2021

Genetics and Alcohol Interventions in Youth

Zoe E. Neale
Virginia Commonwealth University

GENETICS AND ALCOHOL INTERVENTIONS IN YOUTH

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

by

ZOË E. NEALE
Master of Science, Clinical Psychology,
Virginia Commonwealth University, 2016

Director: Danielle M. Dick, Ph.D.
Commonwealth Professor
Departments of Psychology and Human & Molecular Genetics

Virginia Commonwealth University
Richmond, Virginia
April 2021

Acknowledgement

The completion of this dissertation would not have been possible without a tremendous amount of support from a number of individuals throughout my training. First, I would like to thank my graduate mentor and dissertation chair, Danielle Dick. I am forever grateful to you for providing me with my first job in research and supporting me in every pursuit that followed. I have learned so much from you and I am thankful that I had the chance to grow as a researcher, clinician, and professional under your mentorship. I would also like to extend a huge thank you to the members of my dissertation committee, Fazil Aliev, Nathan Gillespie, Josh Langberg and Dace Svikis for their helpful feedback, patience, and support throughout the process of completing this dissertation project. Many thanks as well to the Spit for Science and Project Alliance participants who made this research possible, and to the NIAAA for financially supporting my training through a pre-doctoral fellowship award (F31AA027130).

I am also immensely grateful to a number of faculty, research staff, postdocs, and graduate students who contributed to my learning and the completion of this project. Peter Barr and Nate Thomas are *R* coding masterminds who kindly and patiently problem-solved countless error codes as I tussled my way through learning *R*. Many thanks to Sally Kuo for orienting me to the PAL project and to Jessie Greenlee for learning growth curve modeling in *R* along with me. I am also so thankful to my fellow EDGE Lab graduate students, Rebecca Smith and Morgan Driver, for their kind and collaborative encouragement and the unending stream of pet-related content. To the rest of the EDGE Lab (and alumni) – thank you for providing a stimulating and inspiring place for my training.

Special thanks go to three people who have been alongside me since I began with the Spit for Science project nearly 10 years ago! First, thank you to Jessica Salvatore for your friendship, informal mentorship, efficiency, sense of humor, and sanity-checks. Second, I am so grateful to Kimberly Pedersen for her outstanding job as the Spit for Science project coordinator, but more importantly for the treasured friendship that continues to grow even as our lives evolve. Lastly, thank you to Megan Cooke for being my "grad school guide" in every sense. I truly cannot imagine getting through this without your gracious support and encouragement, day and night.

Finally, I would not be here without the support of my family and friends. Thank you to my mom, Lynne, for loving me and inspiring/expecting me to work to my fullest potential. To my dad, Michael – thank you for being my dad above all else and tempering your scientific expertise with an endless stream of dad joke. Thank you to Hermine for loving me like one of your own. A thousand thanks to each of my five siblings – you have been my cheerleaders, my comfort, my inspiration, and my best friends. Thank you to my amazing cohort who have cheered me along in every success and failure. You all are brilliant, and I am honored to call you my friends and colleagues. Finally, thank you to my spouse, Stephen. I am grateful for your steadiness, humor, and affable nature that makes life with you so much fun – and thank you for bringing Henson into our lives. I had second thoughts about getting a dog during my first year of graduate school, but he turned out to be the greatest gift I have ever received.

TABLE OF CONTENTS

List of Tables

List of Figures

Abstract

GENETICS AND ALCOHOL INTERVENTIONS IN YOUTH

Zoe Elizabeth Neale, M.S.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2021

Director: Danielle Dick, Ph.D.
Commonwealth Professor
Departments of Psychology, and Human & Molecular Genetics

Alcohol is the most commonly used substance among youth, and risky alcohol use is associated with harmful consequences such as accidents, academic consequences, and physical and emotional health problems. Alcohol use disorders are approximately 50% heritable, yet most efforts to prevent and intervene upon youth alcohol use focus only on environmental factors. Furthermore, current prevention and intervention programs tend to have modest effects and are not uniformly effective for all individuals. Gene-by-intervention (GxI) studies offer an opportunity to expand current understanding of interventions by examining whether underlying genetic risk may contribute to differential program effects. Much of the current GxI literature on alcohol and substance use outcomes is limited in scope due to reliance on candidate gene methods, focus on youth prevention samples, and lack of understanding of mediators or mechanisms through which genetics may contribute to differential intervention effects. To

address these gaps in the research, the present study aimed to 1) determine if polygenic risk for externalizing problems moderated the effectiveness of an alcohol intervention, and 2) to examine whether peer deviance and drinking motives mediated intervention effects for those at greater genetic risk. To explore whether findings were consistent across different types of interventions and developmental timing, the present study used data from two samples: a college prevention intervention program conducted with a genetically informed sample (Spit for Science; S4S), and a middle school-based prevention program targeting adolescent problem behavior with longitudinal follow up and genetic data (Project Alliance; PAL). In the S4S sample, multilevel growth curve analyses showed no evidence of interactions between polygenic risk for externalizing problems (EXT PRS) and the intervention on alcohol consumption and alcohol use disorder (AUD) symptoms across time; however, there was evidence of short-term GxI effects on AUD symptoms in post-hoc analyses. Individuals with lower EXT PRS in the intervention condition reported significantly greater reduction in AUD symptoms than individuals with higher EXT PRS and control. In the PAL sample, we observed no significant GxI effects on trajectories of alcohol consumption across time or AD symptoms. There was also no evidence of mediation via peer deviance or drinking motives in either sample. Due to limitations of statistical power, the lack of replication across studies, and the possibility of measurement error, the significant GxI effects in S4S are viewed conservatively. Larger, more well-powered studies in diverse samples are needed to explore the presence or absence of very small (f^2 = .005) GxI effects and determine whether genetics can be harnessed to develop novel interventions to better address alcohol-related problems. Opportunities for attaining larger, more diverse samples are discussed.

Introduction to the Study

Alcohol is one of the most easily accessible and commonly used substances in the United States (Schulenberg et al., 2017). Despite known associations between alcohol consumption and costly health consequences, alcohol use and misuse remains prevalent (Griswold et al., 2018). Alcohol misuse is particularly concerning in emerging adulthood, during which peak use of alcohol occurs and heavy drinking is often normalized (Grant et al., 2017). Emerging adulthood represents a period of substantial transition and increasing independence, as well as a confluence of psychosocial and biological factors that predispose individuals in this age group to risky behavior (Arnett, 2000; Sussman & Arnett, 2014). Among individuals 18-25 years old, 57% report drinking alcohol in the past month and more than a third (38.4%) endorsed past month binge drinking (Schulenberg et al., 2018). Most college and non-college attending young adults report having been drunk, and an estimated 10.7% of individuals age 18-25 met criteria for an alcohol use disorder (AUD) in 2016 (Schulenberg et al., 2017; Substance Abuse and Mental Health Services Administration, 2017). Rates of AUDs are highest among emerging adults, with approximately 7% of 18-29 year-olds meeting criteria for a severe AUD (Grant et al., 2004, 2015). Risky alcohol use is associated with immediate and long-term consequences to the individual and the larger community, including costly personal effects (health, emotional wellbeing, relationship problems, and academic success) and community problems (accidents, assault, crime) (Arria et al., 2013; Hingson et al., 2009; Wechsler et al., 2002; White & Hingson, 2013). Alcohol use behaviors that develop in young adulthood set the stage for patterns that last throughout adulthood (Gotham et al., 1997; Jackson et al., 2001). Thus, adolescence and young

adulthood are important developmental periods in which to focus prevention and intervention efforts.

Prevention and Intervention Methods

Existing methods for prevention and intervention. Evidence-based methods for preventing AUDs and alcohol-related harms typically involve school-based, family-based, or multi-component interventions for adolescents and brief, motivational interventions with normative feedback for young adults (Smit et al., 2008; Spoth et al., 2008; Stigler et al., 2011). Schools provide an efficient means of accessing large numbers of youth through universal intervention programs, as well as targeted programs for high-risk students. School-based programs for youth are most commonly delivered during middle school, capitalizing on the period during which many youth have their first experiences with alcohol and other substance use. The most effective programs are delivered across multiple years, target multiple behaviors, and incorporate social norms surrounding peer substance use, skill-building, interactive activities (e.g., role play), peer leaders, and culturally and developmentally appropriate content (MacArthur et al., 2018; Stigler et al., 2011). Fewer programs are delivered in elementary school, but those that do typically focus on the prevention of externalizing behaviors and other risk factors for future substance misuse (Spoth et al., 2008; Stigler et al., 2011). Despite the prevalence of alcohol use among high school students, there are also few programs that address alcohol misuse among high schoolers (Spoth et al., 2008).

Family-based interventions work with either parental figures exclusively or parental figures and children together to enhance family management, improve bonding, build prosocial skills, increase monitoring and communication, and minimize aggressive/problem behaviors (Griffin & Botvin, 2010; Lochman & van den Steenhoven, 2002). Parenting skills and family

bonding appear to be important active ingredients in family-based interventions, but effectiveness is often limited by challenges engaging parental figures of the most high-risk youth (Griffin & Botvin, 2010). Two recent systematic reviews of family-based prevention programs observed little to no effect on frequency, quantity, or prevalence of alcohol use in youth, with small effects emerging only for universal programs and those that target racial/ethnic minority groups (Gilligan et al., 2019; MacArthur et al., 2018).

Multi-component programs deliver intervention content across multiple settings, combining elements of school and family-based interventions to broaden the impact of programming. One such program, named FAST (Families and Schools Together) Track, targets disruptive/aggressive students to prevent conduct problems and externalizing behavior into adulthood (Conduct Problems Prevention Research Group, 1992). FAST Track implements a comprehensive program comprised of universal school curriculum, tutoring, home visits, group skills training, mentoring and individual services for children in FAST Track. Systematic review of multi-component interventions suggest small effects that decrease the burden of alcohol use problems over time (Foxcroft & Tsertsvadze, 2011).

