all, human behaviors are heritable, but those same be-

[7] Note that these are the same experiences as listed as examples of shared-environmental experiences.

[8] When subjects are of the same developmental age as is the case with twins, between-family sources of the environment cannot be distinguished from shared-environmental sources. This can hold in cases where there are no age effects.

haviors can also have large environmental influences (most often non-shared, Turkheimer, 2000).

There are two types of heritability:

- Broad heritability, which encompasses all genetic variance, including non-additive and nonlinear processes such as dominance and epistatic effects; and
- Narrow heritability, which focuses on additive genetic variance.

In non-human animal experiments, the majority of genetic variance is additive (Neale & Maes, 2004). Additive genetic effects reflect the "sum of the average effects of individual alleles". Accordingly, early models of heritability in humans only incorporated the additive components. Methodological advancements have incorporated other genetic variance components, including dominance and epistatic interactions (Lykken, McGue, Tellegen, & Bouchard, 1992). Regardless, classic genetic models focus on narrow heritability, as does this paper. For more on modeling heritability, see Neale and Maes (2004).

(Mis)Interpreting Heritability. Unlike environmental influences, heritability is prone to be misinterpreted. Such misconceptions can be construct specific, such as SES (Rowe & Rodgers, 1997), or they can be universal to all constructs. Visscher et al. (2008) reviewed five common misconceptions of heritability.

1. "Heritability is the proportion of a phenotype that is passed on to the next generation"
2. "High heritability implies genetic determination"
3. "Low heritability implies no additive genetic variance"
4. "Heritability is informative about the nature of between-group differences"
5. "A large heritability implies genes of large effect"

I will discuss misconception #2 because of its relevance to Social Causation, and I direct the reader to Visscher et al. (2008) for more thorough explanations of all the misconceptions.

Genetic determination is not implied by high heritability. (Nor does low heritability imply the absence of genetic causation.) Rather, high heritability suggests that parents with a specific trait are likely to pass that trait onto their offspring. Misconceptions about genetic determination lead one to assume a strong link between genes and outcome, either directly (*i.e.*, genetic fixity; Dennett, 2015) or indirectly (*i.e.*, innate capacity; Lewontin, 1991). Such attitudes can do more harm than good when it comes to human behavior, as they imply that heritable outcomes cannot be influenced by social policy (Goldberger, 1979).[9]

Yet, contrary to these attitudes, medical/policy interventions can affect fully-heritable disorders. For example, phenylketonuria (PKU) is a genetic metabolic disorder, where phenylalanine cannot be metabolized by the liver (Fölling, 1934; National Institutes of Health, 2000). Without treatment, phenylalanine accumulates in the tissue and blood; infants with the disorder fail to reach developmental milestones, suffer from brain abnormalities, and tend to have severe intellectual disabilities. The symptoms of PKU can be drastically reduced by monitoring phenylalanine levels and restricting the amount of phenylalanine consumed (*e.g.*, soy beans, whale, nuts). Prior to widespread screening in the 1960s, the symptoms of PKU were fully heritable; now, the expression of PKU is almost fully environmental (Plomin, DeFries, & McClearn, 1990) because it can be controlled by dietary behaviors.

1.6 Prior Work

Some studies have decomposed SES-Health differences into between- and within-family variance (*e.g.*, Monden, 2010; Søndergaard et al., 2012, 2013). They found sizable between-family variance. For example, 20% of education and self-assessed Health was explained by between-family variance (Monden, 2010), and education and all-cause mortality (Søndergaard et al., 2012) hazard ratios were reduced by 10-40% when controlling for between-family

[9] Goldberger (1979) went so far as to advocate that no one estimate heritability as it does more harm than good.

variance. These studies use components of SES, but not the overall construct; nevertheless, these findings are suggestive of shared-environmental effects and/or genetic effects. Because this method aggregates shared-environmental and genetic effects, I cannot eliminate theories that suggest genetic pathways (Social Confound theories) or shared-environmental pathways (Social Causation theories).

To date, Lichtenstein et al. (1993) is the only identified study that decomposes the SES-Health gradient into its biometrical sources of variance. They used the Swedish Adoption/Twin Study of Aging (SATSA; n = 785 pairs), and identified 398 twin pairs that were raised apart, as well as 387 twin pairs raised together, ranging in age from 26 to 86 in 1984. The raised-together twins were matched on age, gender, and county of birth. 82% of raised-apart twins were separated before their 5^{th} birthday. The primary reasons for separation were parental health (*e.g.*, parental illness, parental death) or "economic problems." This cross-sectional analysis used multiple components of SES:

- Material resources (Index of modern conveniences, homeownership, savings *etc.*),
- Perceived standard of living,
- Education (4 levels, with a maximum of university attendance), and
- Occupational status;

and two measures of health:

- Number of organ systems affected by a chronic health problem, (SUMILL; Harris, Pedersen, McClearn, Nesselroade, & Plomin, 1992), and
- Self-rated health (4 items, including retrospective reports).

They identified a moderate genetic covariance (mean $r_a = .5$) across all components of SES and health. The environmental effects were inconsistent. Lichtenstein et al. (1993) found a shared-environmental effect in self-rated health, but not in the SUMILL, whereas

non-shared environmental effects were small (mean $|r_e| = .07$), but the direction of these effects was inconsistent. The inconsistency of their results might be due to limitations of their study. They did not create a composite measure of SES, instead relying on the components. This method has its merits, but it does not allow researchers to test broad theories about causes of the gradient. The authors did not address potential selection effects caused by health- or wealth-related twin separations, or survivor effects given the age of the subjects ($\text{age}_{mean} = 58.6$, according to Harris et al., 1992, who also noted that these measures were influenced by age). Regardless, their results were a promising first step into identifying the sources of the gradient, and helped motivate the current study.

Chapter 2

Current Study

To summarize, the current study examines the relationship between socioeconomic status and health, using biometrical modeling and data from a nationally representative sample, the National Longitudinal Survey of Youth 1979 (NLSY79; described later). This examination extends the SES-health gradient literature in several key ways. First, I decompose the relationship between SES and health into genetic and environmental components, using a national probability sample, and thus have findings that are representative of all levels of socioeconomic status. Second, I employ a dataset with kinship pairs, a more general representative of those in U.S. families, which naturally accounts for more levels of relatedness than traditional twin or adoption designs. Third, I distinguish between mental and physical health in these analyses, and provide analytic results from each domain separately and also a combined analysis of both domains together. Finally, I evaluate various theories within the context of these result, identifying ones that are inconsistent with my empirical results.

The results from Lichtenstein et al. (1993) are my only basis for making predictions. However, Lichtenstein et al. (1993) is a fairly old and data-limited sample. In general, they support a genetic component and possibly a shared-environmental component influencing the bivariate relationship between SES and physical health. They provide no guidance for mental health. Therefore, rather than making specific predictions, I will discuss the broad analytic plan employed in this paper.

First, I will conduct univariate analyses on SES, physical, and mental health. The best fitting models will be determined by series of nested model comparisons. The univariate results from the best fitting models will guide model fitting for later stages. Second, I will conduct bivariate analyses on SES with physical health, and SES with mental health.

These models will be guided by the univariate results. Specifically, all tests of whether a correlation has a heritable component are contingent on both variables having heritable components at the univariate level. For example, if SES does not have a heritable component, then there can be no common heritable component between SES and any measure of health. This same logic applies to shared-environmental components. Again, the best fitting models will be determined by series of nested model comparisons. Finally, a trivariate model will be run, linking SES, physical health, and mental health. The best fitting bivariate models will be merged to create this trivariate model. Again, the best fitting model will be determined by a series of nested model comparisons.

Chapter 3

Method

3.1 Model

3.1.1 Assumptions

There are five major modeling assumptions that behavior genetic models employ in path diagrams. Neale and Maes (2004, pg. 114) summarized them succinctly:

1. "No genotype-environment correlation, *i.e.*, latent genetic variables [(a)] are uncorrelated with latent environmental variables [(c)] and [(e)];"

2. "No genotype environment interaction, so that the observed phenotypes are a linear function of the underlying genetic and environmental variables;"

3. "Random mating, *i.e.*, no tendency for like to marry like, an assumption which is implied by fixing the covariance of the additive genetic deviations of [...] full sib[ling]s to 0.5$[\text{Cov}]_a$," half siblings to .25Cov_a, and cousins to .125Cov_a;

4. "Random placement of adoptees, so that the rearing environments of separated twin pairs are uncorrelated;" and

5. Equal environments, siblings of different levels of relatedness (full siblings vs half siblings) are equally exposed to "environmental events of etiologic importance" (Kendler, Neale, Kessler, Heath, & Eaves, 1993).

All assumptions but random adoption placement were relevant to the behavior genetic models employed in this paper.

3.1.2 Specification, Estimation, and Missing Data

All models were specified using Mplus (version 7.4, Muthén & Muthén, 2014), and estimated using full information maximum likelihood (FIML) to account for missing data. The univariate ACE model specification is illustrated with the path diagram in Figure 3.1. Sample Mplus syntax, adapted from Prescott (2004), is provided in section C.1 of Appendix C. The bivariate correlated factors ACE model specification is illustrated with the path diagram in Figure 3.2. The diagram is simplified to only include one member of the kin pair. Sample Mplus syntax is provided in section C.2 of Appendix C.

3.1.3 Terminology

Throughout this paper, I refer to the covariance between genetic factors (cov_a), the genetic correlation (r_a), and the heritability of the (phenotypic) correlation. These terms are related, but not identical constructs. The covariance between genetic factors describes the degree of overlap between the genetic variance in PCS with the genetic variance SES. The genetic correlation reflects the correlation between PCS and SES that is attributable to the covariance between genetic factors. The heritability of the correlation is the proportion of correlation between PCS and SES that is attributable to the covariance between genetic factors.

Similarly, I refer to the covariance between shared-environmental factors (cov_c), the shared-environmental correlation(r_c), and the environmentality of the (phenotypic) correlation. Again, these terms are related, but not identical constructs. The covariance between shared-environmental factors describes the degree of overlap between the shared-environmental variance in MCS with the shared-environmental variance SES. The shared-environmental correlation reflects the correlation between MCS and SES that is attributable to the covariance between shared-environmental factors. The environmentality of the correlation (between SES and MCS) is the proportion of correlation between MCS and SES

that is attributable to the covariance between shared-environmental factors.

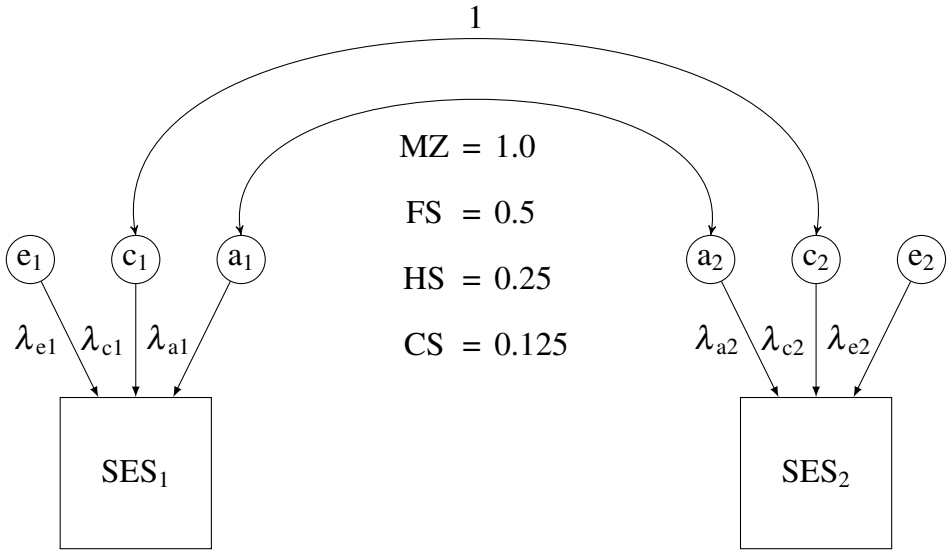

Figure 3.1: Path diagram illustrating model specification for a univariate ACE model. *a*, additive genetic factor; λ_{a}, additive genetic loading; *c*, shared-environmental factor; λ_{c}, shared-environmental loading; *e*, non-shared-environmental factor; λ_{e}, non-shared-environmental loading; $\mathrm{SES}_{1,2}$, socioeconomic measure for kin 1 and kin 2; MZ, monozygotic twins; FS, full siblings; HS, half siblings; CS, cousins.

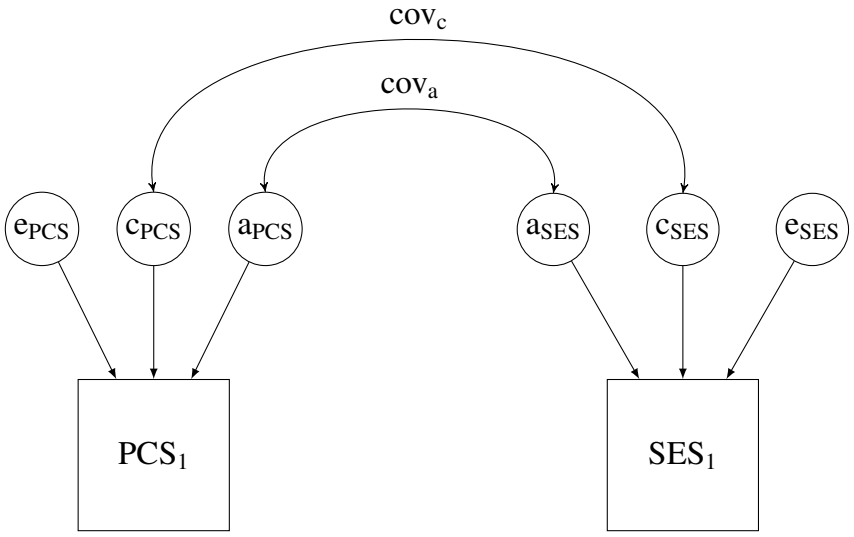

Figure 3.2: Path diagram illustrating model specification for a bivariate correlated factors ACE model. The diagram is simplified to only include one member of the kin pair. *a*, additive genetic factor; λ_a, additive genetic loading; cov_a covariance between additive genetic factors; *c*, shared-environmental factor; λ_c, shared-environmental loading; cov_c covariance between shared-environmental factors; *e*, non-shared-environmental factor; λ_e, non-shared-environmental loading.