Among college students brief motivational interventions (BMIs), personalized normative feedback (PNF), and skills training are the most commonly used empirically-supported approaches for preventing risky alcohol use (Cronce & Larimer, 2011; Larimer & Cronce, 2007). BMIs typically involve an assessment of alcohol use behaviors, beliefs about risks and benefits of alcohol, and goals for college, followed by the delivery of in-person or computerized feedback in a motivational interviewing (MI) style. The feedback aims to elucidate discrepancies between drinking behaviors and student goals for college, and to increase motivation to change using a non-judgmental, client-centered tone. PNF involves collecting information about an individual's

alcohol use behaviors and providing feedback about how one's own drinking compares to others in their population and healthy standards for alcohol use (Lewis & Neighbors, 2006). PNF operates by correcting misperceptions (usually overestimates) of peer rates of drinking, which results in small but consistent reductions in alcohol use behaviors in college students (Carey et al., 2007; Larimer & Cronce, 2007; Lewis & Neighbors, 2006). Skills training employs strategies from cognitive behavioral therapy to help individuals identify personal drinking cues, build drink refusal skills, set limits, and manage triggers such as stress and depression (Baer et al., 1992; Kivlahan et al., 1990). Across reviews, brief alcohol interventions demonstrate small effects on alcohol consumption in the short-term (less than 1 year), but most effects dissipate over time (Cronce & Larimer, 2011; Huh et al., 2015; Larimer & Cronce, 2007; Samson & Tanner-Smith, 2015).

Limitations of prevention and intervention. Substantial resources have been dedicated to the design, evaluation, and implementation of alcohol prevention programs, yet small effect sizes persist in both adolescent and young adult samples (Huh et al., 2015; MacArthur et al., 2018; Sandler et al., 2014; Strøm et al., 2014). Alcohol prevention and intervention programs also are not equally effective for all individuals, but the reasons driving differential response have not yet been resolved. A prominent example of this was Project MATCH (Matching Alcoholism Treatments to Client Heterogeneity), a longitudinal, multi-site clinical trial of three psychosocial interventions for alcohol use disorders that attempted to identify factors to match patients to the most effective program (Project MATCH Research Group, 1993). Results showed little evidence that any specific patient-level factors (e.g., sex, anger, alcohol dependence versus abuse) predicted better response in one treatment over another (Project MATCH Research Group, 1997). In prevention research, much of the research on improving outcomes has focused

on implementation and program fidelity, with less attention to individual, person-level factors that may differentially influence both susceptibility to alcohol-related problems and prevention/intervention outcomes (Belsky & van Ijzendoorn, 2015). The incorporation of biological factors, such as genetic risk, represents an opportunity to increase understanding of why some individuals respond to prevention programming and others do not (Dick & Hancock, 2015).

The Genetics of Alcohol Use and Dependence

History of the genetics of alcohol use and dependence. Genetics play an important role in the development of alcohol problems. The genetics of alcohol were first explored through twin and family data, which allow researchers to parse the effect of genetics versus environmental influences on alcohol use behaviors (Schuckit, 2009). Through twin studies, which compare concordance rates for a given outcome between monozygotic twins (who share all of their genetic variance) and dizygotic twins (who share on average 50% of their genetics), researchers have established that the development of AUDs is partly due to genetics rather than environment alone (Tawa et al., 2016). Alcohol use disorders are approximately 50% heritable, meaning at least half of the liability for AUDs is due to genetic factors (Verhulst et al., 2015). Furthermore, twin and family studies demonstrated that the genetic risk for substance use disorders (including AUDs) is largely conveyed through a broad vulnerability to all SUDs (and related behaviors) with much smaller contributions for single substances (Kendler et al., 2003; Kendler, Jaffee, et al., 2011). Extensive efforts are now underway to identify the specific genetic factors that contribute to risk for alcohol misuse and AUDs (Clarke et al., 2017; Hart & Kranzler, 2015; Kranzler et al., 2019; Liu et al., 2019; Walters et al., 2018).

Methods for gene identification have progressed substantially over the past two decades, due in part to advancements in genetic technology and our understanding of the human genome (International HapMap Consortium, 2003; Lander et al., 2001). Genotyping was previously very expensive and there were very few known genetic markers across the genome. This limited genetic analyses to linkage and candidate gene studies of known polymorphisms. As the science progressed, it became clear that thousands of genetic variants (rather than individual candidate genes) influence alcohol misuse and related emotional and behavioral traits (Kendler et al., 2003; Krueger et al., 2002). Genome-wide association studies (GWAS) were developed to better address the polygenic nature of complex traits and behaviors. In GWAS, researchers test for associations between millions of SNPs and an outcome of interest, encompassing genetic influence across the entire genome. To accommodate multiple testing effects, very large samples and stringent p-value thresholds are used to obtain adequate statistical power (Hong & Park, 2012).

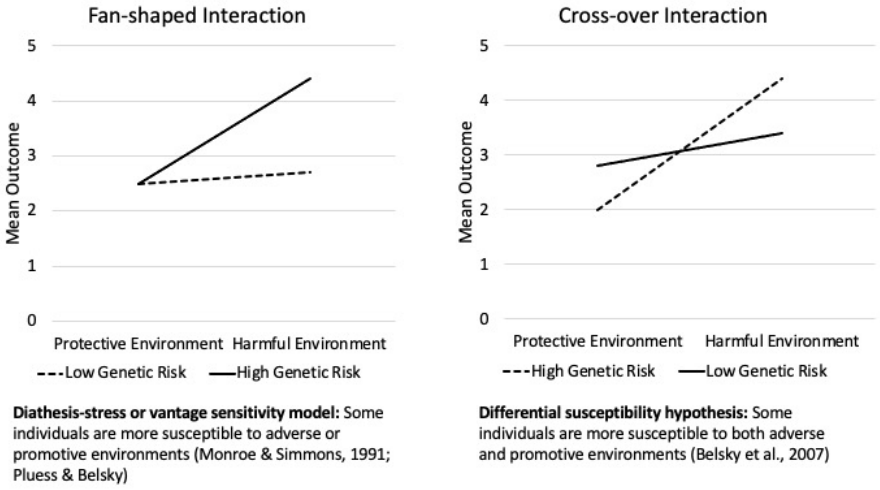
Polygenic scores capitalize on genome-wide data from GWAS to account for the many variants involved in genetic risk for alcohol and substance use (Salvatore et al., 2014). These scores are created by using large, independent samples to identify the genome-wide SNPs associated with an outcome of interest at a more liberal p-value threshold than typically required for GWAS significance (International Schizophrenia Consortium et al., 2009). The standard significant p-value threshold for GWAS is $p < 5 \times 10^{-8}$, whereas for polygenic scores the p-value threshold is adjusted based on posterior effect sizes for each SNP in the GWAS summary statistics (Ge et al., 2019; So & Sham, 2017). Given that polygenic scores include a mixture of true genetic signals and noise, using a more liberal p-value threshold allows for inclusion of a larger amount of the true genetic signals (Maher, 2015). The number of alleles for each identified

SNP is weighted according to effect size for a given outcome established in the GWAS. Researchers then use the weighted values from the independent sample to create polygenic scores in their own samples. Higher polygenic scores indicate a greater genetic predisposition for the outcome of interest. As sample sizes for GWAS discovery samples increase, the reliability of polygenic risk prediction also increases (Maher, 2015). With the decreasing cost of genotyping and collaborative efforts of consortia to combine samples, polygenic scores are likely to have increasing utility in understanding how genetic predispositions may relate to differential response to prevention and intervention programming.

Gene-environment interaction. Although an individual's genetic code is set at conception, the relative impact of genetic influence on behavioral outcomes can vary as a function of the environment. Gene-environment interaction (GxE) studies emerged as a way to understand how genes and environment dynamically interact to contribute to the development of various outcomes (Kendler, Jaffee, et al., 2011). The nature of GxE interactions can be challenging to interpret, and thus are often illustrated through graphical depictions. As discussed by Dick & Kendler (2012), a "fan-shaped" GxE interaction (Figure 1A) might occur when there is little difference in a given outcome under protective environments as a function of genotype; however, under increasingly harmful environments, the difference between genotypes becomes more apparent. Fan-shaped GxE interactions are often associated with the diathesis-stress framework, which posits that genetic influences are more pronounced under increasingly harmful environments (Dick, 2011; Monroe & Simons, 1991). Fan-shaped interactions might also result from differing effect of genotype under increasingly positive environmental experiences, a phenomenon characterized in the vantage sensitivity model (Pluess & Belsky, 2013). Another type of interaction is the "cross-over" interaction (Figure 1B), in which those who are most at

risk at one end of the environmental spectrum are least at risk at the other end of the spectrum, and vice versa. Cross-over interactions underscore the importance of measuring GxE interactions because the effect of genetic risk might otherwise be masked (Belsky et al., 2007; Dick & Kendler, 2012). The differential susceptibility hypothesis provides a framework for understanding cross-over interactions, suggesting that some individuals are more sensitive to both promotive and harmful environments based on their genotype (Belsky & Pluess, 2009).

Figure 1. Graphical depiction of gene-by-environment interactions.



Note: Graphical depiction of gene-by-environment interactions. A) Fan-shaped interaction of the effect of genetic risk on an outcome in intervention and control participants. B) Cross-over interaction of the effect of genetic risk on an outcome in intervention and control participants.

Although GxE models are useful for understanding the way biological and contextual factors come together to influence alcohol use outcomes, they also present some challenges. A common concern with GxE research is that genes and environment are often correlated (rGE), which confounds the interpretation of an interaction relationship between these factors (Jaffee & Price, 2007). Gene-environment correlation can occur through three different processes: passive,

evocative, and active rGE (Plomin et al., 1977; Scarr & McCartney, 1983). Passive rGE refers to the relationship between the genotype passed on from a biological parent and the rearing environment determined by the biological parent. Evocative rGE refers to the relationship between an individual's behavior (influenced by genetics) and the response evoked from individuals in their environment. Active rGE occurs when individuals self-select into their environments as a function of their genotype. For example, individuals with greater genetic predisposition for sensation-seeking self-select into peer groups with higher alcohol use, and in turn consume more alcohol (Yanovitzky, 2006). The question becomes whether genetic predisposition for sensation-seeking or exposure to heavy-drinking peers is the causal factor driving increased alcohol use. Ultimately, both are important, but the influence of heritable traits on both risky environments and the outcome of interest can reduce the ability to discern causal effects in GxE models. Scientific designs that permit random assignment to environments (i.e., GxE experiments) can resolve issues of rGE, and provide a clearer understanding of the way genetic and environmental factors come together to influence complex behaviors (van Ijzendoorn et al., 2011).

Gene-by-Intervention Studies

Gene by intervention (GxI) interaction studies offer one way to circumvent the limitations of rGE by experimentally manipulating the environment through random assignment to treatment conditions. When individuals are randomly assigned to conditions, their environment and genotype are inherently uncorrelated. Thus, GxI studies offer a strong design through which to understand how genetic factors may play a role in differential response to prevention and intervention. Although there are other epidemiological approaches to account for rGE, such as Mendelian Randomization (Davey Smith & Ebrahim, 2003), incorporating genetics

into randomized-controlled designs offer the most feasible opportunity to bring together genetics and interventions. Given the heritability of about 50% for AUD, it seems that genetics may indeed be important to consider in the emergence of alcohol-related problems after exposure to preventive programming. Some researchers have suggested that without the inclusion of genetics into intervention research, intervention findings may be misinterpreted by "overestimating it for some (less susceptible) individuals and underestimating it for other (more susceptible) individuals," (van Ijzendoorn et al., 2011).

The earliest GxI studies emerged in the mid-2000s, when candidate gene studies were growing in popularity. One of the first GxI studies found that a parenting intervention was effective for reducing externalizing behavior in children with the dopamine *D4* receptor (*DRD4*) 7-repeat allele, but not in children without the *DRD4* 7-repeat allele (Bakermans-Kranenburg, Van IJzendoorn, Pijlman, Mesman, & Juffer, 2008). Subsequently, a series of papers found a similar pattern of results with individuals carrying specific polymorphisms on *SLC6A4*(5HTT*)* and *DRD4* that conveyed risk for higher alcohol use under control conditions, but predicted enhanced preventive effects (i.e., lower alcohol use) under intervention conditions (Beach et al., 2010; Brody, Beach, et al., 2009; Brody, Chen, et al., 2009). The evolving popularity of the GxI design spurred two meta-analyses of GxE experimental studies (including GxI) focused on evidence for the differential susceptibility hypothesis. The authors of these papers observed relatively consistent evidence that certain genotypes predicted poorer outcomes under control conditions and improved outcomes under intervention conditions across a range of psychological outcomes (Bakermans-Kranenburg & van IJzendoorn, 2015; van Ijzendoorn & Bakermans-Kranenburg, 2015). However, most of these studies relied on candidate gene methods, which are no longer consistent with the state of the science for genetic research (Auwera et al., 2018;

Border et al., 2019; Johnson et al., 2017). The limitations of candidate gene methods are discussed further below.

GxI studies of alcohol use. Although GxI studies have grown in popularity, there are to date fewer than 20 GxI studies that focus on substance use outcomes (Neale et al., 2020). Systematic review of GxI studies on substance use outcomes identified total of 14 papers with significant GxI effects on alcohol use outcomes, including studies that used composite substance use scores. Among these studies, a variety of different genetic markers (e.g., SNPs, variable number tandem repeats) were found to moderate the effects of prevention and intervention programs. Most studies reported that individuals with higher genetic risk had higher alcohol use in control conditions and lower use in intervention conditions. The most commonly studied genetic variants were on *DRD4* and the 5-HTT linked polymorphic region on *SCL6A* (5-HTTLPR). *DRD4* is responsible for coding amino acids in the dopamine D4 receptor (McGeary, 2009a). 5-HTTLPR is a promoter sequence involved in the expression of *SCL6A,* which codes for serotonin transporters (Heils et al., 1996; Homberg & van den Hove, 2012). Improved intervention response was observed in individuals with the *DRD4* 7-repeat allele in four studies (Beach et al., 2010; Brody et al., 2014, 2015; Cleveland et al., 2015). The pattern of results for 5-HTTLPR studies was less consistent, with some studies showing the highest rates of substance misuse among *short* allele carriers in the control condition (Schlomer et al., 2017), while another study found that *long* allele carriers in the control had the highest rates of alcohol initiation and drunkenness (Cleveland et al., 2015). A third study found that individuals with higher cumulative genetic risk (including higher risk for *short* allele carriers on 5-HTTLPR) had more days abstinent after a combined batterer intervention program + brief alcohol intervention compared to standard batter intervention alone (Stuart et al., 2016). However, given that the 5-HTTLPR

genotype was collapsed into a cumulative genetic score along with a variant on *MAOA*, it is difficult to determine the degree to which 5-HTTLPR drove the GxI effects. The differing findings and lack of replication across 5-HTTLPR studies are characteristics of some of the broader limitations of candidate gene studies.

Limitations and opportunities in GxI research.

Measurement of genotype. Almost all GxI studies of alcohol and other substance use outcomes have used candidate gene methods for measurement of genotype. Although once believed to be the key to understanding genetic influences on complex behaviors, candidate gene studies are now known to be quite problematic (Latendresse et al., 2018a; Musci & Schlomer, 2018a). Across review studies, candidate gene studies have been characterized as difficult to interpret, underpowered, susceptible to publication bias, and rarely replicable across studies (Dick et al., 2015; Duncan & Keller, 2011). Candidate genes were often selected for studies based on hypotheses about their biological role in substance use and/or addictive behaviors. However, results of genome-wide association studies have rarely supported the hypothesized role of these genes (Dick et al., 2015; Duncan & Keller, 2011). Alcohol dehydrogenase (ADH) gene variants are to date one of the only candidate genes with reliable evidence for a relationship with alcohol dependence in genome-wide association studies (Clarke et al., 2017; Tawa et al., 2016; Walters et al., 2018). Substance use behaviors are also polygenic in nature, such that they are influenced by a variety of genetic markers (Kendler et al., 2003; Krueger et al., 2002). Aside from the ADH genes, which code for metabolism of alcohol, it was overly optimistic to hope that individual genetic variants would produce large effects (Dick et al., 2018). As such, candidate gene studies are insufficient to address the questions about the relationships between genetic predispositions and response to prevention and intervention programs.

Three existing GxI studies of substance use outcomes have used polygenic scores, and only one of those studies examined alcohol use behaviors as an outcome (Kuo et al., 2019; Musci et al., 2015, 2018). Musci et al. examined whether polygenic scores for smoking quit success (Uhl et al., 2010) affected the results of a school-based prevention program on age of first tobacco use (Musci et al., 2015) and marijuana use (Musci et al., 2018). However, these studies used a discovery sample of 550 European-Americans to derive their polygenic scores, which we now know is severely underpowered (Hong & Park, 2012). Current standards estimate that samples in the hundreds of thousands are necessary to reliably identify genetic variants associated with complex behaviors (Dudbridge et al., 2018).

Kuo et al. (2019) published first paper to examine alcohol use outcomes in a GxI framework using polygenic scores derived scores from a published GWAS study of alcohol dependence in of 16,087 European American subjects was published in 2019 (Gelernter et al., 2014a; Kuo et al., 2019). Findings indicated that a preventive intervention moderated the effect of polygenic risk for alcohol dependence, such that higher polygenic scores were associated with increased risk of alcohol dependence diagnosis in the control condition but not in the intervention condition. Further research is needed to determine if similar findings are observed in other samples and across other alcohol use outcomes (e.g., alcohol consumption).

Developmental factors. Most GxI studies of alcohol have focused on samples that delivered prevention intervention programs to children and adolescents (Neale et al., 2020). Only two studies tested GxI effects in adult clinical samples (Bauer et al., 2007; Stuart et al., 2016), and only one prior study focused on college-aged students (Feldstein Ewing et al., 2009). Emerging adulthood is a known period of elevated risk for the development of substance use problems (Skidmore et al., 2016; Sussman & Arnett, 2014). Although substantial resources have

been dedicated to the prevention and intervention of substance use problems in college students, the influence genetic predispositions have not yet been incorporated. The relative influence of genetic risk for alcohol and substance use varies across the lifespan, with robust evidence that genetic factors affecting substance dependence are less influential in early adolescence and become more influential across adolescence into adulthood (Dick et al., 2007; Dick, Cho, et al., 2014; Edwards & Kendler, 2013; Kendler, Schmitt, et al., 2008; Meyers et al., 2014). Additional research in emerging adults will help to clarify the relationship between genetic predispositions and response to prevention during this important developmental period.

Mechanisms of GxI effects. There is considerable need to further identify mechanisms through which genetic risk influences intervention outcomes. Studying mediators of GxI effects allows researchers to identify factors that could be harnessed to enhance intervention effects. For example, selection of peers is influenced by genetics (Kendler et al., 2007; Kendler & Baker, 2007; Tarantino et al., 2014), but can also be targeted in prevention programming (Dodge et al., 2006; Hansen & Graham, 1991; Larimer & Cronce, 2007). Two existing GxI studies have integrated mediators into their analyses (Brody et al., 2014, 2015). The first study tested mediated moderation by parenting practices targeted in the Strengthening African American Families-Teen program, and found that increased positive parenting partially mediated the GxI effect on substance use. The second study found that changing thoughts related to susceptibility to drug use through the Adults in the Making program also partially accounted for the GxI effects on youth substance use. These papers took an important step toward answering the question of how the intervention differentially affects individuals with different genotypes. The inclusion of mediators such as these is increasingly important, particularly as prevention trials begin to integrate more genome-wide methods (e.g., polygenic scores), which are by design

hypothesis-free. Polygenic scores tap into all known genetic factors observed in GWAS that predispose an individual for substance use problems, such as sensation seeking, sociability, internalizing, etc. Therefore, mediators are necessary to uncover the specific mechanisms driving the relationship between polygenic scores and prevention/intervention outcomes.

Peer deviance and drinking motives are possible mechanisms of GxI interaction effects. Peer group deviance significantly predicts alcohol and substance use (Leung et al., 2014; Stone et al., 2012). It was initially thought that this effect was largely environmentally driven; however, GxE research suggests a correlated genetic liability for substance use and peer group deviance, such that individuals carrying genetic risk for substance use problems may self-select into higher risk peer groups (Gillespie et al., 2009). Therefore, it is critical to understand the way that peers may mediate genetic risk in the context of a prevention program. Similarly, drinking motives are also heritable (specifically drinking to cope with negative emotions, drinking to enhance positive feelings) and there is evidence that they mediate genetic risk (as measured by family history) for alcoholism (Agrawal et al., 2008; Beseler et al., 2008). Examining peer deviance and drinking motives, which have environmental and genetic influences on substance use, may shed light on the mechanisms by which interventions influence alcohol use outcomes for those who are genetically at risk.