3.2 Subject Characteristics

The National Longitudinal Survey of Youth 1979 dataset (NLSY79) (described elsewhere in Garrison & Rodgers, 2016), is based on a nationally representative household probability sample. The NLSY79 was jointly sponsored by the U.S. Bureau of Labor Statistics and the U.S. Department of Defense. On December 31, 1978, 12,686 adolescents were sampled within a household probability sample from 8,770 households. The initial sample consisted of three subsamples:

- a cross-sectional household probability sample of 6,111 non-institutionalized adolescents residing in the United States on December 31st of 1978;
- a separate over-sampled civilian subsample of 5,295 racial minorities and disadvantaged whites;

- a representative sample of 1,280 youth serving in the U.S. Military on September 30^{th}, 1978.

In the two civilian samples, subjects' birthdates ranged from January 1, 1957 to December 31, 1964, and were between the ages of 14 and 21 on December 31, 1978; military subject's birthdates ranged from January 1, 1957 to December 31, 1961, and were between 17 and 21 years old. Participants were surveyed annually until 1994, and then surveyed biennially to the present. Two waves of planned attrition occurred, due to budgeting restrictions. After the 1984 interview, all but 201 randomly selected members of the military sample were dropped. After the 1990 interview, all 1,643 disadvantaged whites from the oversample were dropped. Note that because there are no siblings within the military subsample, it is irrelevant for the current research, as all military respondents are screened out of the analyses by the requirement of having siblings within the sample. More information about the sampling process and the data can be found on the Bureau of Labor Statistics (BLS) website: `http://www.bls.gov/nls/nlsy79.htm`

To conduct this study using the requisite within-family information, I require kinship pairs. In the original NLSY79 survey, there was no explicit identification of level of sibling relatedness. NLSY79 twins, full siblings, half siblings, and adoptive siblings were distinguishable indirectly from respondent and maternal information about birthdates and the biological father(s). In 2006, the survey included explicit indicators of the level of sibling relatedness. Our research team has recently completed a multi-year project to reliably and validly identify the kinship pairs (J. L. Rodgers et al., 2016), using both indirect and direct ascertainment of kinship relatedness.

3.2.1 Sample Selection

Monozygotic-twin, dizygotic-twin, full-sibling, half-sibling, and cousin pairs (who lived together in the same household in 1979) are used in the current study. Table 3.1 provides sample sizes for each level of kinship-relatedness, which varied because they are approxi-

mately representative of the distribution of kinship pairs in the population.

Throughout the text, I refer to two samples in the NLSY79: Full Sample, and National Probability Subsample. The Full Sample consists of all three subsamples described in the previous section (*i.e.*, a cross-sectional household probability sample, a separate oversampled civilian subsample of racial minorities and disadvantaged whites, and a representative sample of youth serving in the U.S. Military). The National Probability Subsample consists only of the cross-sectional household probability subsample.

Table 3.1: Number of Kinship Pairs by Level of Relatedness in the Full Sample

Kinship	Number of Pairs
Cousin	96
Half Sibling	297
Full Sibling	4,006
MZ Twin	11

3.3 Health Measures

The NLSY respondents were given a health "module," a battery of health questions, at age 40. Administration began in 1998, when the oldest respondents (*i.e.*, those born in 1957 and 1958) reached age 40. The module continued to be administered biennially until the youngest respondents reached age 40 in 2006. Approximately half of subjects took the Health 40 module at age 41. The 40+ module included:

- a 7-item version of Center for Epidemiological Studies Depression Scale, (CES-D; see Levine, 2013, for psychometric properties)
- questions on family history and health care history,
- the 12-Item Short Form Health Survey (SF-12; Ware, Kosinski, & Keller, 1995), and
- a health ailment checklist.

More details about these health modules can be found here: `https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/health`

I selected two measures for health: the physical component summary score of the SF-12 (PCS), and the mental component summary score of the SF-12 (MCS). Higher scores correspond to greater health. I focused on these two measures because they are established scales with documented psychometric properties and are norm-referenced. The other measures were created/adapted for use in the NLSY79, and have no external validation.

3.3.1 Physical Component Summary

PCS scores varied by sex and race.[1] These differences were significant, but not large in magnitude. Univariate t-tests were conducted without assuming equal variances by adjusting degrees of freedom with the Welch-Satterthwaite formula. Average scores for women (51.4, SD = 8.7) were significantly lower than average scores for men (52.6, SD = 7.3); $t(8216.49) = -6.88$, $p < .001$. Black respondents (51.3, SD = 8.4) and Hispanic respondents (51.8, SD = 8.1) had significantly lower average scores ($t(8379.9) = -5.54$, $p < .001$) than non-Black, non-Hispanic respondents (52.5, SD = 7.8). Throughout this paper, race has been dictomized into a minority status variable. These two groups consist of (0) Non-Black, Non-Hispanic respondents; and (1) Black and/or Hispanic respondents. Whenever possible, I have reverted to the traditional three group variable denoted as race.

Figure 3.3 characterizes the distribution of PCS, grouped by sex and race with smoothed density plots. As noted, although the group differences are significant, they do not differ greatly in magnitude.

[1] Race was defined by the original investigators based on a combination of self-identification, interviewer-report, and inference from household reports (Ward & Burich, 1978). This method resulted in three groups: Hispanic, Black, and non-Black-non-Hispanic respondents.

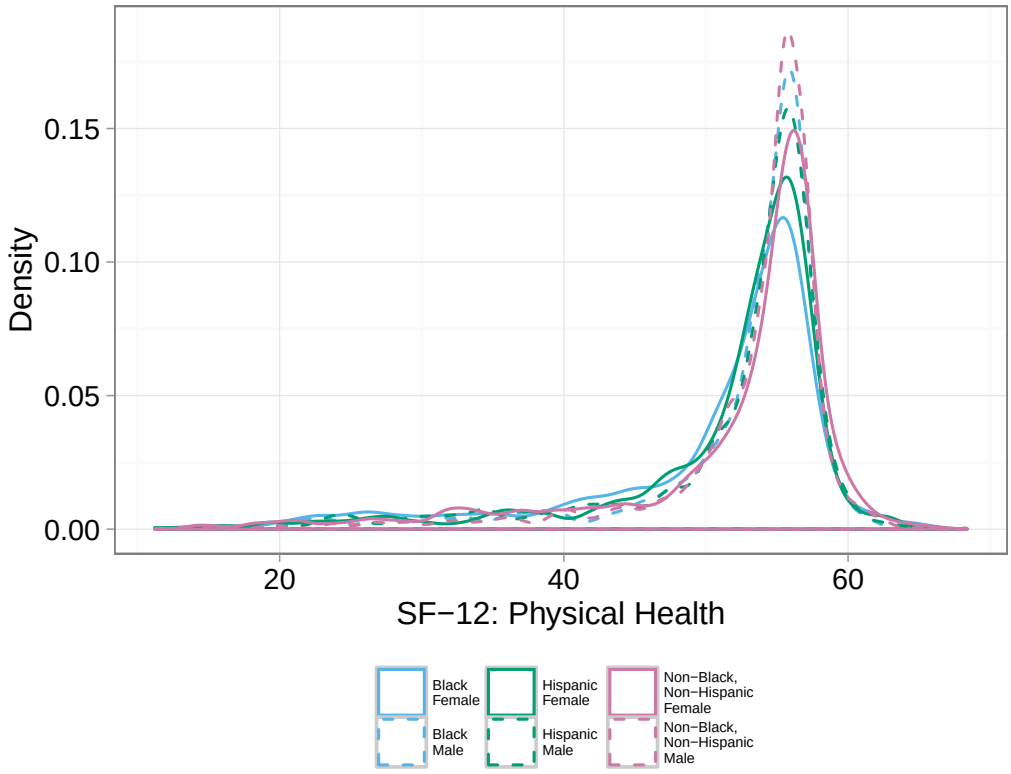

Figure 3.3: Density Plot of SF-12: Physical Health at Age 40 by Sex and Race

3.3.2 Mental Component Summary

MCS scores varied by sex, but not race. Figure 3.4 characterizes the distribution of MCS, grouped by sex and race with smoothed density plots. Average scores for women (51.9, SD = 8.9) were significantly lower than average scores for men (54.0, SD = 7.7); $t(8294.65) = -11.76$, $p < .001$. The Black respondent mean (53.0, SD = 8.7) and Hispanic respondent mean (52.9, SD = 8.7) did not differ significantly ($t(8351.06) = 0.18$, $p = 0.86$) from the non-Black-non-Hispanic respondent mean (52.9, SD = 8). Although sex differences were significant, they were not large in magnitude.

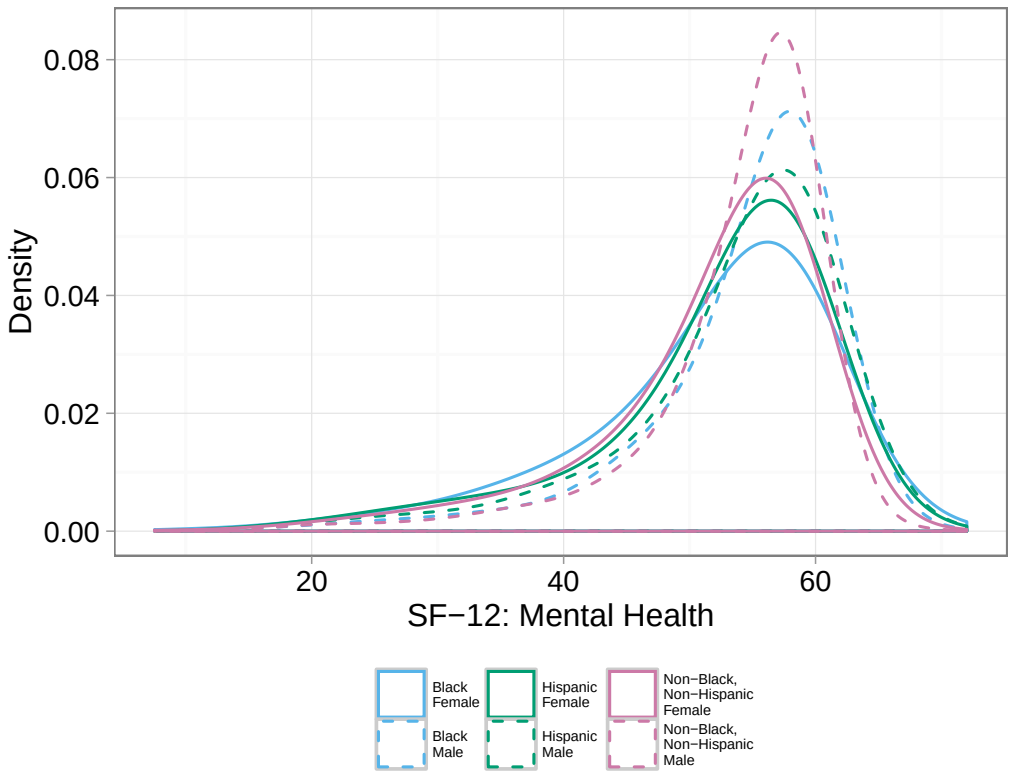

Figure 3.4: Density Plot of SF-12: Mental Health at Age 40 by Sex and Race

3.3.3 Normality

The distributions of both PCS and MCS were highly negatively skewed (see Figures 3.3 and 3.4), and were not normally distributed; see the QQplots in Figure 3.5. I transformed the SF-12 measures using Box-Cox transformations (Box & Cox, 1964).[2] The transformed variables' QQplots are displayed in Figure 3.6, and were much closer to normal. Analyses using the untransformed variables gave very similar heritability estimates, albeit with poorer model fits.

[2] I used the powerTransform function from the car package (J. Fox & Weisberg, 2010), which employs a maximum-likelihood adaptation of Box-Cox to estimate a transformation that normalizes a distribution.

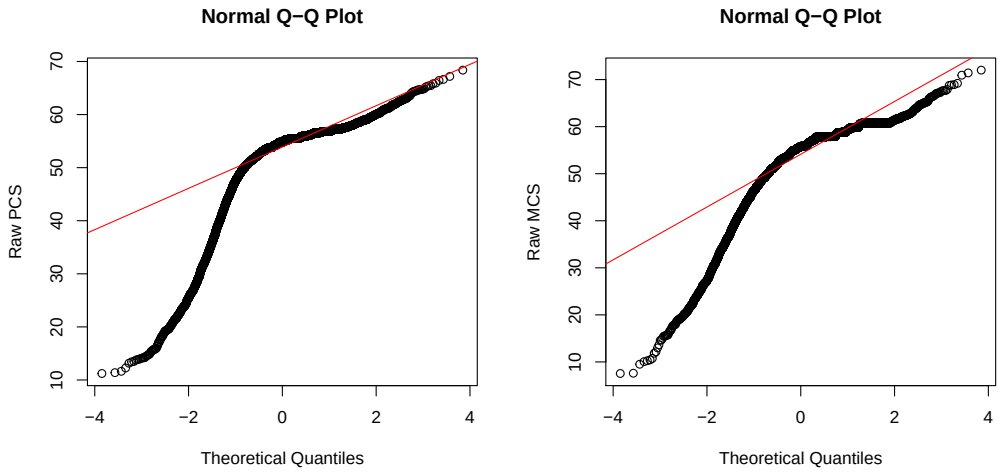

Figure 3.5: QQplots of raw score PCS and MCS

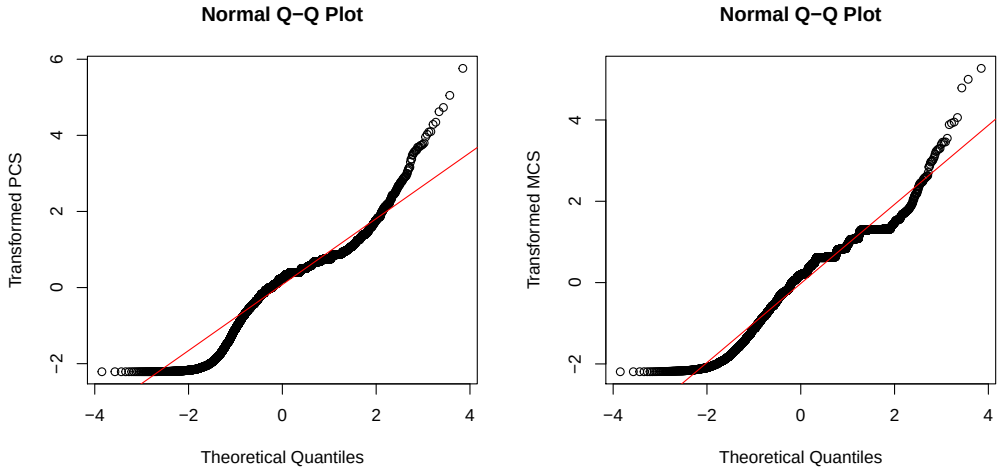

Figure 3.6: QQplots of transformed PCS and MCS

3.3.4 Heritability, Past Studies

There is limited information on the heritability of the SF-12. The Danish Twin Registry collected PCS and MCS from the SF-12 (Johnson et al., 2010; Steenstrup, Pedersen, Hjelmborg, Skytthe, & Kyvik, 2013) at age 45 (SD = 13.7). It is unclear whether the scale was translated into Danish, or the authors relied upon the fact that 86% of Danes speak English (European Commission, 2012). PCS and MCS distributions were skewed in the same manner as our sample (Johnson et al., 2010; Steenstrup et al., 2013). Median values

for PCS (55.9) and MCS (54.4) were comparable to non-minorities in the NLSY79 sample (PCS_{median} = 55.3, MCS_{median} = 55.6). Neither Johnson et al. (2010) nor Steenstrup et al. (2013) reported univariate ACE estimates. However, Johnson et al. (2010, Figure 3) illustrated the variance components for PCS moderated by educational attainment. At a 7^{th} grade education, a^2 was 0.63 for men (0.76 for women), c^2 was 0.02 (0.02), and e^2 was 0.36 (0.22). At high school education, a^2 was 0.43 for men (0.58 for women), c^2 was 0.04 (0.03), and e^2 was 0.53 (0.39). At more than 4 years of education beyond highschool, a^2 was 0.15 for men (0.23 for women), c^2 was 0.07 (0.05), and e^2 was 0.78 (0.71).