Statement of the Problem

Alcohol use and alcohol-related problems are common, and the efficacy of current alcohol prevention and intervention programs is limited (Huh et al., 2015; Schulenberg et al., 2017; Spoth et al., 2008). Across alcohol prevention programs for youth and emerging adults, some participants respond positively while others do not (Smit et al., 2008; Stigler et al., 2011; Tanner-Smith & Risser, 2016). Alcohol use and dependence are influenced by both genes and

environment, with heritability estimated at about 50% (Verhulst et al., 2015). As genes explain some of the risk for AUDs, GxI studies have the potential to improve understanding of why prevention intervention programs are differentially effective across individuals (Belsky & van Ijzendoorn, 2015). Existing GxI studies are limited in scope by predominantly using candidate gene methods, which are inconsistent with the state of the science (Dick et al., 2015; Duncan & Keller, 2011; Neale et al., 2020). Most GxI studies have been conducted in adolescent samples, and few examined mediators of GxI effects to help explain mechanisms through which genetics differentially affect prevention outcomes. Additional research that integrates genome-wide methods, emerging adult samples, and mediators of GxI effects may help to expand understanding of both genetic factors associated with outcomes and mechanisms that contribute to intervention effects. Completion of this research may lead to improvements of prevention and intervention programs by presenting opportunities to develop more effective, tailored programs.

Current Study

The present study investigated whether intervention outcomes vary as a function of genetic risk for alcohol problems and explored factors that may explain how those differences may occur. The present study also examined whether findings generalized across samples, which differed on treatment modality and developmental timing. This research capitalized on data from two existing resources: a college prevention intervention program focused on level of response to alcohol conducted with a genetically-informed sample (Study One: Spit for Science), and a middle school-based prevention program targeting adolescent problem behavior with longitudinal follow up and genetic data (Study Two: Project Alliance). Alcohol misuse is influenced by genetic and environmental factors and these factors are correlated. In order to develop more effective, tailored prevention intervention programs, it is important to understand

how response to programs may differ according to genetic predispositions. Completion of this research may result in improvement of prevention and intervention programs through increased understanding of both individual factors associated with outcomes and mechanisms that contribute to intervention effects. The following sections of the paper are organized by study, beginning with the research aims and hypotheses, methods, results, and discussion for Study One (Spit for Science). The research aims and hypotheses, methods, results, and discussion for Study Two (PAL) are then presented, and the paper concludes with a global discussion section.

Research Aims and Hypotheses: Study One

Guided by the extant literature and relevant theoretical models, the proposed research has the following aims and hypotheses:

1. The first aim is to determine if genetic predispositions toward alcohol misuse moderate the effectiveness of an alcohol intervention in a college student sample using polygenic risk scores (PRS) associated with externalizing behaviors.
    a. Informed by the differential susceptibility hypothesis, I hypothesized that individuals with higher PRS will show greater reductions in alcohol use and alcohol use disorder symptoms post-intervention.
2. The second aim is to examine whether peer deviance and drinking motives mediate changes in alcohol use for those at greater genetic risk in a college student intervention sample.
    a. I hypothesize that the intervention will lead to lower peer deviance among those genetically at risk, which will partially account for lower levels of alcohol use and problems in intervention participants.
    b. Secondly, I hypothesize that the intervention will lead to lower levels of drinking to cope and drinking to enhance positive feelings among those genetically at risk, which will partially account for lower alcohol use and problems among intervention participants.

Methods: Study One

Spit for Science Sample

Parent study participants. Spit for Science is a longitudinal study of genetic and environmental factors that influence alcohol use, other substance use, and emotional health outcomes in college students at a large, mid-Atlantic university (Dick, Nasim, et al., 2014). The study invited all first-time freshman college students aged 18 and older to complete an online survey at the start of fall semester. New cohorts were recruited each year in 2011-2014 and 2017. The overall participation rate in the study was 64% (N= 12,370). Of those who picked up survey compensation, 97% (N=11,147) also provided a DNA sample. Every spring semester, participants were invited to complete a follow-up survey to assess changes in alcohol use, other substance use, and emotional health across the college years and beyond. Participants were compensated $10 for each survey and an additional $10 for providing a DNA sample. Genotypic data is currently available for cohorts 1-3.

Prevention study subsample. In collaboration with researchers at the University of California, San Diego, we designed a spin-off study to measure the effectiveness of tailoring an alcohol prevention program to low level of response (LR) to alcohol (Schuckit et al., 2012, 2015). Low LR predisposes an individual to heavy drinking because he/she must consume more alcohol to feel the same effects as the average drinker (Schuckit et al., 1997). To carry out the study, we recruited a subset of the 2013 cohort of S4S participants to take part in an alcohol prevention study (referred to as S4S-LR) in September of their freshman year. Eligibility was limited to students who 1) completed the S4S survey within the first two weeks of data collection (83% of 2,022 individuals), 2) endorsed drinking alcohol at least five or more times in their life (72.7%), and 3) scored 0.25 standard deviations above or below the mean for the Self-Rating of

the Effects of Alcohol (SRE) scale (44.6%). The SRE score is calculated by taking the average of four items asking how many drinks it takes to feel certain intoxication effects of alcohol (Schuckit et al., 1997). Higher SRE scores equate to a lower level of response to alcohol. Of the 572 students invited to participate, 231 (56.5%) enrolled in the S4S-LR prevention study (n=104 with low LR and n=127 with high LR). The sample was mostly female (n=165, 71.4%), and 0.4% American Indian/Native American, 6.9% Asian, 11.3% Black/African American, 5.6% Hispanic/Latino, and 75.8% White.

Prevention program and procedure.

Prevention program. The S4S parent study and the S4S-LR spin-off study were approved by the university's Institutional Review Board. Following electronic informed consent, S4S-LR participants (n=231) were paired according to similar alcohol use and demographic characteristics but mismatched SRE classification (i.e., high or low LR). Pairs were then randomly assigned to one of two online alcohol prevention programs: a standard, "one size fits all" approach ("state of the art", or SOTA) and a tailored approach based on level of response to alcohol ("level of response based" or LRB). Participation in the study involved completion of online video modules (SOTA or LRB) once a week for four weeks, as well as a 30-day follow up assessment for up to $100 in compensation (Schuckit et al., 2012). The annual S4S spring semester follow-up surveys were used to further assess long-term effects of the SOTA and LRB programs. The primary goal of the study was to determine if the tailored approach was more effective in reducing alcohol consumption and problems among college students who differed in their LR to alcohol. Initial findings showed that students with low LR who completed the tailored LRB program reported lower levels of maximum drinks in 24 hours than low LR students in the SOTA program approximately five months after the intervention (Savage et al.,

2015). Our results provided robust evidence that overall, individuals with the riskier, low LR showed greater reductions in alcohol use over time than high LR individuals in either program. Compared with students who received no alcohol prevention, we observed strong effects of either prevention program (LRB or SOTA) on risky outcomes (maximum drinks and AUD symptoms). Based on these findings, we plan to collapse across S4S-LR prevention groups to compare any prevention (LRB or SOTA) to a no prevention control group in the proposed analyses.

Creation of comparison group. Both conditions of S4S-LR sample (LRB and SOTA) received a relatively intensive alcohol prevention program, the results of which were fairly similar. To address the research questions of the present study, a comparison group of participants who received no intervention was needed. The comparison group was derived from the pool of S4S participants with genotypic data who were not invited to the LR study. Propensity score matching (PSM) was used to derive scores that imitate the similarity of baseline characteristics between treatment groups achieved in a randomized controlled trial (Austin, 2011; Guo & Fraser, 2010). The propensity score is considered the probability of consenting to a treatment based upon a set of baseline covariates (Austin, 2011; Rosenbaum & Rubin, 1983). When individuals are matched using the propensity score, the effects of selection bias should be mitigated. Following calculation of the propensity score for treated and untreated groups, individuals can be matched using a one-to-one approach or one-to-many, in which one treated individual is matched to several control individuals with similar propensity scores (Austin, 2011). Although the one-to-many matching approach can result in a small increase in bias, it also yields greater precision relative to one-to-one matching and increased sample size (Rassen et al., 2012). Given these relative strengths and the abundance of potential controls available through

the parent S4S sample (N=4199 with genotypic data available), the one-to-many PSM approach was used to derive a comparison group.

Measures. Processed and cleaned genotypic data as well as longitudinal phenotypic data from the S4S surveys are included in the analyses. Phenotypic data is accessed through secure data sharing procedures managed by the S4S Registry Coordinator. Relevant measures for the present study include alcohol use behaviors, peer deviance, drinking motives, and covariates. Genotypic data is housed on a secure server at Virginia Commonwealth University. The sections below provide further detail on the data available and how it was used for the present study.

Alcohol use and problems. Typical quantity and frequency of alcohol consumption in the past 30 days was measured at each wave using two items from the Alcohol Use Disorder Identification Test (DeMartini & Carey, 2012). Alcohol consumption will be measured by grams of ethanol per month, which is calculated by multiplying typical quantity * typical frequency * 14 (the number of grams of ethanol in a standard drink). See Salvatore et al. (2016) for additional information about creation of the grams of ethanol per month measure. The Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) was adapted to measure DSM-5 criteria for AUD at each wave of the S4S survey (Bucholz et al., 1994). SSAGA items assess how often students have experienced alcohol-related consequences (e.g., drinking in dangerous situations, alcohol-related arrests), symptoms of dependence (e.g., withdrawal, tolerance, desire to quit), and impact on daily functioning (e.g., interference with work/school, relationship problems). Response options were "Never," "1–2 times," "three or more times," or "don't know."

Peer deviance. Peer deviance was measured using a set of six items previously operationalized in the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders

(Kendler, Jacobson, et al., 2008). Participants reported on the proportion of high school friends (at baseline) and college friends (at follow-up) who engaged in certain deviant and problem behaviors, such as getting drunk, getting in trouble with the law, or using marijuana. Response options ranged from "none" (0) to "all" (5). A mean score is computed for all participants who responded to 50% or more of the items. For inclusion in mediation analyses, change scores were calculated by subtracting the baseline measure (pre-intervention) from the first follow-up measure of peer deviance (post-intervention).

Drinking motives. Four subscales of drinking motives (Social, Coping, Enhancement, Conformation) were measured using the Drinking Motives Questionnaire at each wave of S4S data collection (Cooper, 1994). Responses were coded such that higher scores conveyed greater motivation to drink for that reason. A mean score for each subscale was created by averaging the score for each item within each scale.

Covariates. Covariates include sex, ancestry principal components, prevention program (LRB or SOTA), and LR to alcohol. Ancestry principal components (PCs) account for variation in allele frequency across different population structures. They are derived through Principal Components Analysis (PCA), which is further detailed below. LR was measured using the Self-Rating of the Effects of Alcohol (SRE) scale (Schuckit et al., 1997). The SRE score is calculated as an average of four items measuring how many drinks it takes to 1) feel tipsy or have a buzz, 2) dizzy or slur speech, 3) stumble or find it hard to walk properly, and 4) pass out. Prevention program (LRB or SOTA) was collinear with intervention group; after confirming that it was not significantly associated with the outcomes of interest, prevention program was removed from analyses.

Genotypic data. DNA samples were collected using Oragene kits and extracted using

standard procedures. Samples were then sent to the Rutgers University Cell and DNA Repository where genotyping was completed. Samples from the first three cohorts were genotyped on the Affymetrix Axiom BioBank Array Version 2, which includes a 296K single nucleotide polymorphisms (SNPs) GWAS imputation grid, 197K non-synonymous coding SNPs, 18K insertions/deletions, novel exome/loss of function variants, and 16K eQTL markers. Quality control procedures removed SNPs missing more than 5% of genotype, samples missing greater than 2% of genotypes, and SNPs missing more than 2% of genotypes after filtering. Genotypes were imputed using SHAPEIT2/IMPUTE2 (Delaneau et al., 2013; Howie et al., 2009) and the 1000 genomes phase 3 reference panel (Sudmant et al., 2015; The 1000 Genomes Project Consortium, 2015). Additional details about genotyping the Spit for Science sample are available in Webb (2017).

Data Analysis Plan: Study One

All analyses were conducted using R, which is a flexible statistical computing program with several available methods for handling missing data. R's functionality is expanded through various packages built to run advanced statistical techniques described below.

Data preparation. All phenotypic variables were examined for normality. Log+1-transformations were computed when appropriate to reduce the effects of non-normality. Due to the longitudinal nature of the study, some missing data is expected. However, the inclusion criteria for the intervention sample and comparison group required complete data on key baseline variables: gender, alcohol use, and LR to alcohol. For the growth curve models, individuals missing greater than one time point were excluded list-wise. Complete data was required for mediation analyses.

Creation of comparison group. Propensity score matching (PSM) was conducted using the *MatchIt* package in R (D. Ho et al., 2011). Following procedures outlined by Guo & Fraser (2014), the first step in PSM was to identify covariates that may influence the likelihood of receiving treatment. Informed by existing research on factors that influence research participation and relevance to the present study, the following covariates were included in the propensity score calculation: sex, race/ethnicity, age, socioeconomic status, participant occupational status, SRE score, alcohol consumption, AUD symptoms, family history of alcohol or substance use problems, depression, anxiety, and antisocial behavior (Kelpin et al., 2018; Krueger et al., 2002; Patel et al., 2003; Savage et al., 2015). These variables were entered into a multivariate logistic regression predicting intervention participation, and then used to derive a probability of intervention participation for each individual. The estimated propensity scores are the predicted probability of treatment given the observed covariates. The second step of PSM involves resampling to obtain a set of control participants that optimally balances the covariates. The matching procedure conducted pair matching based on similar propensity scores with a variable ratio of one intervention participant to two control participants. Post-matching analysis compared intervention and control participants on key demographic and baseline variable. Any significant differences between intervention and control participants are reported and controlled for accordingly.

Creation of polygenic risk scores (PRS). PRS were calculated using prioritized SNPs identified in an independent genome-wide association study (GWAS) of externalizing behavior in approximately 1.5 million subjects (Karlsson Linnér et al., 2020). The GWAS of externalizing behaviors measures genetic factors associated with a collection of related phenotypes (ADHD, alcohol dependence, alcohol consumption, cannabis use, age at first sex,

number of sexual partners, general risk tolerance, and tobacco use) characterized by behavioral disinhibition, impulsive behaviors, and/or deficits in self-regulation (Karlsson Linnér et al., 2020). In adolescence, broader externalizing risk has a greater influence on alcohol misuse than genetic risk specific to alcohol and genetic risk for externalizing behaviors is more likely to manifest earlier than alcohol-specific risk (Kendler, Gardner, et al., 2011; Meyers et al., 2014). As a result, the Karlsson Linnér et al., (2020) externalizing GWAS was selected for use in this study.

Polygenic scores for this study were derived with PRS-CS, which uses continuous shrinkage priors within a Bayesian regression framework to adjust shrinkage based on the strength of a SNP's association in GWAS (Ge et al., 2019). Scores are calculated by multiplying the number of score alleles by the $\log_{10}$ of the weighted SNP effect (beta). For example, consider SNP 1 (A/G). If the externalizing GWAS summary statistics indicate that increasing copies of the A allele are associated with externalizing behavior (OR = 1.80) then a S4S participant carrying two copies of the A allele will have a score of $2 * \log_{10}(1.80) = 0.51$ for SNP 1. The values of all prioritized alleles are then summed to create the polygenic score, with higher polygenic scores indicating greater predisposition for risky behaviors. PRS-CS then uses linkage disequilibrium (LD) patterns observed in the 1000 Genome Phase 3 European Ancestry reference panel with 500kb physical distance and an LD threshold of $r^2 >= 0.25$ (The 1000 Genomes Project Consortium, 2015). The final PRS scores were calculated using the PLINK 1.9 *score* procedures, averaging across the total number of non-missing SNPs in the sample (Chang et al., 2015; Purcell et al., 2007). PRS-CS has been shown to improve prediction over other pruning and clumping methods for polygenic score calculation, with the added benefit of providing a single p-value threshold to optimize prediction in the target sample. Of note, the discovery

sample for the PRS was composed of individuals of European-descent, so additional steps are needed to accommodate other ancestry groups in analyses. The recommended approach is to conduct analyses separately within adequately sized ancestry groups (Tian et al., 2008); however, the S4S-LR study sample is 75% White. The other ancestry groups were inadequately powered to conduct growth curve analyses, and as a result they were excluded from the analyses.

Principal Components Analysis (PCA). Principal components (PCs) were included as covariates to account for population stratification as well as admixture. Following quality control procedures, PCA was conducted for ancestry-specific PCs using EIGENSOFT and SmartPCA. Regions of high LD were excluded using PLINK 1.9, so as to ensure relative independence of SNPs. The 1000 Genome Project phase 3 reference panel was used in the PCA for European ancestry. Ancestry group assignment for S4S participants is described in detail in Peterson et al., (2017).

Evaluation of Externalizing PRS in S4S. To evaluate the association between the externalizing genome-wide polygenic score (EXT PRS) and relevant outcomes (alcohol consumption, AUD symptoms, peer deviance, and antisocial behavior), a hierarchical multiple regression was conducted estimating the effect of the EXT PRS on alcohol consumption and AUD symptoms over and above the effect of ancestry PCs. The purpose of these analyses was to validate use of the EXT PRS in the S4S sample by ensuring that the polygenic score was predictive of relevant phenotypes in the expected direction of effect. To account for multiple testing, *p*-values were adjusted using the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).

Aim 1: Tests of GxI effects on Alcohol Consumption and AUD Symptoms. Aim 1 analyses were conducted using growth curve modeling within a multilevel framework to estimate

the effect of intervention and PRS on trajectories of alcohol consumption and AUD symptoms across time. The multilevel framework accommodated the nested structure of repeated measures within individuals, while also allowing for between individual and between-group variation in intercepts and slopes of the dependent variables. A model building approach was employed to compare the goodness of fit between the unconditional (null) model, unconditional growth model, and conditional growth model with time-invariant covariates (ancestry PCs, LR to alcohol, propensity score, and sex). Analyses were conducted using the *lme4* package for R, with maximum likelihood estimation. Model fit was evaluated using AIC, intraclass correlation (ICC), and pseudo R^2. Alcohol consumption and AUD symptoms were examined in separate models.

Unconditional models. First, we constructed an unconditional null model to provide an estimate of the within-person (Level 1) and between-person (Level 2) variance components. The unconditional null model serves as a base model with no predictors to test whether model fit improves with the addition of Level-2 effects in subsequent models. Next, an unconditional growth model was fit. The unconditional growth model estimates the trajectory of the outcome across four time points, with Time as a predictor at Level 1. As there were no significant differences between intervention and control groups at baseline, Time was centered at first follow-up (Time 1), with each successive follow-up coded to account for approximately equal time between follow-up assessments (Time 1=0, Time 2=1, Time 3=2, and Time 4=3). Linear and curvilinear effects for time were tested, with the quadratic effect providing a better fit for the data. A comparison of model fit between the unconditional null model and the unconditional growth model confirmed there was sufficient individual variability to warrant advancing to conditional models, which include predictors to estimate variation in intercept and/or slope.

Conditional models. The first conditional growth model included Time as the Level 1 variable, and intervention group and EXT PRS as Level 2 variables. Fixed and random effects for slope and intercept were tested, with random slope and intercept providing the best fit for the data. In the second conditional model, 2-way and 3-way interaction terms were added to examine the degree to which interactions between Time, EXT PRS, and intervention group contributed to variation in the outcomes. In the final model, sex, LR to alcohol, propensity score, and ancestry PCs were added as covariates to account for their potential impact on the resulting models. The *p*-values for the Level 1 variables (intercept, Time), Level 2 variables (intervention group, EXT PRS), and Interaction components were adjusted for multiple testing using the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).

Post-hoc analyses. Visual examination of the data suggested potential differences between intervention and control on AUD symptoms at the first follow-up when intervention effects were most robust. Post-hoc analyses were conducted to explore more proximal effects of GxI interaction on change in AUD symptoms from baseline (Time 0) to the first follow-up assessment (Time 1). Hierarchical multiple regression was used to estimate the effects of polygenic risk, intervention group, and their interaction on change in alcohol consumption and AUD symptoms from baseline to first follow-up. Hierarchical multiple regression analyses were conducted using the *lm* function in the *stats* package for R, and regions of significance were explored using the *interactions* package for *R.* Analyses controlled for sex, LR to alcohol, propensity score, and ancestry PCs.