Steenstrup et al. (2013) provided enough information to derive univariate values, using Falconer's formula (1952). The following calculations are for subjects between ages 35-54. For MCS, a^2 was 0.17 for men (0.31 for women), c^2 was 0.09 for men (-0.01), and e^2 was 0.73 for men (0.69). Given that the samples were derived from Nordic populations and I am unable to provide confidence intervals for these calculations, I consider these findings to be only partially informative for our US sample – at least slightly suggestive of an AE model for PCS and an ACE model for MCS.

3.4 Socioeconomic Status

I constructed an SES measure from the NLSY79 based on Myrianthopoulos and French (1968) and used more recently by Turkheimer et al. (2003). Each subject was given an aggregate score based on the mean of their total net family income, education, and occupation quantile scores. Subjects with missing data were not excluded – instead, their aggregated scores were created from their non-missing components. Higher scores correspond to higher socioeconomic status.

SES was computed for the same year as the Health 40 module. The mean SES for wave Health 40 was 52.1 (SD = 21.8). The distribution of this index by race and sex is illustrated in Figure 3.7. Average scores for women (52.8, SD = 21.9) were significantly higher than average scores for men (51.3, SD = 21.6); $t(7322.04) = 2.87$, $p < .001$. Black

respondent means (47.9, SD = 21.7) and Hispanic respondent means (48.2, SD = 21.6) were significantly lower than the mean for non-Black-non-Hispanic respondents (56.1, SD = 21.1), $t(7325.12) = -16.07$, $p < .001$.

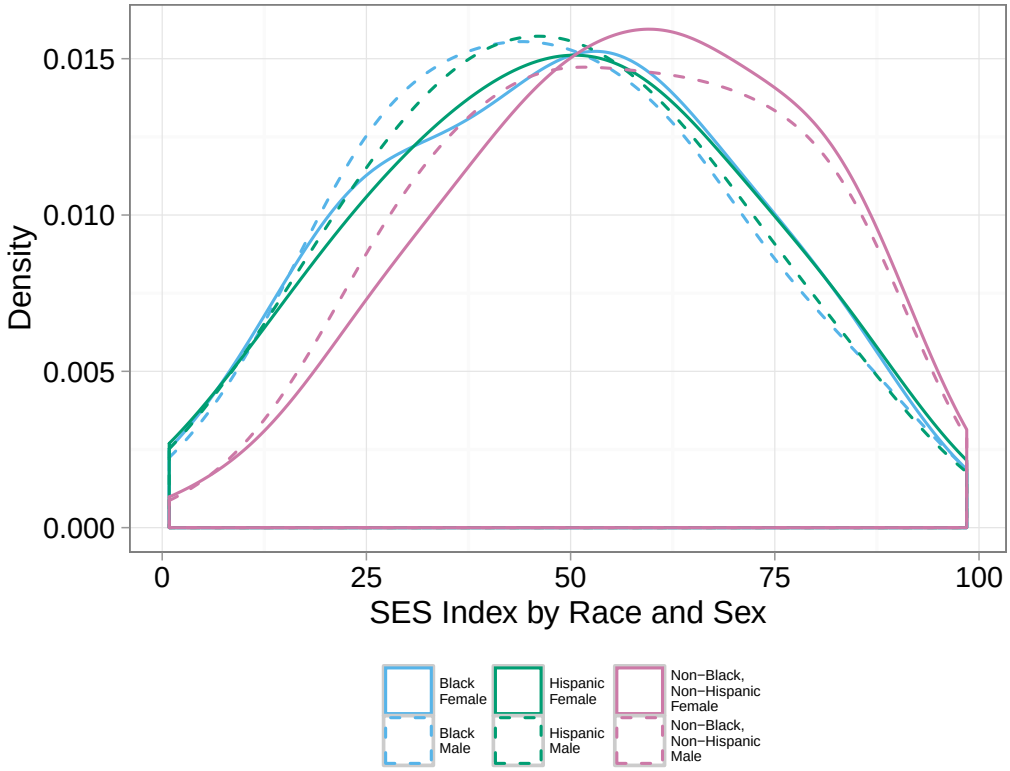

Figure 3.7: Density Plot of SES Index at Health 40 by Sex and Race

3.4.1 Heritability

I found no instances of Myrianthopoulos and French's 1968 system decomposed into its biometric components. However, recent reviews in economics provide evidence for both genetic ($a^2 \approx .5$) and shared-environment effects ($c^2 \approx .1$) for the components commonly used in SES composites (Benjamin et al., 2012), as do findings in psychology (Plomin & Bergeman, 1991; Rowe, Vesterdal, & Rodgers, 1998). Moreover, Lichtenstein et al. (1993) found evidence for genetic and environmental effects for their multiple components of SES.

Chapter 4

Results

In the full sample, the observed correlation between SES and the non-linearly rescaled PCS was 0.2 (the correlation for raw PCS was 0.2), whereas the observed correlation between SES and the transformed MCS was 0.08 (raw MCS was 0.11). Table 4.1 displays the correlation matrix for the Full Sample NLSY79 (see Table 4.2 for the National Probability Subsample correlation matrix). The diagonal indicates the sample size for each variable, and upper triangle reveals the number of respondents with viable scores for both respective variables. Kinship correlations are presented in Tables A.1 - A.5 in Appendix A. Because of the data missingness, all models were estimated using full information maximum likelihood (FIML). FIML was restricted to subjects with the same levels of relatedness employed in the models.

Table 4.1: Correlations and Sample Sizes, in Full Sample

	SES	PCS	MCS
SES	8465	8402	8402
PCS	0.198∗	8402	8402
MCS	0.083∗	-0.006	8402

Table 4.2: Correlations and Sample Sizes, in National Probability Subsample

	SES	PCS	MCS
SES	4789	4755	4755
PCS	0.213∗	4755	4755
MCS	0.089∗	-0.055∗	4755

4.1 Univariate Results

Table 4.3 reports the estimated variance components and model fit statistics for univariate ACE models of SES, PCS, and MCS, controlling for sex and minority status. I have included TLI and CFI in our summary statistics, but I did not use them to evaluate model fit because they penalize models heavily for complexity (Marsh, Hau, & Grayson, 2005).

4.1.1 SES

The best fitting model for the heritability of SES, controlling for minority status and sex, was an ACE model. In general, model fit statistics were excellent. The χ^2 test of the overall model was not significant ($p = 0.19$), indicating good model fit. RMSEA indicated that the model fit closely (0.01, 90% CI [0, 0.03]) and that perfect fit could not be rejected. SRMR was below .07, also indicating good model fit.

Minority status and sex accounted for 4% of the total variance in SES. After adjusting for minority status and sex, $a^2 = 48\%$ ($p < .001$; 95% CI [.20, .87]), $c^2 = 12\%$ ($p < .001$; 95% CI [.1, .34]), and the non-shared environment accounted for the remainder, $e^2 = 40\%$ ($p < .001$; 95% CI [.25, .60]).

4.1.2 PCS

The best fitting model for the heritability of PCS was an AE model. It was statistically indistinguishable from an ACE model ($\chi^2(1) = 0$, p = 1.00) because the c^2 estimate was zero in both models. In general, model fit statistics were good. Without controlling for covariates in the AE model, the χ^2 test of model fit (p = 0.35) was not significant and RMSEA (0.01, 90% CI [0, 0.03]) could not reject perfect fit. SRMR (0.07) was on the higher side, but still approximately .07. Controlling for covariates in the AE model, the χ^2 test of model fit (p = 0.02) was significant and RMSEA (0.02, 90% CI [0.01, 0.04]) could reject perfect fit, but not close fit. SRMR was lower (0.04) and below the .07 threshold.

For consistency, I have reported the fits and estimates for the AE model with covariates in Table 4.3. Taken together, the models paint a consistent finding of an AE model.

The variance estimates between the models with and without covariates were not significantly different. Minority status and sex accounted for 1% of the total variance in PCS. After adjusting for minority status and sex, $a^2 = 17\%$ ($p < .001$; 95% CI [.11, .25]), whereas the non-shared environment accounted for the remainder, $e^2 = 83\%$ ($p < .001$; 95% CI [.76, .90]).

4.1.3 MCS

The best fitting model for MCS was a CE model. It was statistically indistinguishable from an ACE model ($\chi^2(1) = 0$, p = 1.00) because the a^2 estimate was zero in both models. The χ^2 test was not significant ($p = 0.26$), indicating good model fit. RMSEA indicated that the model fit closely (0.01, 90% CI [0, 0.03]) and that perfect fit could not be rejected. SRMR was below .07, also indicating good model fit.

Minority status and sex accounted for 2% of the total variance in MCS. After adjusting for minority status and sex, the shared environment (c^2) accounted for 9% ($p < .001$; 95% CI [.06, .13]) of the variance in MCS, while the non-shared environment accounted for the remainder, $e^2 = 91\%$ ($p < .001$; 95% CI [.87, .95]).

Table 4.3: Estimates and Fit Statistics of Best Fitting Univariate Models

	ACE SES	AE PCS	CE MCS
A	0.48	0.17	
p(A)	0	0	
C	0.120		0.089
p(C)	0.004		0.000
E	0.40	0.83	0.91
p(E)	0	0	0
p(ChiSqM)	0.195	0.018	0.258
ChiSqM DF	35	36	36
p(ChiSqBaseline)	0	0	0
ChiSqBaseline DF	28	28	28
CFI	0.99	0.85	0.97
TLI	0.99	0.88	0.98
p(RMSEA) $<.05$	1	1	1
RMSEA 90CI LB	0.00	0.01	0.00
RMSEA	0.014	0.024	0.012
RMSEA 90CI UB	0.028	0.035	0.026
SRMR	0.025	0.043	0.029

4.2 Bivariate Results

Both meaningful genetic and shared environmental variance exist in the measure of SES. The question remains whether the genetic variance in PCS and the shared-environmental variance in MCS overlap with the equivalent variance source in SES. Thus, a series of

correlated-factors models were run to identify the best fitting models of the relationship between SES and a single measure of health. All models controlled for minority status and sex.

4.2.1 Bivariate models of SES and PCS

A nested model comparison found that allowing the genetic effects to covary improved the fit of the model dramatically; $p(\chi^2(1) < .001)$. Model fit statistics are provided in Table 4.4 and were generally good. Although the χ^2 test of model fit was significant ($p < .001$), RMSEA (0.03, 90% CI [0.03, 0.04]) indicated close fit. SRMR = 0.05 was low and below the .07 threshold. Standardized model parameter estimates are displayed in Figure 4.1 with standard errors in parentheses.

After adjusting for minority status and sex, the estimated parameters for SES were a^2 = 49% ($p < .001$; 95% CI [.21, .88]), c^2 = 9% ($p = 0.04$; 95% CI [.002, .33]), e^2 = 42% ($p < .001$; 95% CI [.27, .62]). The estimated parameters for PCS were a^2 = 11% ($p < .001$; 95% CI [.05, .19]) and e^2 = 89% ($p < .001$; 95% CI [.82, .97]). Unadjusted a^2, c^2, and e^2 estimates can be found by squaring the factor loadings in Figure 4.1. The bivariate estimates were nearly identical to the univariate estimates.

The estimated covariance between the genetic components of SES and PCS was 0.65 ($p < .001$) [0.31, 0.99], which means that 65% of the genetic influence was common to both measures. I calculated the genetic correlation ($\hat{r}_a$) between SES and PCS with path tracing rules (Wright, 1923, 1934).

$$\hat{r}_a = \beta_{PCS,a_{PCS}} * \psi_{a_{PCS},a_{SES}} * \beta_{SES,a_{SES}} \tag{4.1}$$

$$\hat{r}_a = 0.32 * 0.65 * 0.68 = 0.14 \tag{4.2}$$

I found that the genetic correlation ($\hat{r}_a$) was 0.14, 95% CI [0.12, 0.16]. Consequently, the proportion of the latent phenotypic correlation ($\hat{r}$ = 0.16) attributable to genetics was

0.91, 95% CI [0.73, 1]. As a more conservative test, I substituted the latent phenotypic correlation with the observed correlation between SES and PCS in the full sample (r = 0.2); the proportion explained by genetics was 0.73, 95% CI [0.59, 0.81]. Both proportions indicated that a large component of the gradient was heritable, and the remaining phenotypic variance was explained by the non-shared environment.

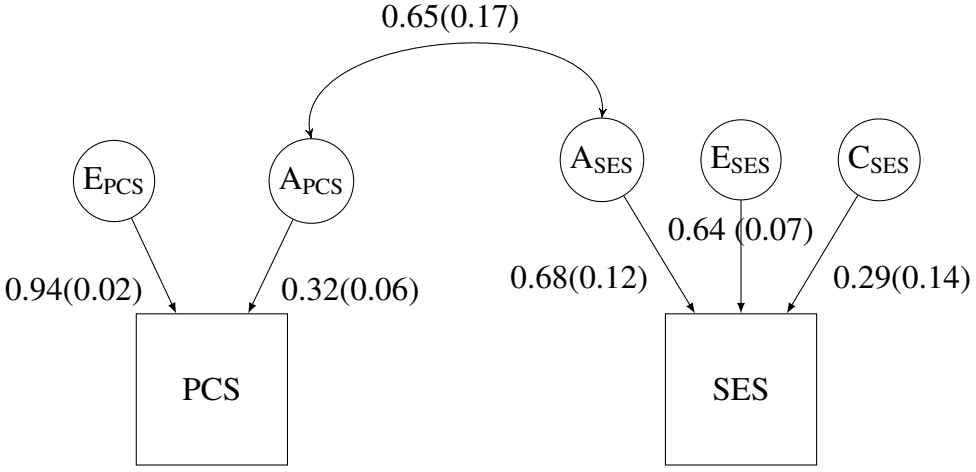

Figure 4.1: Correlated Factors of SES and PCS

Table 4.4: Bivariate Model Fit for PCS and SES

Parameters	18
ChiSqM Value	188.00
ChiSqM DF	86
p(ChiSqM)	0
p(ChiSqBaseline)	0
CFI	0.91
TLI	0.92
RMSEA 90CI LB	0.03
RMSEA	0.03
RMSEA 90CI UB	0.04
p(RMSEA) <.05	1
SRMR	0.05

4.2.2 Bivariate models of SES and MCS

A nested model comparison found that allowing the shared-environmental effects to covary improved the fit of the model dramatically, $p(\chi^2(1) < .001)$. Model fit statistics are provided in Table 4.5 and were generally excellent. The χ^2 test of model fit was not significant (p = 0.05), and RMSEA (0.02, 90% CI [0, 0.02]) could not reject perfect fit. SRMR = 0.03 was low and below the .07 threshold. Standardized model parameter estimates are displayed in Figure 4.2 with standard errors in parentheses.

After adjusting for minority status and sex, the estimated parameters for SES were a^2 = 46% ($p < .001$; 95% CI [.18, .86]), c^2 = 13% ($p < .001$; 95% CI [.1, .35]), e^2 = 42% ($p < .001$; 95% CI [.26, .61]). The estimated parameters for MCS were c^2 = 8% ($p < .001$;

95% CI [.05, .12]) and e^2 = 92% ($p < .001$; 95% CI [.88, .95]). Unadjusted a^2, c^2, and e^2 estimates can be found by squaring the factor loadings in Figure 4.2. I note that the bivariate estimates were nearly identical to the univariate estimates.

The covariance between the shared-environmental components of SES and MCS was 0.79 (p = 0.01; 95% CI [0.2, 1]), which means that 78.8% of the shared-environmental influences was common to both measures. I calculated the shared-environmental correlation ($\hat{r}_c$) between SES and MCS with path tracing rules (Wright, 1923, 1934).

I found that the shared-environmental correlation $\hat{r}_C$ = 0.08, 95% CI [0.06, 0.1]. Consequently, the proportion of the latent phenotypic correlation ($\hat{r}$ = 0.06) explained by the shared-environmental correlation was 1, 95% CI [0.97, 1]. As a more conservative test, I substituted the latent phenotypic correlation with the observed correlation between SES and MCS in the full sample (r = 0.08); this proportion explained by the shared environment was 0.95, 95% CI [0.69, 1].

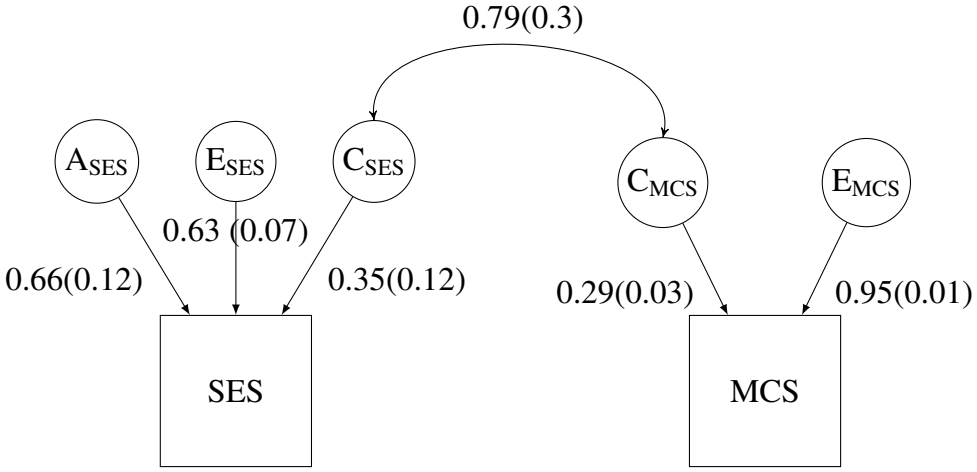

Figure 4.2: Correlated Factors of SES & MCS

Table 4.5: Bivariate Model Fit for MCS and SES

Parameters	18
ChiSqM Value	108.00
ChiSqM DF	86
p(ChiSqM)	0.05
p(ChiSqBaseline)	0
CFI	0.98
TLI	0.98
RMSEA 90CI LB	0
RMSEA	0.02
RMSEA 90CI UB	0.02
p(RMSEA) $<.05$	1
SRMR	0.03

4.3 Trivariate Results

Combining the bivariate models, I examined the influence of SES on both mental and physical health. A series of nested model comparisons found the following:

- allowing the shared-environmental effects between mental health and SES to covary improved the fit of the model $(p(\chi^2(1)) < .001)$;
- allowing the genetic effects between physical health and SES to covary improved the fit of the model $(p(\chi^2(1)) < .001)$;
- allowing the genetic effects between physical health and SES to covary after already

covarying the shared-environmental effects between mental health and SES improved the fit of the model ($p(\chi^2(1)) < .001$); and

- allowing the shared-environmental effects between mental health and SES to covary after already covarying the genetic-environmental effects between physical health and SES improved the fit of the model ($p(\chi^2(1)) < .001$).

Notably, in sensitivity analyses the effect on the χ^2 contribution from MZ twins was disproportionately high across all models. For example, in the final model, they contributed 40% of the total χ^2 (132.590 of 330.13), even though there were only 11 pairs (0.27%) of 4018. This proportion mismatch motivated us to run additional models, where I excluded MZ twins from the analysis. In the spirit of transparency, I have reported those results here as well, and note when the two model results differ. Model fit statistics are provided in Table 4.6 and were excellent. Fits were universally improved by the exclusion of MZ twins, but the estimates themselves did not differ greatly.

Table 4.6: Trivariate Model Fits

	Without Twins	With Twins
Parameters	24	27
ChiSqM Value	197	330
ChiSqM DF	111	153
p(ChiSqM)	0	0
p(ChiSqBaseline)	0	0
CFI	0.94	0.88
TLI	0.94	0.89
AIC	58471	58630
BIC	58622	58800
aBIC	58545	58714
RMSEA 90CI LB	0.019	0.029
RMSEA	0.024	0.034
RMSEA 90CI UB	0.030	0.039
p(RMSEA) <.05	1	1
SRMR	0.035	0.050

4.3.1 With MZ Twins

Model fit statistics were generally good. Although the χ^2 test of model fit ($p < .001$) was significant, RMSEA (0.03, 90% CI [0.03, 0.04]) indicated close fit. SRMR = 0.05 was low and below the .07 threshold.

After adjusting for minority status and sex, the estimated parameters for SES were a^2 = 48% ($p < .001$; 95% CI [.21, .85]), c^2 = 9% ($p = 0.03$; 95% CI [.01, .32]), e^2 = 43%

($p < .001$; 95% CI [.28, .62]). The estimated parameters for PCS were a^2 = 11% ($p < .001$; 95% CI [.05, .19]) and e^2 = 89% ($p < .001$; 95% CI [.82, .96]). The estimated parameters for MCS were c^2 = 9% ($p < .001$; 95% CI [.05, .12]) and e^2 = 91% ($p < .001$; 95% CI [.88, .95]). Standardized model parameter estimates are displayed in Figure 4.3 with standard errors in parentheses. Unadjusted a^2, c^2, and e^2 estimates can be derived by squaring the factor loadings.

The covariance between the genetic components of SES and PCS was 0.66 ($p < .001$; 95% CI [0.32, 1]). I calculated the genetic correlation ($\hat{r}_a$) between SES and PCS with path tracing rules (Wright, 1923, 1934). I found that the genetic correlation ($\hat{r}_a$) was 0.15, 95% CI [0.12, 0.16]. Consequently, the proportion of the latent phenotypic correlation ($\hat{r}$ = 0.16) attributable to genetics was 0.91, 95% CI [0.73, 1]. As a more conservative test, I substituted the latent phenotypic correlation with the observed correlation (r = 0.2) between SES and PCS in the full sample; the proportion explained by genetics was 0.74, 95% CI [0.59, 0.81]. Both proportions indicated that a large component of the gradient was heritable, and the remaining proportion was explained by the non-shared environment.

The covariance between the shared-environmental components of SES and MCS was 0.95 ($p = 0.04$; 95% CI [0.06, 1]). I found that the shared-environmental correlation ($\hat{r}_C$) was 0.08, 95% CI [0.06, 0.1]. Consequently, the proportion of the latent phenotypic correlation ($\hat{r}$ = 0.06) explained by the shared-environmental correlation was 1, 95% CI [0.97, 1]. As a more conservative test, I substituted the latent phenotypic correlation with the observed correlation (r = 0.08) between SES and MCS in the full sample; this proportion explained by the shared environment was 0.96, 95% CI [0.7, 1].

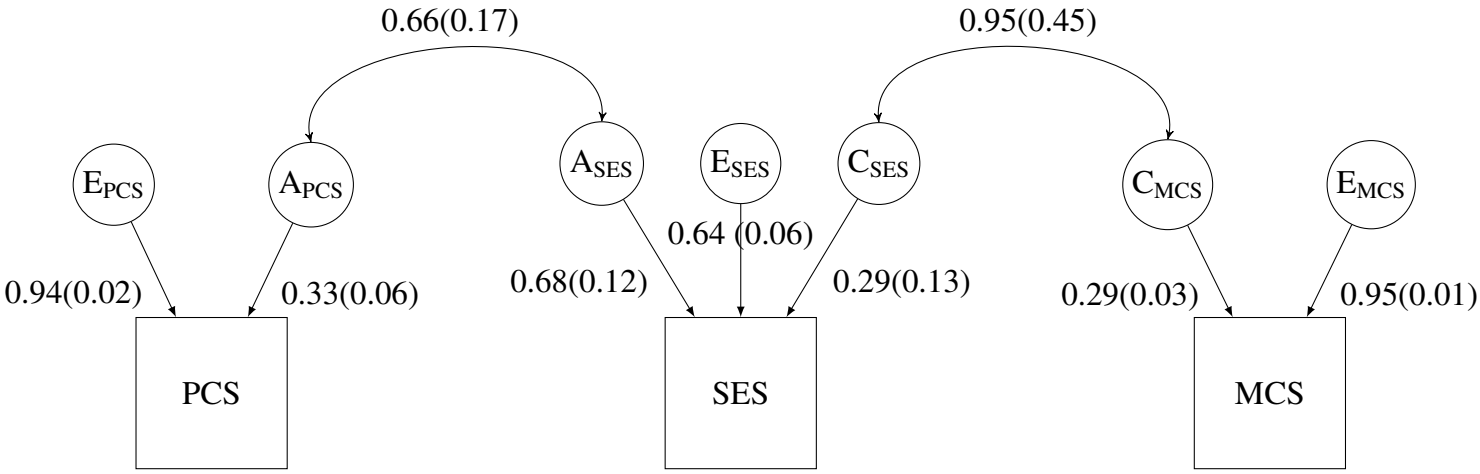

Figure 4.3: Correlated Factors of SES, Physical Health, & Mental Health at Age 40

4.3.2 Without MZ Twins

Model fit statistics were generally excellent. Although the χ^2 test of model fit ($p < .001$) was significant, RMSEA (0.02, 90% CI [0.02, 0.03]) indicated close fit. SRMR was low (0.04) and below the .07 threshold.

After adjusting for minority status and sex, the estimated parameters for SES were a^2 = 49% ($p < .001$; 95% CI [.37, 64]), c^2 = 8 % ($p < .001$; 95% CI [.04, .15]), e^2 = 42% ($p < .001$; 95% CI [.34, .52]). The estimated parameters for PCS were a^2 = 10% ($p < .001$; 95% CI [.05, .19]) and e^2 = 90% ($p < .001$; 95% CI [.82, .97]). The estimated parameters for MCS were c^2 = 8% ($p < .001$; 95% CI [.05, .12]) and e^2 = 92% ($p < .001$; 95% CI [.88, .95]). Standardized model parameter estimates are displayed in Figure 4.4 with standard errors in parentheses. Unadjusted a^2, c^2, and e^2 estimates can be derived by squaring the factor loadings.

The covariance between the genetic components of SES and PCS was 0.65 ($p < .001$; 95% CI [0.37, 0.94]). I calculated the genetic correlation ($\hat{r}_a$) between SES and PCS with path tracing rules (Wright, 1923, 1934). I found that the genetic correlation ($\hat{r}_a$) was 0.14, 95% CI [0.12, 0.16]. Consequently, the proportion of the latent phenotypic correlation ($\hat{r}$ = 0.16) attributable to genetics was 0.88, 95% CI [0.72, 0.98]. As a more conservative

test, I substituted the latent phenotypic correlation with the observed correlation between SES and PCS in the full sample (r = 0.2); the proportion explained by genetics was 0.73, 95% CI [0.59, 0.81]. Both proportions indicated that a large component of the gradient was heritable, and the remaining proportion was explained by the non-shared environment.

The covariance between the shared-environmental components of SES and MCS was 1 ($p < .001$; 95% CI [0.89, 1]). I found that the shared-environmental correlation ($\hat{r}_C$) was 0.08, 95% CI [0.06, 0.1]. Consequently, the proportion of the latent phenotypic correlation ($\hat{r}$ = 0.06) explained by the shared-environmental correlation was 1, 95% CI [0.94, 1]. As a more conservative test, I substituted the latent phenotypic correlation with the observed correlation between SES and MCS in the full sample (r = 0.08); this proportion explained by the shared environment was 0.96, 95% CI [0.71, 1].