Aim 2: Examining peer deviance and drinking motives as mediators of gene-by-intervention effects in the S4S-LR sample. Aim 2 analyses test whether changes in peer deviance and drinking motives mediated an effect of the intervention on AUD symptoms at Time

1 for those at varying levels of genetic risk. For peer deviance, drinking to cope, and drinking to enhance, change scores were calculated by subtracting the Time 0 score from the Time 1 (first follow-up) score. Correlations of all mediating, moderating, predictor, and outcome variables were conducted to determine whether a multiple mediator model was warranted. Observing no significant correlation between mediators, we proceeded in constructing three separate mediated moderation models to estimate the direct and indirect effect of peer deviance, drinking to cope, and drinking to enhance on AUD symptoms varying levels of EXT PRS and intervention status. Analyses were conducted using the *mediation* package in *R*. In the first step, we regressed the effect of EXT PRS * intervention and covariates (gender, LR to alcohol, ancestry PCs, and propensity score) onto the centered mediating variable. Next, we estimated the effect of EXT PRS * mediator, EXT PRS * intervention, and covariates on the outcome variable (change in AUD symptoms). In the final step, we specified the levels of the moderator (EXT PRS) at which to calculate the mediation function, setting the values of EXT PRS at 1 standard deviation above and below mean. Finally, we tested for significant differences in the total, direct and indirect moderating effects of EXT PRS and intervention status on AUD symptoms through peer deviance, drinking to cope, and drinking to enhance, using a bias-corrected and accelerated bootstrap resample of 2,000 to calculate the 95% confidence interval for the indirect effects.

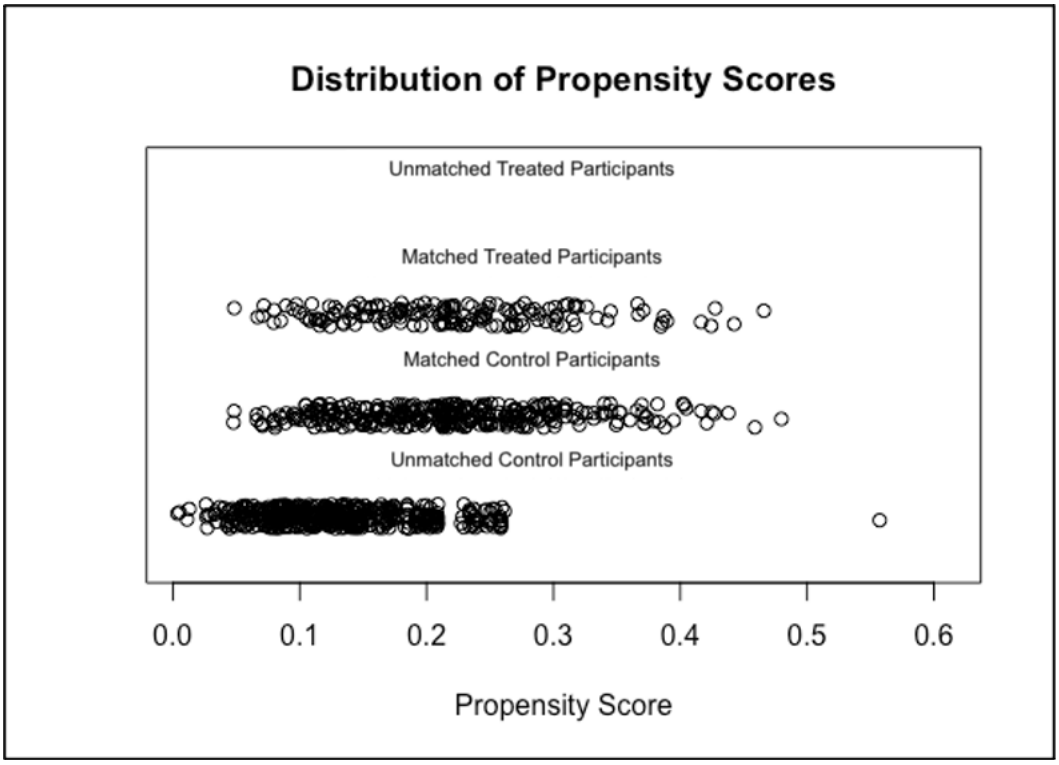
Results: Study One

Preliminary Analyses

Propensity Score Matching. In order to approximate the effect of random assignment, we conducted propensity score matching to identify a group of untreated "controls" with similar propensity to participate in the intervention. We included the following variables in the propensity score calculation: gender, socioeconomic status (as measured by parental education), occupation status, level of response to alcohol, grams of ethanol consumed per month, alcohol use disorder symptoms, familial risk for alcohol or substance use problems, depression symptoms, anxiety symptoms, and high school antisocial behavior. The matching procedure was limited to individuals of European ancestry with genotypic data to align with the characteristics of the intervention study participants included in analyses. Individuals were eligible for the intervention study if they had initiated alcohol use and responded to two or more items in the Self-Rating of the Effects of Alcohol scale. After applying these limitations, there were a total of 740 available controls for 161 treated individuals. Using a 2:1 ratio of controls to treated individuals, the propensity score matching successfully matched 322 untreated controls to 161 treated individuals. Figure 2 shows the distribution of propensity scores for unmatched treated participants, matched treated participants, matched control participants, and unmatched control participants. All treated individuals were successfully matched; thus, the distribution of unmatched treated participants is empty. The unmatched controls have higher density in the lower range of propensity scores, and as a result were not matched to the treated individuals.

Figure 2.

Distribution of Propensity Scores for Unmatched Treated, Matched Treated, Matched Control, and Unmatched Control Participants.



T-tests were conducted to compare the difference between means for the matched treated participants and matched control participants. There were no significant differences between the matched treated and matched controls on the variables displayed in the table.

Table 1.

T-tests comparing matched treated participants and matched control participants.

Predictor	t	df	*p*	Mean Diff	SE	95% CI
Gender (female)	.000	481	1.00	.000	.044	[-.086, .086]
College educated parent	-.374	481	.709	-.016	.042	[-.097, .066]
Full-time employed	.451	481	.652	.006	.014	[-.021, .033]
Part-time employed	-.236	481	.814	-.009	.039	[-.087, .068]
LR to Alcohol	.017	481	.986	.004	.258	[-.502, .511]
Time 0 Alcohol consumption	.220	481	.826	8.416	38.33	[-66.92, 83.75]
Time 0 AUD symptoms	-.117	481	.907	-.025	.212	[-.442, .392]
Familial risk for alcohol/drug problems	.491	481	.624	.0318	.0648	[-.096, .159]
Depression	-.496	481	.620	-.171	.344	[-.847, .506]
Anxiety	.298	481	.766	.087	.292	[-.487, .661]
Antisocial Behavior (HS)	.000	481	1.00	.000	.220	[-.432, .432]
Age	-.474	481	.635	-.0157	.033	[-.080, .049]

Evaluation of Externalizing PRS in S4S. Results of hierarchical multiple regression analyses indicated that the EXT PRS was significantly associated with alcohol consumption, AUD symptoms, peer deviance, and antisocial behavior at baseline in the S4S sample. After controlling for ancestry PCs and correcting for multiple testing, the externalizing PRS significantly predicted AUD symptoms and alcohol consumption at most time points, accounting for approximately 1-1.5% of the variance in baseline measures of all phenotypes examined (Table 2, Figure 3). The strength of the effect diminished across follow-up time points for all phenotypes examined, which may be in part due to sample attrition. Baseline (Time 0) measures of Peer Deviance and Antisocial Behavior indexed lifetime report of these constructs, whereas Time 1-4 indexed report of these behaviors in the time since last assessment. Accordingly, Baseline (Time 0) measures account for a longer period of time than subsequent follow-up

assessments of the same measures. One possible alternative approach to assessing EXT PRS at each wave would be to take the highest value of each outcome across waves, which would provide the benefit of using all of the available information for each individual participant. However, that approach would also result in censoring of data, which can occur when there are incomplete observations due to attrition. Prior analyses in the Spit for Science sample have indicated that students who consumed more alcohol were more likely to withdraw early from the university and thus were not retained for follow-up data collection efforts (Ho et al., 2016). Taking the highest value across waves would compare alcohol use behaviors for individuals who were retained across all four years, to those who only have baseline measures despite possible unmeasured escalation of use across the following years. After comparing the relative benefits and drawbacks of each approach, we proceeded with assessing EXT PRS at each wave for the four outcomes of interest to reduce the impact of attrition of these analyses.

Table 2.

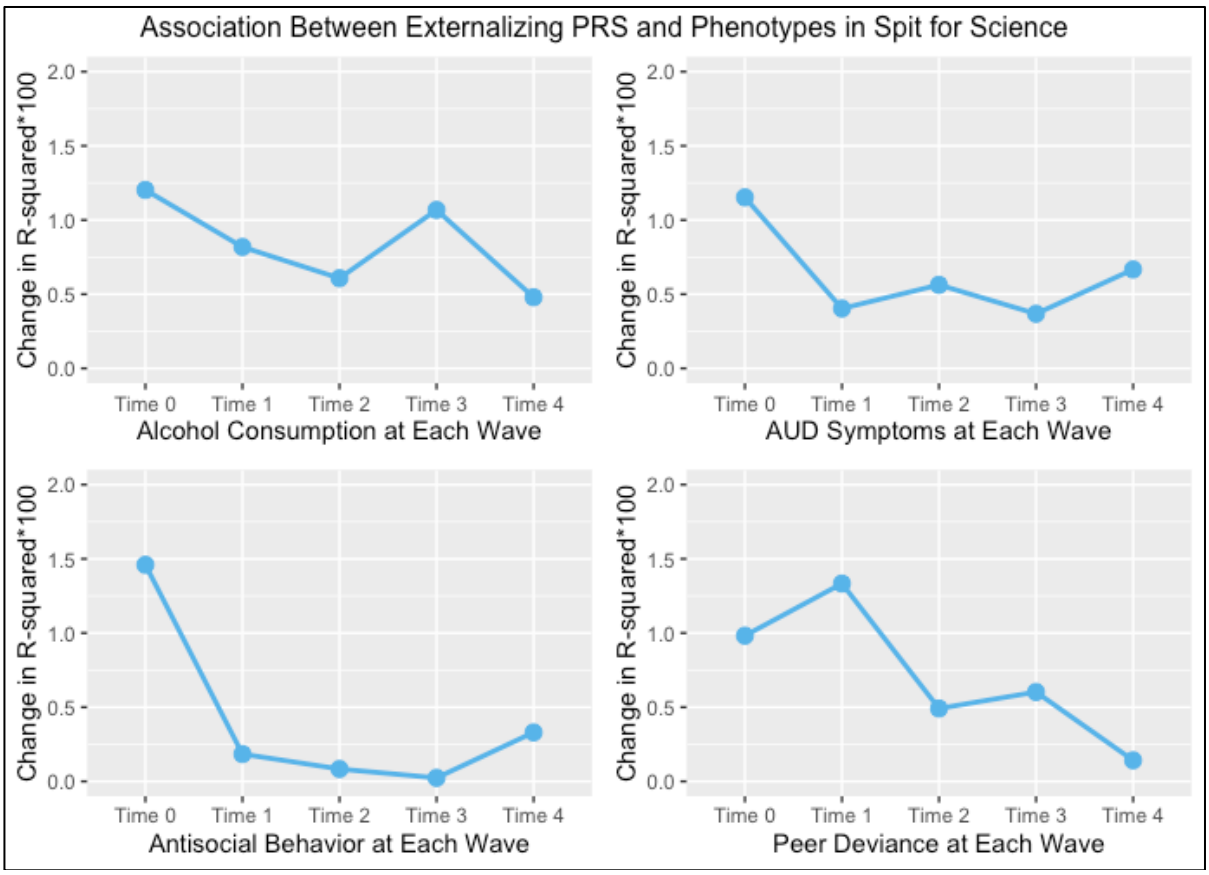
Variance in relevant phenotypes accounted for by the externalizing PRS across five time points.