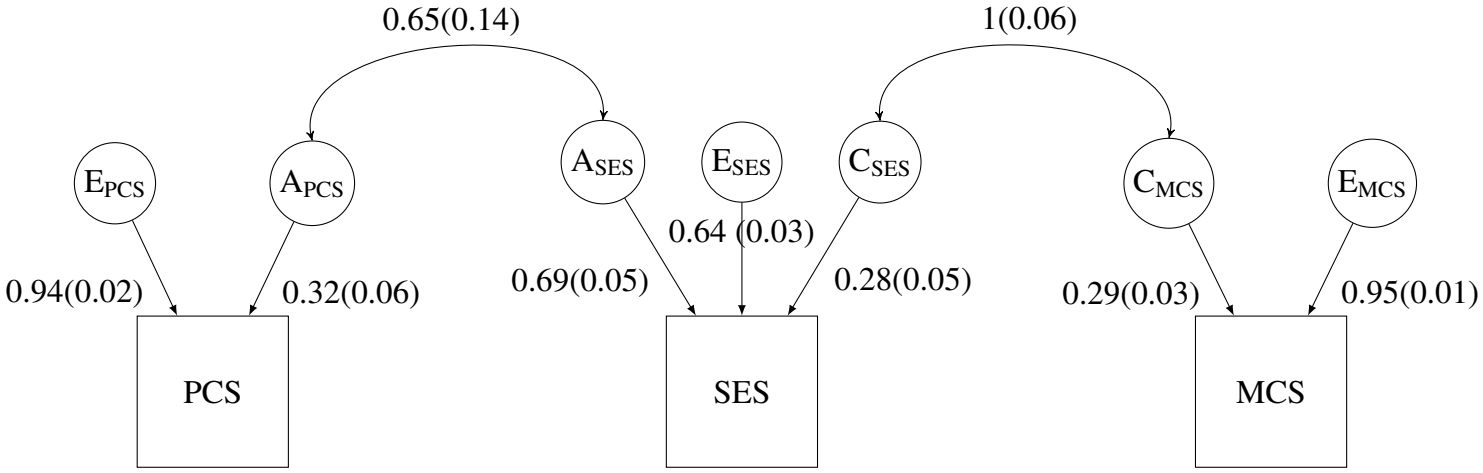

Figure 4.4: Correlated Factors of SES, Physical Health, & Mental Health at Age 40 (without MZ twins)

4.3.3 Model Differences

Model fits were noticeably improved on all indices with the exclusion of MZ twins. The only notable parameter difference was the shared-environmental covariance between SES and mental health. The MZ twin model estimated a correlation of 0.95 ($p = 0.04$), whereas the MZ-less model provided an estimated shared-environmental correlation of 1

($p < .001$). Practically, the magnitude of the effect did not change, only the precision of the estimate. With MZ twins, the 95% confidence interval was [0.06, 1], whereas without MZ twins, the 95% confidence interval was [0.89, 1].

Chapter 5

Discussion

This article presents a biometrical decomposition of the bivariate relationship between SES and Health – often called the SES-health gradient – for both mental and physical health at age 40. The identified general pathways of influence of the SES-Health gradient differ by aspect of health. SES links to mental health through the shared-environment. In contrast, SES links to physical health through genetic and non-shared environmental pathways.

For physical health, our results are consistent with the Lichtenstein et al. (1993) general finding of a genetic pathway linking physical health and SES. In this study, the genetic pathway explained the bulk of the gradient, even using the most conservative proportion I could devise by substituting the latent phenotypic correlation with the observed correlation. For mental health, however, my findings diverge. I observed no genetic effect and instead found the shared-environment explain practically all of the mental health gradient at age 40. Again, this relationship held even when using the most conservative proportion I could devise.

5.1 Physical Health

Across models, I consistently found genetic variance to underlie the gradient for physical health. This percentage ranged from 70% genetic and 30% non-shared environment in conservative calculations to 90% genetic and 10% non-shared environment in more traditional calculations. The consistency of this finding was comparable to those of Lichtenstein et al. (1993)'s Nordic sample (they found 67% genetic variance) – even though they used a different type of sample, twins raised apart – lends more support to this result.These results suggest that Adler and Stewart (2010) and Gottfredson (2004) are both correct – the environment matters as do genes, but the nature of the health outcome is critical.

As discussed in the (Mis)interpreting heritability section, this finding does not mean that the gradient is fixed or cannot respond to intervention. On the contrary, the high proportion of genetic variance simply identifies the source of overlap between SES and health as emergent from biological sources. Both gene-environment interaction and epigenetic sources provide many explanations for why such genetic overlap does not imply deterministic outcomes.

5.2 Mental Health

Across models, I consistently found the shared-environment to be the primary source of the gradient for mental health at age 40. This percentage ranged from 90% shared-environment and 10% non-shared environment in conservative calculations to 100% shared environmental in more traditional calculations. Regardless of the method of calculation, the consistency of the result indicates that the shared-environment explains practically all of the small correlation between SES and mental health, suggesting that early experiences in the home may be important influences on the mental health aspect of the gradient. This result maps cleanly onto indirect models of Social Causation (and Selection). Further research is needed to explicitly test the specific theories within Social Causation, such as whether early home stressors mediate the mental health gradient (Repetti et al., 2002).

5.3 Theoretical Implications

Theorists (notably, Adler et al., 1994; Adler & Ostrove, 1999; Adler & Stewart, 2010) assert that Social Causation and Selection theories act exclusively through the environment, whereas Social Confound theories act exclusively through genes. Using their framework, the genetic effect of physical health with SES provides definitive support for Social Confound theories and all but eliminates Social Causation (and Selection) theories; whereas the shared-environmental effect of mental health with SES maps onto Social Causation (and Selection) theories, accordingly eliminating Social Confound explanations. Such an

interpretation may be overly simplistic.

In reality, the relationship between SES and health is dynamic, temporal – in other words, complex. For example, schizophrenia is often used as an illustration of selection (Goldberg & Morrison, 1963). Accordingly, theories of selection would predict that schizophrenia is not heritable and predominantly explained by environmental sources. Yet, schizophrenia is heritable (estimates range from 64% to 89%, National Institute of Mental Health's Genetics Workgroup, 1998; Tsuang, 2000). How can SES-health gradient theory and behavior genetics appear so incompatible? They are not. The estimates given by behavior genetics refer to distal causes, whereas the gradient theories we have refer to proximal causes. Instead, we ought to think of these two pieces as giving predictions at different points in the causal stream. The proximal causes are conditioned on distal causes. For example, given an individual genetically predisposed to developing schizophrenia, Social Selection theories predict the following chain of events:

- The at-risk individual experiences schizophrenia, then
- a downward drift into poverty occurs;

whereas Social Causation theories predict an alternative chain of events:

- The at-risk individual experiences poverty, then
- onset of schizophrenia occurs (*i.e.*, a decline in health).

In both models, the distal cause, (*e.g.*, genetic predisposition toward schizophrenia), is a necessary step in the causal stream. However, in neither model, is having the genetic predisposition sufficient. In the Causation Model, being vulnerable to schizophrenia is insufficient without exposure to poverty; whereas in the Selection Model, being vulnerable to schizophrenia is insufficient without the actual onset of schizophrenia. In other words, only if the distal cause is present, can the proximal causes proposed by Selection, Causation, and Confounding theories be relevant within the causal stream. It is in this sense

that I suggest interpreting the findings presented in this paper. The illustration above used schizophrenia – the logic holds for all heritable traits. However, Adler and others have framed genetic effects as caused by third variables. There are many traits with substantial heritabilities that could be "third variable" more proximate causal influences on health. Accordingly, more research should be dedicated to identifying what these third variables are. A reasonable starting point would be examining whether (and which) individual differences in personality and cognitive ability explain the identified genetic overlap. Recent work by Arden et al. (2016) and Trzaskowski et al. (2014) both identify that genes are a large source of covariation for health and wealth with intelligence. The next step would be to test whether those common genetic sources are common to one another, using models and analytic approaches similar to those in the current study.

If individual differences in personality are the third variables at the heart of the gradient, then recent advances in research on personality change and growth-mindset interventions can potentially provide the groundwork for larger scale interventions (*e.g.*, Hudson & Fraley, 2015; Magidson, Roberts, Collado-Rodriguez, & Lejuez, 2014; Paunesku et al., 2015). Just as the impact of PKU can be circumvented with changes in diet, the gradient can potentially be affected by adapting personality change interventions. As a result, the genetic proportion of the gradient should decline with time because post-intervention, there would be more homogeneity in people acting in health-preserving and wealth-generating manners.

5.4 Caveats and Conclusions

5.4.1 Measures

Health. The measures of health are not diagnosis-specific, but are overall measures of health. Accordingly, specific diagnoses may have different distributions of genetic and environmental influences, which may in turn map onto different theoretical pathways.

For example, although mental health symptoms (*e.g.*, depressive symptomatology; Byers, Levy, Kasl, Bruce, & Allore, 2009) and many mental health diagnoses (Burt, 2009) have a shared-environmental component, genetic influences vary – specifically, depressive symptomatology (Byers et al., 2009) may lack a genetic influence, whereas clinical diagnoses of major depressive disorder have a well-established genetic component (Sullivan, Neale, & Kendler, 2000). Thus, depressive symptomatology may share variance through the shared-environment, whereas major depressive disorder may share variance through various genetic confounds. Again, further research is needed to untangle the mechanisms for more disorders to determine whether these findings generalize beyond general health.

SES. There is no definitive measure for socioeconomic status – although numerous task forces have attempted to create such a measure (*e.g.*, Saegert et al., 2006). Many studies referenced throughout this paper have employed a single component of SES (*e.g.*, household income, education, *etc.*), and used it as a proxy for the entire construct. Other studies, notably (Lichtenstein et al., 1993), report results using multiple components of SES. Combined, these studies cast a nomological net, allowing us to understand the interrelated nature of SES. This paper's purpose, however, was to understand the biometrical underpinnings of the SES-health gradient – not the SES-components-gradient. Thus, I made a deliberate choice to focus on the higher-level construct. Hence, I employed an overall index (similar to the indices in Myrianthopoulos & French, 1968; Turkheimer et al., 2003), rather than to conduct repeated analyses on multiple components of SES, as did Lichtenstein et al. (1993). Further research is needed to examine which components of SES contribute to the overall relationship between SES and health.

5.4.2 Modeling

Assumptions. Like all work in behavior genetics, these biometrical models impose simplifying assumptions. We explicitly assumed no interaction between genotype and environment. Robustness checks that we did not report here did not find gene-by-environment

interactions; however, as discussed previously, genetic distal causes may interact with environmental proximal causes. Indeed, recent work by Adler (2013) has acknowledged that genetics (particularly, epigenetics) may be important elements underlying the gradient. Future work that explicitly models gene-by-environment interactions could do much to further untangle the gradient, beyond our initial decompositions presented herein.

The models assumed random mating. However, partners tend to marry and mate with individuals alike on many characteristics (Buss, 1985), including SES (Watkins & Meredith, 1981) and health (Meyler, Stimpson, & Peek, 2007). Accordingly, the genetic similarity between kin may be higher than what we modeled. Thereby our heritability estimates may be deflated and our estimates of environmentality may be inflated. Particularly, the inflation of the shared-environmental influence may explain why the proportion of the mental health gradient was larger than typical estimates in the literature. Other designs, such as intergenerational designs or the inclusion of spouses can untangle the influence of non-random assortment (*e.g.*, Cardon, Fulker, & Jöreskog, 1991; Heath & Eaves, 1985). Moreover, simulation studies find that typical levels of non-random mating are not large enough to drastically impact biometrical models (Fernando & Gianola, 1990), as do analytic derivations (Reeve, 1961).

Finally, the major assumption of equal environments assumed that siblings of different levels of relatedness (full siblings vs half siblings) are equally exposed to "environmental events of etiologic importance" (Kendler et al., 1993). When equal environments are incorrectly assumed, genetic effects will be overestimated. Much of the debate over the validity of the equal environments assumption focuses on differential treatment between monozygotic and dizygotic twins, and accordingly, tests of equal environments use twin-specific situations, including misdiagnosis of zygosity (Matheny, 1979). Our sample included a wider range of kin relationships. We assumed that monozygotic twins, full-siblings, half-siblings and cousins were treated the same with respect to both etiological precursors of later health and SES. However, fairy tales such as Cinderella (step-siblings, Grimm, 1955)

or Ye Xian (half-siblings, Louie, 1982) and sagas such as Harry Potter (cousins, Rowling, 1997) provide us with illustrations of differential treatment. Still, for the equal environments assumption to be violated, the differential experiences must be of etiological influence to either later SES or later health – not merely being differentially treated with respect to those constructs. Although this is a complicated domain, we do note that our estimates using multiple levels of kin pairs were similar to those with twins of varying zygosity (Lichtenstein et al., 1993).

Alternative Approaches. Alternative modeling approaches such as Cholesky models impose more assumptions – mainly temporal ordering. We explicitly focused on a cross-sectional snapshot of the USA, which inhibited us from imposing a direction. We recommend that future papers employ this family of models to untangle the direction of the relationship between SES and health, though remaining in a behavior genetic framework.

5.4.3 Design

Our findings offer a snapshot of the United States population at age 40, between 1998 and 2006, as portrayed by the NLSY79 data. The effects may vary by age. The shared environment (and the theories explained by the shared environment) may be a more potent influence on the gradient for younger individuals because their childhood experiences are more salient. I plan to replicate and incorporate repeated measures once the BLS finishes collecting health at age 50 in 2016 (and releases it by 2020). At that point, I can explicitly untangle whether the environmental influences are driven by selection or causation.

5.4.4 Final Remarks

This paper decomposed the relationship between socioeconomic status and health into its biometrical components. The results differed by measure of health. Physical health's relationship with SES was primarily explained through genes, whereas mental health's relationship with SES was primarily explained through the shared environment. If we interpret

these findings through the genes versus environment framework that theorists (notably, Adler et al., 1994; Adler & Ostrove, 1999; Adler & Stewart, 2010) impose, then these results imply the following:

- the physical health gradient is a product of Social Confounding, and is caused by third variables, such as intelligence and personality;
- the mental health gradient is a product of Social Causation, and suggests that socioeconomic status causes disparities in mental health.

Such an interpretation is overly simplistic. Rather, the results suggest that there are genetic precursors common to both physical health and SES, and that there are shared-environmental influences common to both mental health and SES. These genetic precursors do not necessarily imply that third variables are the cause of the SES-physical health gradient. Just as these shared-environmental influences do not necessarily imply that socioeconomic status causes disparities in mental health. We integrate Adler and Stewart (2010)'s interpretations with behavior genetics to conclude the following: At age 40,

- the physical health gradient has genetic precursors, that potentially are explained by third variables, such as intelligence and personality;
- the mental health gradient has shared environmental sources, and are suggestive of a social causation model.

References

Abramson, C. (2015, apr). Unequal until the end. *The Atlantic*. Retrieved from http://www.theatlantic.com/business/archive/2015/04/unequal-until-the-end/389910/

Adler, N. E. (2013). Health disparities: Taking on the challenge. *Perspectives on Psychological Science*, *8*(6), 679–681.

Adler, N. E., Boyce, W. T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist*, *49*(1), 15–24.

Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: What we know and what we don't. *Annals of the New York Academy of Sciences*, *896*(1), 3–15.

Adler, N. E., & Stewart, J. (2010). Health disparities across the lifespan: Meaning, methods, and mechanisms. *Annals of the New York Academy of Sciences*, *1186*(1), 5–23.

Alfert, E. (1969a). Comment on: The Promotion of Prejudice. *Journal of Social Issues*, *25*(4), 206–212.

Alfert, E. (1969b). Response to Jensen's rejoinder. *Journal of Social Issues*, *25*(4), 217–219.

Ameri, M., Schur, L., Adya, M., Bentley, S., McKay, P., & Kruse, D. (2015). *The disability employment puzzle: A field experiment on employer hiring behavior.* Retrieved from http://www.nber.org/papers/w21560

Angadi, S. (2015, jul). We can't all go to fancy yoga classes. *Time Magazine*. Retrieved from http://time.com/3975136/healthy-living-ideas/

Antonovsky, A. (1967). Social class, life expectancy and overall mortality. *Milbank Memorial Fund Quarterly*, *45*(2), 31–73.

Arden, R., Luciano, M., Deary, I. J., Reynolds, C. A., Pedersen, N. L., Plassman, B. L., ... Visscher, P. M. (2016). The association between intelligence and lifespan is mostly genetic. *International Journal of Epidemiology*, *45*(1), 178–185.

Batty, G. D., Deary, I. J., & Gottfredson, L. S. (2007). Premorbid (early life) IQ and later mortality risk: Systematic review. *Annals of Epidemiology*, *17*(4), 278–288.

Baum, A., Garofalo, J. P., & Yali, A. M. (1999). Socioeconomic status and chronic stress. Does stress account for SES effects on health? *Annals of the New York Academy of Sciences*, *896*, 131–144.

Behling, O. (1998). Employee selection: Will intelligence and conscientiousness do the job? *Academy of Management Perspectives*, *12*(1), 77–86.

Beier, M. E., & Ackerman, P. L. (2003). Determinants of health knowledge: An investigation of age, gender, abilities, personality, and interests. *Journal of Personality and Social Psychology*, *84*(2), 439–448.

Benjamin, D. J., Cesarini, D., Chabris, C. F., Glaeser, E. L., Laibson, D. I., Gunason, V., ... Lichtenstein, P. (2012). The promises and pitfalls of genoeconomics. *Annual Review of Economics*, *4*(1), 627–662.

Bergen, S. E., Gardner, C. O., & Kendler, K. S. (2007). Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: A meta-analysis. *Twin Research and Human Genetics*, *10*(3), 423–433.

Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, *130*(6), 887–919.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 211–252.

Braveman, P. A., & Gottlieb, L. (2014). The social determinants of health: It's time to consider the causes of the causes. *Public Health Reports*, *129*.

Burt, S. A. (2009). Rethinking environmental contributions to child and adolescent psychopathology: A meta-analysis of shared environmental influences. *Psychological Bulletin*, *135*(4), 608–637.

Buss, D. M. (1985). Human mate selection: Opposites are sometimes said to attract, but

in fact we are likely to marry someone who is similar to us in almost every variable. *American Scientist*, 47–51.

Byers, A. L., Levy, B. R., Kasl, S. V., Bruce, M. L., & Allore, H. G. (2009). Heritability of depressive symptoms: A case study using a multilevel approach. *International Journal of Methods in Psychiatric Research*, *18*(4), 287–96.

Cardon, L. R., Fulker, D. W., & Jöreskog, K. G. (1991). A LISREL 8 model with constrained parameters for twin and adoptive families. *Behavior Genetics*, *21*(4), 327–350.

Catroppa, C., Anderson, V. A., Morse, S. A., Haritou, F., & Rosenfeld, J. V. (2008). Outcome and predictors of functional recovery 5 years following pediatric traumatic brain injury (TBI). *Journal of Pediatric Psychology*, *33*(7), 707–718.

Chapman, B. P., Fiscella, K., Kawachi, I., & Duberstein, P. R. (2010). Personality, socioeconomic status, and all-cause mortality in the United States. *American Journal of Epidemiology*, *171*(1), 83–92.

Chapman, B. P., Roberts, B. W., & Duberstein, P. R. (2011). Personality and longevity: Knowns, unknowns, and implications for public health and personalized medicine. *Journal of Aging Research*.

Chirikos, T. N., & Nestel, G. (1985). Further evidence on the economic effects of poor health. *Review of Economics and Statistics*, *67*(1), 61–69.

Christensen, G. T., Mortensen, E. L., Christensen, K., & Osler, M. (2016). Intelligence in young adulthood and cause-specific mortality in the Danish Conscription Database A cohort study of 728,160 men. *Intelligence*.

Collins, S. D. (1926). Economic status and health. A review and study of the relevant morbidity and mortality data. *Public Health Bulletin*(165).

Conti, G., Heckman, J., & Urzua, S. (2010). The education-health gradient. *American Economic Review*, *100*(2), 234.

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors

by education. *Journal of Health Economics*, *29*(1), 1–28.

Davies, L. E. (1969, may). Harassment charged by author of article about negroes' I.Q.'s. *The New York Times*, 33.

Deary, I. J. (2010). Cognitive epidemiology: Its rise, its current issues, and its challenges. *Personality and Individual Differences*, *49*(4), 337–343.

Deary, I. J. (2012). Looking for 'system integrity' in cognitive epidemiology. *Gerontology*, *58*(6), 545–553.

Deary, I. J., & Johnson, W. (2010). Intelligence and education: Causal perceptions drive analytic processes and therefore conclusions. *International Journal of Epidemiology*, *39*(5), 1362–1369.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21.

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest*, *11*(2), 53–79.

Dennett, D. C. (2015). *Elbow room: The varieties of free will worth wanting*. MIT Press.

Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development*, 649–665.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology*, *91*(1), 40–57.

Duesenberry, J. S. (1967). *Income, saving, and the theory of consumer behavior* (Vol. 180). Oxford University Press.

Education Testing Services. (2016). *2016-17 GRE Guide to the Use of Scores - Ta-*

ble 4 (Tech. Rep.). Retrieved from https://www.ets.org/s/gre/pdf/gre_guide_table4.pdf

Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers in Psychology*, *5*, 36.

European Commission. (2012). *Europeans and Their Languages* (Tech. Rep.). Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf

Evans, G. W., & English, K. (2002). The environment of poverty: Multiple stressor exposure, psychophysiological stress, and socioemotional adjustment. *Child Development*, *73*(4), 1238–48.

Evans, W., Wolfe, B., & Adler, N. E. (2012). The SES and health gradient: A brief review of the literature. In *Biological consequences of socioeconomic inequalities.*

Falconer, D. S. (1952). The problem of environment and selection. *American Naturalist*, *86*(830), 293–298.

Fernando, R. L., & Gianola, D. (1990). Statistical inferences in populations undergoing selection or non-random mating. In *Advances in statistical methods for genetic improvement of livestock* (pp. 437–453). Springer.

Fölling, A. (1934). Über Ausscheidung von Phenylbrenztraubensäure in den Harn als Stoffwechselanomalie in Verbindung mit Imbezillität. *Hoppe-Seyler's Zeitschrift für physiologische Chemie*, *227*(1-4), 169–181.

Fox, J., & Weisberg, S. (2010). *An R companion to applied regression*. Sage.

Fox, J. W. (1990). Social class, mental illness, and social mobility: The social selection-drift hypothesis for serious mental illness. *Journal of Health and Social Behavior*, 344–353.

Gale, C. R., Batty, G. D., Cooper, C., & Deary, I. J. (2009). Psychomotor coordination and intelligence in childhood and health in adulthood–testing the system integrity

hypothesis. *Psychosomatic Medicine*, *71*(6), 675–681.

Garrison, S. M., & Rodgers, J. L. (2016). [In Press] Casting doubt on the causal link between intelligence and age at first intercourse: A cross-generational sibling comparison design using the NLSY. *Intelligence*.

Goldberg, E. M., & Morrison, S. L. (1963). Schizophrenia and social class. *British Journal of Psychiatry*, *109*, 785–802.

Goldberger, A. S. (1979). Heritability. *Economica*, *46*(184), 327–347.

Goldsmith, H. H. (1993). Nature-nurture issues in the behavioral genetic context: Overcoming barriers to communication. In G. McClearn & R. Plomin (Eds.), *Nature, nurture, and psychology* (pp. 325–339). Washington, DC: American Psychological Association.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*(1), 13–23.

Gottfredson, L. S. (2004). Intelligence: Is it the epidemiologists' elusive "fundamental cause" of social class inequalities in health? *Journal of Personality and Social Psychology*, *86*(1), 174–199.

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science*, *13*(1), 1–4.

Grimm, W. (1955). *Grimm's fairy tales* (Vol. 12). Hayes Barton Press.

Haas, S. A. (2006). Health selection and the process of social stratification: The effect of childhood health on socioeconomic attainment. *Journal of Health and Social Behavior*, *47*(4), 339–354.

Hampson, S. E., Goldberg, L. R., Vogt, T. M., & Dubanoski, J. P. (2007). Mechanisms by which childhood personality traits influence adult health status: Educational attainment and healthy behaviors. *Health Psychology*, *26*(1), 121–125.

Harris, J. R., Pedersen, N. L., McClearn, G. E., Nesselroade, J. R., & Plomin, R. (1992). Age differences in genetic and environmental influences for health from the Swedish

Adoption/Twin Study of Aging. *Journal of Gerontology*, *47*(3), P213–P220.

Heath, A. C., & Eaves, L. J. (1985). Resolving the effects of phenotype and social background on mate selection. *Behavior Genetics*, *15*(1), 15–30.

Hernstein, R. J., & Murray, C. (1994). *The bell curve: The reshaping of American life by differences in intelligence*. New York: Free Press.

Hudson, N. W., & Fraley, R. C. (2015). Volitional personality trait change: Can people choose to change their personality traits? *Journal of Personality and Social Psychology*, *109*(3), 490–507.

Hummer, R. A., Rogers, R. G., & Eberstein, I. W. (1998). Sociodemographic differentials in adult mortality. *Population and Development Review*, *24*(3), 553–578.

Jackson, J. J., Connolly, J. J., Garrison, S. M., Leveille, M. M., & Connolly, S. L. (2015). Your friends know how long you will live: A 75-year study of peer-rated personality traits. *Psychological Science*, *26*(3), 335–340.

Jensen, A. R. (1969a). Counter response. *Journal of Social Issues*, *25*(4), 219–222.

Jensen, A. R. (1969b). How much can we boost IQ and scholastic achievement. *Harvard Educational Review*, *39*(1), 1–123.

Jensen, A. R. (1969c). Rejoinder: Promotion of dogmatism. *Journal of Social Issues*, *25*(4), 212–217.

Johnson, W., Kyvik, K. O., Mortensen, E. L., Skytthe, A., Davids Batty, G., & Deary, I. J. (2010). Education reduces the effects of genetic susceptibilities to poor physical health. *International Journal of Epidemiology*, *39*(2), 406–414.

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, *52*(3).

Judge, T. A., Martocchio, J. J., & Thoresen, C. J. (1997). Five-factor model of personality and employee absence. *Journal of Applied Psychology*, *82*(5), 745.

Kell, H. J., Lubinski, D., & Benbow, C. P. (2013). Who rises to the top? Early indicators.

Psychological Science, *24*(5), 648–659.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A test of the equal-environment assumption in twin studies of psychiatric illness. *Behavior Genetics*, *23*(1), 21–7.

Keynes, J. M. (1936). The general theory of employment, interest and money. *Quarterly Journal of Economics*, *51*(2), 209–223.

Kuncel, N. R., Ones, D. S., & Sackett, P. R. (2010). Individual differences as predictors of work, educational, and broad life outcomes. *Personality and Individual Differences*, *49*(4), 331–336.

Levine, S. Z. (2013). Evaluating the seven-item Center for Epidemiologic Studies Depression Scale short-form: A longitudinal US community study. *Social Psychiatry and Psychiatric Epidemiology*, *48*(9), 1519–1526.

Lewontin, R. C. (1991). *Biology as ideology: The doctrine of DNA*. New York: Harper Collins.

Lichtenstein, P., Harris, J. R., Pedersen, N. L., & McClearn, G. E. (1993). Socioeconomic status and physical health, how are they related? An empirical study based on twins reared apart and twins reared together. *Social Science & Medicine*, *36*(4), 441–450.

Lievens, F., Ones, D. S., & Dilchert, S. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology*, *94*(6), 1514–1535.

Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social Science Research*, *37*(3), 888–902.

Lodi-Smith, J., Jackson, J. J., Bogg, T., Walton, K., Wood, D., Harms, P., & Roberts, B. W. (2010). Mechanisms of health: Education and health-related behaviours partially mediate the relationship between conscientiousness and self-reported physical health. *Psychology and Health*, *25*(3), 305–319.

Louie, A.-L. (1982). *Yeh-Shen, A Cinderella Story*. New York: Philomel

books. Retrieved from http://www.unc.edu/$\sim$rwilkers/resource-china.htm

Lubinski, D. (2009). Cognitive epidemiology: With emphasis on untangling cognitive ability and socioeconomic status. *Intelligence*, *37*(6), 625–633.

Lykken, D. T., McGue, M., Tellegen, A., & Bouchard, T. J. (1992). Emergenesis. Genetic traits that may not run in families. *American Psychologist*, *47*(12), 1565–1577.

Lynch, J. W., Smith, G. D., Harper, S., Hillemeier, M., Ross, N., Kaplan, G. a., & Wolfson, M. (2004). Is income inequality a determinant of population health? Part 1. A systematic review. *Milbank Quarterly*, *82*(1), 5–99.

Magidson, J. F., Roberts, B. W., Collado-Rodriguez, A., & Lejuez, C. W. (2014). Theory-driven intervention for changing personality: expectancy value theory, behavioral activation, and conscientiousness. *Developmental psychology*, *50*(5), 1442–50.

Major, J., Johnson, W., & Deary, I. J. (2014). Trait complexes of cognitive abilities and interests and their relations to realized occupation. *Personality and Individual Differences*, *60*, S46.

Mandleberg, I. A., & Brooks, D. N. (1975). Cognitive recovery after severe head injury. 1. Serial testing on the Wechsler Adult Intelligence Scale. *Journal of Neurology, Neurosurgery, and Psychiatry*, *38*(11), 1121–1126.