Outcome	N	Change in R^2	*p*	*p* [a]	Ancestry PCs R^2
Time 0 (Baseline) AUD symptoms	2622	.012	**<.001**	**<.001**	.005
Time 1 AUD symptoms	1698	.004	**0.009**	**.019**	.016
Time 2 AUD symptoms	1335	.006	**0.006**	**.015**	.008
Time 3 AUD symptoms	1033	.004	0.051	.071	.014
Time 4 AUD symptoms	843	.007	**0.017**	**.030**	.029
Time 0 (Baseline) Grams of ethanol per month	2037	.012	**<.001**	**<.001**	.006
Time 1 Grams of ethanol per month	2038	.008	**<.001**	**<.001**	.013
Time 2 Grams of ethanol per month	1282	.006	**0.005**	**.013**	.008
Time 3 Grams of ethanol per month	1024	.011	**0.001**	**.004**	.020
Time 4 Grams of ethanol per month	837	.005	**0.043**	.063	.032
Time 0 (Baseline) Peer Deviance	2965	.010	**<.001**	**<.001**	.002
Time 1 Peer Deviance	2241	.013	**<.001**	**<.001**	.005
Time 2 Peer Deviance	1393	.005	**0.009**	**.019**	.006
Time 3 Peer Deviance	1077	.006	**0.011**	**.021**	.008
Time 4 Peer Deviance	872	.001	0.264	.292	.020
Time 0 (Baseline) Antisocial Behavior	2967	.015	**<.001**	**<.001**	.007
Time 1 Antisocial Behavior	2241	.002	**0.042**	.063	.006
Time 2 Antisocial Behavior	1391	8.34E-04	0.28	.292	.014
Time 3 Antisocial Behavior	1075	2.35E-04	0.614	.614	.014
Time 4 Antisocial Behavior	876	.003	0.088	.116	.017

Note: Change in R^2 represents the variance accounted for by EXT PRS over and above the effect of Ancestry PCs. Bolded values indicate significant *p*-values less than .05.
[a]*p*-values have been adjusted for multiple testing using the Benjamini–Hochberg procedure.

Figure 3.

Plot of variance in relevant phenotypes explained by externalizing PRS across waves



Aim 1: Results of Multilevel Models Examining GxI Effects on Alcohol Consumption and AUD Symptoms Across Time.

Alcohol Consumption. Results of the multilevel models (MLM) for alcohol consumption are displayed in Table 3. The MLM was first constructed as an unconditional means model (Null Model; Table 3), with an estimated Intra-Class Correlation (ICC) coefficient of .55, indicating that slightly more than half (55%) of the variance in alcohol consumption was due to differences between individuals based on parameters in the model. In the unconditional growth model (Unconditional Growth Model; Table 3), the significant effects for intercept, time, and time^2 indicated there was sufficient within-person variation in trajectories of alcohol

consumption to warrant a multilevel framework. The intercept value for the Unconditional Growth Model indicates that the average value for log-transformed alcohol consumption at Time 1 was 5.21. The positive linear slope, represented by Time (7.04), indicated that on average, alcohol consumption increased over time; however, the quadratic equation provided a significantly better fit for the data. The negative quadratic slope, represented by $Time^2$ of -3.11, indicated that the slope of alcohol consumption was curvilinear in nature with alcohol consumption increasing over time for Times 1-3 before dropping slightly at Time 4 (Figure 4). Results of a likelihood ratio chi-square difference test comparing the Unconditional Growth Model with random intercept and slope (AIC=2928.7, ICC=.64) to the Null Model (AIC=2955.2, ICC=.55) showed significantly improved model fit with the Unconditional Growth Model [$X^2(4, 375) = 34.46$, $p<.001$].

Next, we constructed the Conditional Growth Model with fixed effects for time, intervention group, and EXT PRS, and random effects for slope and intercept at Level 1. The Conditional Growth Model adds to the unconditional growth model through the inclusion of main effect predictors of change in the outcome over time. We compared conditional models with random intercept only, random slopes only, and a combined model with random intercept and slopes, with the latter providing the overall best fit for the data structure (AIC=2701.69, ICC=.63). In the Conditional Model, the intercept, linear slope, and quadratic slope remained significant; however, there was no evidence of significant main effects for intervention group or EXT PRS. Next, an Interaction Model was constructed to evaluate whether there were variations in the slope and intercept of alcohol consumption across time as a function of the interaction between intervention group, EXT PRS, and time. In the Interaction Model, the intercept and linear slope remained significant, but the quadratic slope was no longer significant, suggesting

that the interaction terms may account for some of the variance in the quadratic effect of time. We also observed no significant two-way or three-way interactions Time, Time^2, intervention group, or EXT PRS, indicating there is no evidence suggest that trajectories of alcohol consumption vary as a function of the interaction between intervention group or EXT PRS. We proceeded with the addition of covariates to fully evaluate the final model. The effects for intercept and linear slope remained significant after the addition of covariates and multiple testing corrections. There were also significant main effects for two covariates, propensity score and LR to alcohol, suggesting that these factors significantly influenced trajectories of alcohol consumption among college students. The significant positive association between LR to alcohol and alcohol consumption suggests that individuals with higher LR consumed more alcohol, whereas the significant negative association between propensity score and alcohol consumption suggest that individuals with higher propensity scores consumed less alcohol on average. The propensity score includes a number of indicators used to estimate the likelihood of agreeing to participate in the intervention, which means that this effect indicates individuals with higher likelihood of agreeing to participate in an intervention tended to drink less alcohol on average. No significant main effects for intervention group or EXT PRS were observed, and there remained no evidence to suggest an interaction between intervention group and EXT PRS on the slope of alcohol consumption.

Table 3.

Multilevel Growth Curve Analysis of Alcohol Consumption from Time 1 to Time 4

	Model				
Level and Variable	Null	Unconditional Growth	Conditional Growth	Interaction Model	Interaction Model with Covariates[1]
Level 1					
Intercept	5.12 (.07)**	5.21 (.07)**	5.27 (.09)**	5.26 (.09)**	5.20 (.27)**
Time		7.04 (.1.82)**	6.77 (1.82)**	6.05 (2.33)**	6.27 (2.37)**
Time^2		-3.11 (1.50)*	-3.35 (1.54)*	-3.74 (1.99)	-4.01 (2.02)
Level 2					
Intervention (control=0)			-.15 (.14)	-.15 (.15)	-.13 (.16)
EXT PRS			-.02 (.07)	-.02 (.09)	.03 (.09)
Interaction components					
Time * EXT PRS				-.09 (2.34)	.02 (.03)
Time * Intervention				.82 (3.78)	.00 (.06)
Time^2 * EXT PRS				-.22 (1.99)	-.18 (1.99)
Time^2 * Intervention				.67 (3.21)	.18 (3.22)
EXT PRS * Intervention				-.06 (.15)	.05 (.12)
Time * EXT PRS * Intervention				-5.53(3.92)	-5.31 (3.93)
Time^2 * EXT PRS * Intervention				-.18 (3.30)	-.74 (3.31)
Covariates					
Gender (female=1)					.02 (.21)
LR to Alcohol					.13 (.03)**
Propensity Score					-3.27 (1.02)**
Additional Information					
ICC	.55	.64	.63	.63	.59
-2 log likelihood (FIML)	-1474.6	-1457.4	-1341.8	-1240.1	-1317.6
AIC	2955.2	2928.7	2701.69	2712.15	2693.16
Pseudo R^2 (fixed effects)	.00	-	-	-	-
Pseudo R^2 (total)	.55	-	-	-	-
Number of individuals	483	375	345	345	344
Observations	879	879	816	816	814

Note: Estimates of unstandardized coefficients are presented for fixed effects. Values in parentheses are standard errors. Alcohol consumption was log-transformed to account for skewness and kurtosis. Pseudo R^2 cannot be calculated for models with a quadratic slope. The Interaction Model with Covariates also controlled for the effect of Ancestry PCs 1-10. *$p < .05$, **$p < .01$.

[1] p-values for Level 1, Level 2, and Interaction components were adjusted for multiple testing using the Benjamini–Hochberg procedure in the Interaction Model with Covariates.

AUD Symptoms. Model building procedures for the MLM for AUD symptoms mirrored those described for alcohol consumption. The results are displayed in Table 4. The Null Model estimated an ICC of .54, indicating that slightly more than half (54%) of the variance in alcohol use disorder symptoms was due to differences between individuals. In the Unconditional Growth Model, significant effects for intercept, linear time, and quadratic time indicated sufficient within-person variation in trajectories of AUD symptoms to warrant a multilevel framework. The intercept value of 1.22 for the Unconditional Growth Model indicated that at Time 1, the group mean for log-transformed AUD symptoms was 1.22. The positive linear slope (Time = 3.34) indicated that on average, the slope of AUD symptoms increased over time; however, the significant negative quadratic slope ($Time^2$ = -1.85) suggested a curvilinear slope with rising and leveling off over time (Figure 4). Results of a likelihood ratio chi-square different test comparing the Unconditional Growth Model with random intercept and slope (AIC=1552.5, ICC=.61) to the Null Model (AIC=1590.0, ICC=.54) showed that the Unconditional Model significantly improved model fit [$X^2(4, 377) = 45.54$, $p<.001$].