Marioni, R. E., Davies, G., Hayward, C., Liewald, D., Kerr, S. M., Campbell, A., ... Deary, I. J. (2014). Molecular genetic contributions to socioeconomic status and intelligence. *Intelligence*, *44*(100), 26–32.

Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit in structural equation models. *Contemporary psychometrics: A festschrift for Roderick P. McDonald*, 275–340.

Marwaha, S., & Johnson, S. (2004). Schizophrenia and employment. *Social Psychiatry and Psychiatric Epidemiology*, *39*(5), 337–349.

Matheny, A. P. (1979). Appraisal of parental bias in twin studies. Ascribed zygosity and

IQ differences in twins. *Acta geneticae medicae et gemellologiae: twin research*, *28*(02), 155–160.

Matthews, K. A., Gallo, L. C., & Taylor, S. E. (2010). Are psychosocial factors mediators of socioeconomic status and health connections? A progress report and blueprint for the future. *Annals of the New York Academy of Sciences*, *1186*, 146–173.

McEwen, B. S., & Seeman, T. E. (1999). Protective and damaging effects of mediators of stress: Elaborating and testing the concepts of allostasis and allostatic load. *Annals of the New York Academy of Sciences*, *896*(1), 30–47.

McEwen, B. S., & Stellar, E. (1993). Stress and the individual: Mechanisms leading to disease. *Archives of Internal Medicine*, *153*(18), 2093–2101.

McGuire, S., Dunn, J., & Plomin, R. (1995). Maternal differential treatment of siblings and children's behavioral problems: A longitudinal study. *Development and Psychopathology*, *7*(03), 515–528.

Meyler, D., Stimpson, J. P., & Peek, M. K. (2007). Health concordance within couples: A systematic review. *Social Science and Medicine*, *64*(11), 2297–2310.

Monden, C. W. S. (2010). Do measured and unmeasured family factors bias the association between education and self-assessed health? *Social Indicators Research*, *98*(2), 321–336.

Muthén, L. K., & Muthén, B. O. (2014). *Mplus: Statistical analyses with latent variables. User's guide version 7.4.*

Myrianthopoulos, N. C., & French, K. S. (1968). An application of the U.S. Bureau of the Census socioeconomic index to a large, diversified patient population. *Social Science & Medicine*, *2*(3), 283–299.

National Institute of Mental Health's Genetics Workgroup. (1998). *Genetics and mental disorders: Report of the National Institute of Mental Health's Genetics Workgroup* (Tech. Rep.). National Institute of Mental Health. Retrieved from `https://repository.library.georgetown.edu/bitstream/handle/`

10822/532360/nimhgenetics.pdf?sequence=1&isAllowed=y

National Institutes of Health. (2000). *Consensus Development Conference Statement Phenylketonuria: Screening and Management* (Tech. Rep.). National Institutes of Health. Retrieved from https://www.nichd.nih.gov/publications/pubs/pku/pages/sub3.aspx

Neale, M. C., & Maes, H. H. (2004). *Methodology for Genetic Studies of Twins and Families*. Dordrecht, The Netherlands: Kluwer Academic Publishers B.V.

Nelson, C. A., Zeanah, C. H., Fox, N. A., Marshall, P. J., Smyke, A. T., & Guthrie, D. (2007). Cognitive recovery in socially deprived young children: The Bucharest Early Intervention Project. *Science*, *318*(5858), 1937–1940.

Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, *58*(2), 367–408.

Oakes, J. M., & Rossi, P. H. (2003). The measurement of SES in health research: Current practice and steps toward a new approach. *Social Science & Medicine*, *56*(4), 769–784.

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science*, *26*(6), 784–793.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on 'environmental' measures. *Behavioral and Brain Sciences*, *14*, 373.

Plomin, R., DeFries, J. C., & McClearn, G. E. (1990). *Behavior genetics: A primer*. New York: WH Freeman and Company.

Prescott, C. A. (2004). Using the Mplus computer program to estimate models for continuous and categorical data from twins. *Behavior Genetics*, *34*(1), 17–40.

Ree, M. J., & Earles, J. A. (1990). *Relationships of general ability, specific ability, and job category for predicting training performance* (Tech. Rep.). Brooks Air Force Base,

Texas: Manpower and Personnel Division Air Force Human Resources Laboratory.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science*, *1*(3), 86–89.

Reeve, E. C. R. (1961). A note on non-random mating in progeny tests. *Genetical Research*, *2*(02), 195–203.

Repetti, R. L., Taylor, S. E., & Seeman, T. E. (2002). Risky families: Family social environments and the mental and physical health of offspring. *Psychological Bulletin*, *128*(2), 330.

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 369–381). Guilford Press.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345.

Rochat, L., Beni, C., Billieux, J., Azouvi, P., Annoni, J.-M., & Van der Linden, M. (2010). Assessment of impulsivity after moderate to severe traumatic brain injury. *Neuropsychological Rehabilitation*, *20*(5), 778–797.

Rodgers, B., & Mann, S. L. (1993). Re-thinking the analysis of intergenerational social mobility: A comment on John W. Fox's "Social Class, Mental Illness, and Social Mobility". *Journal of Health and Social Behavior*, 165–172.

Rodgers, J. L., Beasley, W. H., Bard, D. E., Meredith, K. M., Hunter, M., Johnson, A. B., ... Rowe, D. C. (2016). The NLSYKinship Links: Using the NLSY79 and NLSY-Children data to conduct genetically-informed and family-oriented research. *Behavior Genetics*, *46*(4), 538–551.

Rodgers, J. L., Rowe, D. C., & May, K. O. (1994). DF analysis of NLSY IQ/achievement data: Nonshared environmental influences. *Intelligence*, *19*(2), 157–177.

Rowe, D. C., & Rodgers, J. L. (1997). Poverty and behavior: Are environmental measures nature and nurture ? *Developmental Review*, *375*(17), 358–375.

Rowe, D. C., Rodgers, J. L., & Meseck-Bushey, S. (1992). Sibling delinquency and the family environment: Shared and unshared influences. *Child Development*, *63*(1), 59–67.

Rowe, D. C., Vesterdal, W. J., & Rodgers, J. L. (1998). Herrnstein's syllogism: Genetic and shared environmental influences on IQ, education, and income. *Intelligence*, *26*(4), 405–423.

Rowling, J. K. (1997). *Harry Potter and the Philosopher's Stone*. Bloomsbury.

Saegert, S. C., Adler, N. E., Bullock, H. E., Cauce, A. M., Liu, W. M., & Wyche, K. F. (2006). *Report of the APA Task Force on Socioeconomic Status* (Tech. Rep.). American Psychological Association. Retrieved from `https://www.apa.org/pi/ses/resources/publications/task-force-2006.pdf`

Sanger-Katz, M. (2015, mar). *Income inequality: It's also bad for your health.* Retrieved from `http://www.nytimes.com/2015/03/31/upshot/income-inequality-its-also-bad-for-your-health.html`

Schmidt, F. L., & Hunter, J. E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173.

Siddiqi, A., & Hertzman, C. (2007). Towards an epidemiological understanding of the effects of long-term institutional changes on population health: a case study of Canada versus the USA. *Social Science & Medicine*, *64*(3), 589–603.

Silventoinen, K., Kaprio, J., Lahelma, E., & Koskenvuo, M. (2000). Relative effect of genetic and environmental factors on body height: Differences across birth cohorts among Finnish men and women. *American Journal of Public Health*, *90*(4), 627–630.

Singh-Manoux, A., Marmot, M. G., & Adler, N. E. (2005). Does subjective social status

predict health and change in health status better than objective status? *Psychosomatic Medicine*, *67*(6), 855–861.

Smith, J. P. (1999). Healthy bodies and thick wallets: The dual relation between health and economic status. *Journal of Economic Perspectives*, *13*(2), 145–166.

Søndergaard, G., Mortensen, L. H., Andersen, A.-M. N., Andersen, P. K., Dalton, S. O., & Osler, M. (2013). Social inequality in breast, lung and colorectal cancers: A sibling approach. *BMJ*, *3*(3).

Søndergaard, G., Mortensen, L. H., Nybo Andersen, A.-M., Andersen, P. K., Dalton, S. O., Madsen, M., & Osler, M. (2012). Does Shared Family Background Influence the Impact of Educational Differences on Early Mortality? *American Journal of Epidemiology*, *176*(8), 675–683.

Steenland, K., Henley, J., & Thun, M. (2002). All-cause and cause-specific death rates by educational status for two million people in two American Cancer Society cohorts, 1959–1996. *American Journal of Epidemiology*, *156*(1), 11–21.

Steenstrup, T., Pedersen, O. B., Hjelmborg, J., Skytthe, A., & Kyvik, K. O. (2013). Heritability of health-related quality of life: SF-12 summary scores in a population-based nationwide twin cohort. *Twin Research and Human Genetics*, *16*(3), 670–8.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, *35*(5), 401–426.

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *American Journal of Psychiatry*, *157*(10), 1552–1562.

Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., … Plomin, R. (2014). Genetic influence on family socioeconomic status and children's intelligence. *Intelligence*, *42*(100), 83–88.

Tsuang, M. (2000). Schizophrenia: Genes and environment. *Biological Psychiatry*, *47*(3), 210–220.

Turkheimer, E. (2000). Three Laws of Behavior Genetics and What They Mean. *Current Directions in Psychological Science*, *9*(5), 160–164.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B. M., & Gottesman, I. I. (2003). Socioeconomic status modified heritability of IQ in young children. *Psychological Science*, *14*(6), 623–628.

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin*, *126*(1), 78–108.

Vernon, P. A., Jang, K. L., Harris, J. A., & McCarthy, J. M. (1997). Environmental predictors of personality differences: A twin and sibling study. *Journal of Personality and Social Psychology*, *72*(1), 177–183.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era – concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255–266.

Wagstaff, A., & Doorslaer, E. V. (2000). Income inequality and health: What does the literature tell us? *Annual Review of Public Health*.

Ward, P., & Burich, M. C. (1978). *Definitions for identifying 4270 samples from Screener data* (Tech. Rep.). National Opinion Research Center.

Ware, J. E., Kosinski, M., & Keller, S. D. (1995). *How to score the SF12 Physical and Mental Health Summary Scales*. Boston: The Health Institute, New England Medical Center.

Watkins, M. P., & Meredith, W. (1981). Spouse similarity in newlyweds with respect to specific cognitive abilities, socioeconomic status, and education. *Behavior Genetics*, *11*(1), 1–21.

Woods, S. A., & Hampson, S. E. (2010). Predicting adult occupational environments from gender and childhood personality traits. *Journal of Applied Psychology*, *95*(6), 1045–1057.

Wright, S. (1923). The theory of path coefficients: A reply to Niles's criticism. *Genetics*, *8*(3), 239–255.

Wright, S. (1934). The Method of Path Coefficients. *Annals of Mathematical Statistics*, *5*(3), 161–215.

Yoshikawa, H., Aber, J. L., & Beardslee, W. R. (2012). The effects of poverty on the mental, emotional, and behavioral health of children and youth: implications for prevention. *American Psychologist*, *67*(4), 272.

Appendix A

Appendix: Kinship Correlations

Table A.1: Kinship Correlations and Sample Sizes (indivs), for MZ Twins, Full-Siblings, Half-Siblings, and Cousins ($*p < .05$)

	SES_1	SES_2	PCS_1	PCS_2	MCS_1	MCS_2
SES_1	7168	6300	7107	6250	7107	6250
SES_2	0.382*	7168	6250	7107	6250	7107
PCS_1	0.197*	0.128*	7107	6202	7107	6202
PCS_2	0.128*	0.197*	0.089*	7107	6202	7107
MCS_1	0.070*	0.009	0.014	0.029*	7107	6202
MCS_2	0.009	0.070*	0.029*	0.014	0.093*	7107

Table A.2: Kinship Correlations and Sample Sizes (indivs), for MZ Twins (r= 1; $*p < .05$)

	SES_1	SES_2	PCS_1	PCS_2	MCS_1	MCS_2
SES_1	20	18	20	18	20	18
SES_2	0.49*	20	18	20	18	20
PCS_1	0.38	0.27	20	18	20	18
PCS_2	0.27	0.38	0.49*	20	18	20
MCS_1	-0.50*	-0.57*	-0.61*	-0.48*	20	18
MCS_2	-0.57*	-0.50*	-0.48*	-0.61*	0.63*	20

Table A.3: Kinship Correlations and Sample Sizes (indivs), for Full-Siblings (r= .5; $*p < .05$)

	SES_1	SES_2	PCS_1	PCS_2	MCS_1	MCS_2
SES_1	6512	5730	6461	5688	6461	5688
SES_2	0.391*	6512	5688	6461	5688	6461
PCS_1	0.202*	0.133*	6461	5646	6461	5646
PCS_2	0.133*	0.202*	0.092*	6461	5646	6461
MCS_1	0.065*	0.012	0.010	0.036*	6461	5646
MCS_2	0.012	0.065*	0.036*	0.010	0.083*	6461

Table A.4: Kinship Correlations and Sample Sizes (indivs), for Half-Siblings (r= .25)

	SES_1	SES_2	PCS_1	PCS_2	MCS_1	MCS_2
SES_1	477	404	470	399	470	399
SES_2	0.217∗	477	399	470	399	470
PCS_1	0.175∗	0.053	470	396	470	396
PCS_2	0.053	0.175∗	0.042	470	396	470
MCS_1	0.148∗	0.003	0.088	-0.052	470	396
MCS_2	0.003	0.148∗	-0.052	0.088	0.140∗	470

Table A.5: Kinship Correlations and Sample Sizes (indivs), for Cousins (r= .125; $*p < .05$)

	SES_1	SES_2	PCS_1	PCS_2	MCS_1	MCS_2
SES_1	159	148	156	145	156	145
SES_2	0.164∗	159	145	156	145	156
PCS_1	-0.020	0.097	156	142	156	142
PCS_2	0.097	-0.020	-0.032	156	142	156
MCS_1	0.102	-0.024	0.076	0.054	156	142
MCS_2	-0.024	0.102	0.054	0.076	0.195∗	156

Appendix B

Appendix: National Probability Subsample Replication

In the spirit of reproducability, I replicated the final model using the National Probability Subsample, which consists of the cross-sectional household probability subsample described earlier. Model fit statistics were generally excellent. Although the χ^2 test of model fit ($p < .001$) was significant, RMSEA (0.04, 90% CI [0.03, 0.05]) indicated close fit. SRMR was low (0.05) and below the .07 threshold.

After adjusting for minority status and sex, the estimated parameters for SES were a^2 = 33.61% ($p < .001$; 95% CI [.05, .86]), c^2 = 19.79% ($p < .001$; 95% CI [.05, .45]), e^2 = 46.61% ($p < .001$; 95% CI [.28, .70]). The estimated parameters for PCS were a^2 = 7.96% ($p < .001$; 95% CI [.02, .18]) and e^2 = 92% ($p < .001$; 95% CI [.84, 1]). The estimated parameters for MCS were c^2 = 6.67% ($p < .001$; 95% CI [3, 12]) and e^2 = 93% ($p < .001$; 95% CI [.89, .98]). Standardized model parameter estimates are displayed in Figure B.1 with standard errors in parentheses. Unadjusted a^2, c^2, and e^2 estimates can be derived by squaring the factor loadings.

The estimated covariance between the genetic components of SES and PCS was 1.0 ($p < .001$; 95% CI [0.61, 1]). I calculated the genetic correlation ($\hat{r}_a$) between SES and PCS with path tracing rules (Wright, 1923, 1934). I found that the genetic correlation ($\hat{r}_a$) was 0.16, 95% CI [0.12, 0.18]. Consequently, the proportion of the latent phenotypic correlation ($\hat{r}$ = 0.18) attributable to genetics (the heritability) was 0.86, 95% CI [0.67, 0.97]. As a more conservative test, I substituted the latent phenotypic correlation with the observed correlation between SES and PCS in the national probability subsample (r = 0.21); the proportion explained by genetics was 0.74, 95% CI [0.57, 0.83]. Both proportions indicated that a large component of the gradient was heritable, and the remaining proportion was explained by the non-shared environment.

The covariance between the shared-environmental components of SES and MCS was 0.76 ($p = 0.01$; 95% CI [0.22, 1]). I found that the shared-environmental correlation ($\hat{r}_C$) was 0.08, 95% CI [0.05, 0.1]. Consequently, the proportion of the latent phenotypic correlation ($\hat{r}$ = 0.07) explained by the shared-environmental correlation was 1, 95% CI [0.72, 1]. As a more conservative test, I substituted the latent phenotypic correlation with the observed correlation between SES and MCS in the national probability subsample (r = 0.09); this proportion explained by the shared environment was 0.94, 95% CI [0.58, 1].

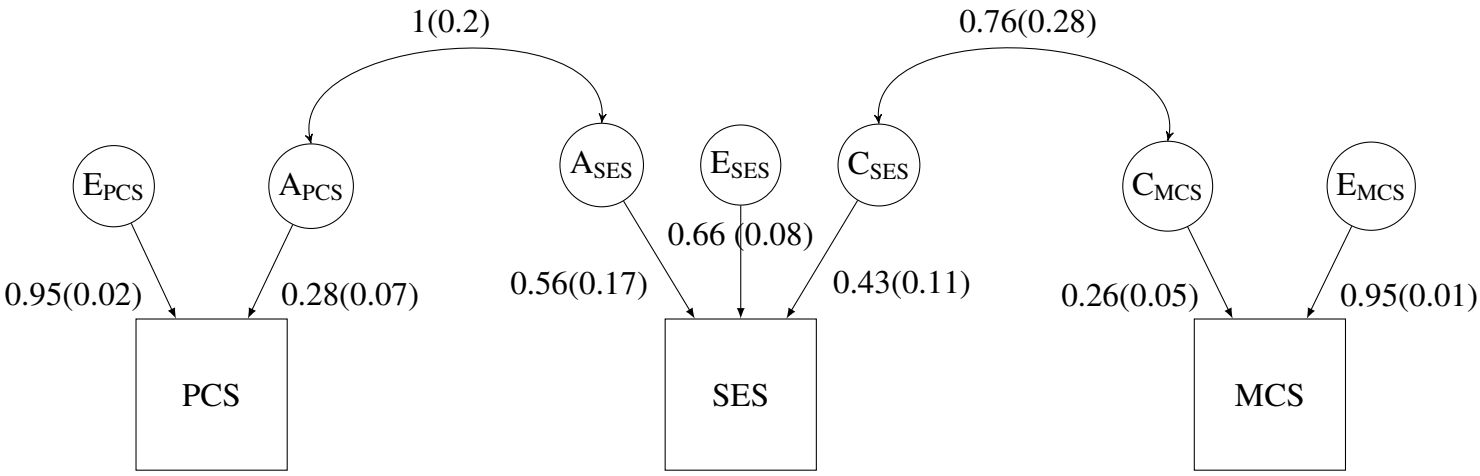

Figure B.1: Correlated Factors of SES, Physical Health, & Mental Health at Age 40 (without MZ twins in the National Probability Subsample)

Appendix C

Appendix: Annotated Mplus Syntax

C.1 Univariate Model

```
USEVAR are
   !Outcomes
         H_1
         H_2

   !Covariates
      Male_1
      Male_2
      RACE_1

   !Grouping Variable on level of relatedness
         g;

   !specify the kinship groups
   grouping=g( 1=R1
            !2=R.75
            3=R.5
            !4=R.375
            5=R.25
            6=R.125
            !7=R.0625
            !8=R0
```

```
      );


DEFINE: ! Link level of relatedness to group
   if (R==1) then g=1; !MZ Twins
   !if (R==.75) then g=2; !Ambig Twins
   if (R==.5) then g=3; !Full Sibs (Includes DZ twins)
   !if (R==.375) then g=4; !Ambig Sibs
   if (R==.25) then g=5; !Half Sibs
   if (R==.125) then g=6; !Full Cousins
   !if (R==.0625) then g=7; !Ambig Cousins
   !if (R==0) then g=8; !Unrelated Housemates

! Define Outcome of Interest
   H_1= MCS_1;
   H_2= MCS_2;

! Restrict Analysis to Full Cousins(6), Half Siblings(5),
!! Full Siblings(3), MZ twins (1); Needed for Smart Missing
   Data
   useobservations = (g eq 6 OR g eq 5 OR g eq 3 OR g eq
      1);!


ANALYSIS:
   MODEL=NOCOVARIANCES;
   ITERATIONS = 25000;
```

```
MODEL: !set up values for all groups
   H_1@0; H_2@0; !fix residual variances to zero
   A1 BY H_1*.5 (a1); A2 BY H_2*.5(a1); !additive genetic
      loadings
   C1 BY H_1*.5 (12); C2 BY H_2*.5 (12); !common envt
      loadings
   E1 BY H_1*.8 (13); E2 BY H_2*.8 (13); !specific envt
      loadings

   !Control for Covariates
   H_1 on MALE_1 (sex); H_2 on MALE_2 (sex);
   H_1 on RACE_1 (race); H_2 on RACE_1 (race);

   !fix latent variable means=0
   [A1@0 A2@0];
   [C1@0 C2@0];
   [E1@0 E2@0];

!fix latent variable vars=1
   A1@1 A2@1 ;
   C1@1 C2@1 ;
   E1@1 E2@1;

!latent variable covs
```

```
   C1 WITH C2@1; !Shared Environment
   E1 WITH E2@0; !Non-Shared Environment

   A1 WITH C1-C2@0; A2 WITH C1-C2@0;
   A1 WITH E1-E2@0; A2 WITH E1-E2@0;
   C1 WITH E1-E2@0; C2 WITH E1-E2@0;


! Set Genetic Covariance based on kin group
MODEL R1:
   [H_1 H_2] (m); !means can vary by group
   A1 WITH A2@1;

MODEL R.5:
   [H_1 H_2] (m5); !means can vary by group
   A1 WITH A2@0.5;

MODEL R.25:
   [H_1 H_2] (m25); !means can vary by group
   A1 WITH A2@0.25;

MODEL R.125:
   [H_1 H_2] (m125); !means can vary by group
   A1 WITH A2@0.125;


    OUTPUT: STDyx TECH1 TECH3 TECH4 CINTERVAL;
```

C.2 Bivariate Correlated Factors Model

```
USEVAR are
   !Outcomes
        H_1
        H_2
        SES_1
        SES_2

   !Covariates
      Male_1
      Male_2
      RACE_1

   !Grouping Variable on level of relatedness
        g;

!specify the kinship groups
   grouping=g( 1=R1
            !2=R.75
            3=R.5
            !4=R.375
            5=R.25
            6=R.125
            !7=R.0625
            !8=R0
      );
```

```
DEFINE: ! Link level of relatedness to group
     if (R==1) then g=1; !MZ Twins
     !if (R==.75) then g=2; !Ambig Twins
     if (R==.5) then g=3; !Full Sibs (Includes DZ twins)
     !if (R==.375) then g=4; !Ambig Sibs
     if (R==.25) then g=5; !Half Sibs
     if (R==.125) then g=6; !Full Cousins
     !if (R==.0625) then g=7; !Ambig Cousins
     !if (R==0) then g=8; !Unrelated Housemates


   !Define Outcome of Interest
      !Health
        H_1= MCS_1;
        H_2= MCS_2;

      !SES
        SES_1= S_40_1;
        SES_2= S_40_2;


! Restrict Analysis to Full Cousins(6), Half Siblings(5),
!! Full Siblings(3), MZ twins (1); Needed for Smart Missing
   Data
   useobservations = (g eq 6 OR g eq 5 OR g eq 3 OR g eq
```

```
      1);!


ANALYSIS:
   MODEL=NOCOVARIANCES;
   ITERATIONS = 25000;


MODEL:

! Create Latent Factors
   !Genetic Effect for SES
      aSES1 BY SES_1;
      aSES2 BY SES_2;

   !Shared Environment Effect for SES
      cSES1 BY SES_1;
      cSES2 BY SES_2;

   !Non-Shared Environment Effect for SES
      eSES1 BY SES_1;
      eSES2 BY SES_2;

   !Genetic Effect for Health
      aH1 BY H_1;
      aH2 BY H_2;
```

```
   !Shared Environment Effect for Health
      cH1 BY H_1;
      cH2 BY H_2;

   !Non-Shared Environment Effect for Health
      eH1 BY H_1;
      eH2 BY H_2;

!Control for Covariates
   !Male
      H_1 on MALE_1 (sexh); H_2 on MALE_2 (sexh);
      SES_1 on MALE_1 (sexs); SES_2 on MALE_2 (sexs);

   !Minority Status
      H_1 on RACE_1 (raceh); H_2 on RACE_1 (raceh);
      SES_1 on RACE_1 (races); SES_2 on RACE_1 (races);


!residual variances
     SES_1@0 SES_2@0; !residual variances on SES to zero
     H_1@0 H_2@0; !residual variances on H to zero



!BIOMETRIC COMPONENTS
     !!SES
         aSES1 BY SES_1*.2 (a1); aSES2 BY SES_2*.2 (a1);
```

```
        cSES1 BY SES_1*.2 (c1); cSES2 BY SES_2*.2 (c1);
        eSES1 BY SES_1*.8 (e1); eSES2 BY SES_2*.8 (e1);

      [aSES1@0 cSES1@0 eSES1@0 aSES2@0 cSES2@0 eSES2@0];
         !fix latent means to 0
      aSES1@1 cSES1@1 eSES1@1 aSES2@1 cSES2@1 eSES2@1; !fix
         latent variances to 1

    !!Health
      aH1 BY H_1*.1 (a2); aH2 BY H_2*.1 (a2);
      cH1 BY H_1*.1 (c2); cH2 BY H_2*.1 (c2);
      eH1 BY H_1*.8 (e2); eH2 BY H_2*.8 (e2);

        [aH1@0 cH1@0 aH2@0 cH2@0 eH1@0 eH2@0]; !fix latent
           means to 0

      aH1@1 cH1@1 aH2@1 cH2@1 eH1@1 eH2@1; !fix latent
         variances to 1


!CORRELATIONS AMONG BIOMETRIC COMPONENTS
    !!SES
      !Cross-Loadings
        aSES1 WITH cSES1-eSES2@0;
        aSES2 WITH cSES1-eSES2@0;
        cSES1 WITH eSES1-eSES2@0;
        cSES2 WITH eSES1-eSES2@0;
```

```
  !Environment
    cSES1 WITH cSES2@1;!Shared Environment
    eSES1 WITH eSES2@0;!Non-Shared Environment

!!Health
  !Cross-Loadings
    aH1 WITH cH1-cH2@0;
    aH2 WITH cH1-cH2@0;
    cH1 WITH eH1-eH2@0;
    cH2 WITH eH1-eH2@0;

  !Environment
     cH1 WITH cH2@1; !Shared Environment
     eH1 WITH eH2@0; !Non-Shared Environment


!!SES with Health
  !Cross-Loadings
     aSES1-aSES2 WITH cH1-cH2@0;
     cSES1 WITH aH2@0; cSES2 WITH aH1@0;
     cSES1-cSES2 WITH aH1-aH2@0;
     eSES1-eSES2 WITH aH1-aH2@0;


  !A Corr
     !aSES1-aSES2 WITH aH1-aH2@0; !No Cor
```

```
      aSES1 WITH aH1 (a1a2); aSES2 WITH aH2 (a1a2);

   !C Corr
      !cSES1-cSES2 WITH cH1-cH2@0;!No Cor
      cSES1 WITH cH1 (c1c2); cSES2 WITH cH2 (c1c2);

   !E Corr
      eSES1-eSES2 WITH eH1-eH2@0; !No Cor

      !eSES1 WITH eH1*.5 (e1e2); eSES2 WITH eH2*.5
         (e1e2);


! Set Genetic Covariance based on kin group
MODEL R1:
     [SES_1 SES_2] (m1); !means can vary by group
     [H_1 H_2] (m2); !means can vary by group

   aSES1 WITH aSES2@1; aH1 WITH aH2@1;

MODEL R.5:
     [SES_1 SES_2] (m1r5); !means can vary by group
     [H_1 H_2] (m2r5); !means can vary by group

   aSES1 WITH aSES2@0.5; aH1 WITH aH2@0.5;

MODEL R.25:
```

```
   [SES_1 SES_2] (m1r25); !means can vary by group
   [H_1 H_2] (m2r25); !means can vary by group

   aSES1 WITH aSES2@0.25; aH1 WITH aH2@0.25;

MODEL R.125:
     [H_1 H_2] (m2r125); !means can vary by group
     [SES_1 SES_2] (m1r125); !means can vary by group

   aSES1 WITH aSES2@0.125; aH1 WITH aH2@0.125;

! total correlations; allows us to get a confidence interval
NEW (rAsh); !Correlation Attributable to Genetics, using
   path tracing
   rAsh=a2*a1*a1a2;
   rAsh<1;

NEW (rCsh); !Correlation Attributable to Shared
   Environment, using path tracing
   rCsh=c2*c1*c1c2;
   rCsh<1;

OUTPUT: STDyx TECH1 TECH3 TECH4 CINTERVAL;
```