Next, we constructed the Conditional Growth Model, which adds to the Unconditional Growth Model through the inclusion of intervention and EXT PRS as main effect predictors of change in the outcome over time. We compared conditional models with random intercept only, random slopes only, and a combined model with random intercept and slopes, with the latter providing the best fit for the structure of the data (AIC=1418.9, ICC=.60). In the Conditional Growth Model, the intercept, linear slope, and quadratic slope remained significant, but there was no evidence of significant main effects for intervention group or EXT PRS. Next, the Interaction Model evaluated whether the intercept and slopes varied as a function of the interaction between intervention group, EXT PRS, and time. There was slightly improved model

fit, as indicated by lower AIC, but there was no evidence that trajectories of AUD symptoms varied as a function of two-way or three-way interactions between Time, Time^2, intervention group, and EXT PRS. We proceeded with the addition of covariates to fully evaluate the final model. The effects for intercept and linear and quadratic slope remained significant after the addition of covariates and multiple testing corrections. However, consistent with the alcohol consumption model, no significant main effects for intervention group or EXT PRS were observed, and there remained no evidence to suggest an interaction between intervention group and EXT PRS on the slope of AUD symptoms.

Table 4.

Multilevel Growth Curve Analysis of Alcohol Use Disorder Symptoms from Time 1 to Time 4

	Model				
Level and Variable	Null	Unconditional Growth	Conditional Growth	Interaction Model	Interaction Model with Covariates[1]
Level 1					
Intercept	1.18 (.03)**	1.22 (.03)**	1.27 (.04)**	1.27 (.04)**	1.14 (.12)**
Time		3.34 (.84)**	3.28 (.84)**	3.11 (1.08)**	3.20 (1.09)**
Time^2		-1.85 (.67)**	-1.92(.69)**	-1.11 (.87)**	-1.21 (.88)
Level 2					
Intervention (control=0)			-.11 (.06)	-.09 (.07)	-.07 (.07)
EXT PRS			.02 (.03)	.02 (.04)	.02 (.04)
Interaction components					
Time * EXT PRS				1.30 (1.07)	1.42 (1.07)
Time * Intervention				.70 (1.71)	.75 (1.74)
Time^2 * EXT PRS				1.61 (.87)	1.58 (.87)
Time^2 * Intervention				-1.31 (1.40)	-1.37 (1.41)
EXT PRS * Intervention				-.01 (.07)	-.01 (.07)
Time * EXT PRS * Intervention				-1.75 (1.78)	-1.77 (1.79)
Time^2 * EXT PRS * Intervention				1.82 (1.44)	1.75 (1.44)
Covariates					
Gender (female=1)					-.03 (.09)
LR to Alcohol					.03 (.01)
Propensity Score					-.03 (.46)
Additional Information					
ICC	.54	.61	.60	.61	.59
-2 log likelihood (FIML)	-792.0	-769.2	-700.4	-692.4	-682.4
AIC	1590.0	1552.5	1418.9	1416.8	1422.7
Pseudo R^2 (fixed effects)	0.00	-	-	-	-
Pseudo R^2 (total)	0.54	-	-	-	-
Number of individuals	377	377	346	346	345
Observations	901	901	834	834	832

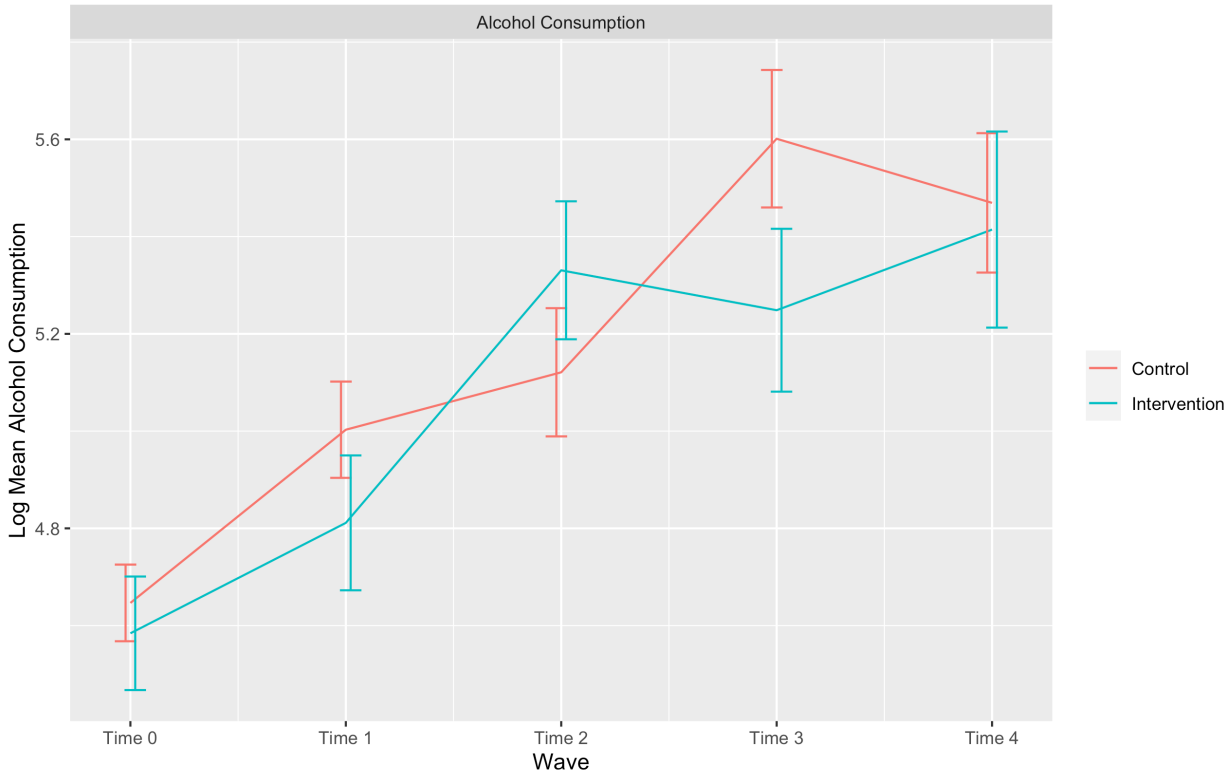
Note: Estimates of unstandardized coefficients are presented for fixed effects. Values in parentheses are standard errors. AUD symptoms was log-transformed to account for skewness and kurtosis. Pseudo R^2 cannot be calculated for models with quadratic growth. The Interaction Model with Covariates also controlled for the effect of Ancestry PCs 1-10. *$p < .05$, **$p < .01$.

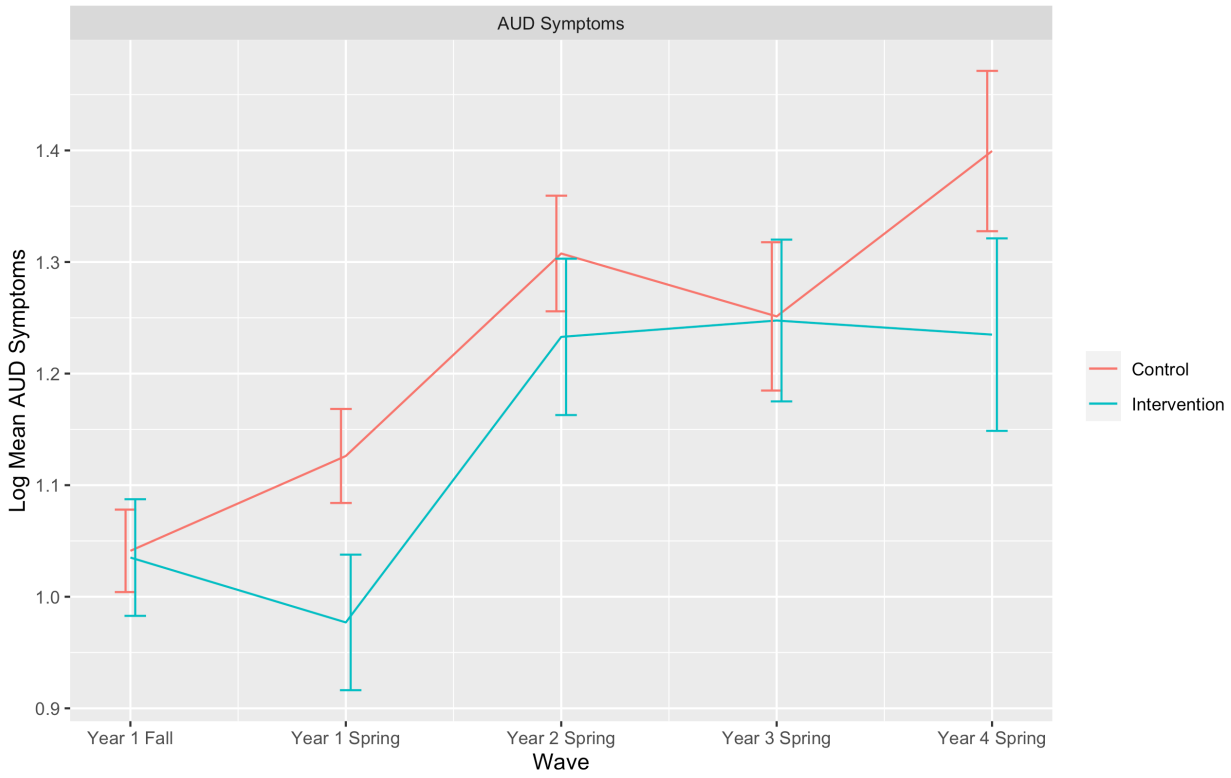
[1] p-values for Level 1, Level 2, and Interaction components were adjusted for multiple testing using the Benjamini–Hochberg procedure in the Interaction Model with Covariates.

Post-hoc Examination of GxI Effects on Proximal Outcomes. Although the primary MLM growth curve analyses described above indicated no evidence an interaction between EXT PRS, intervention group, and time on trajectories of alcohol consumption and AUD symptoms, visual examination of the data suggested that there may have been more proximal, short-term effects of the intervention on alcohol use behaviors. As shown in the second panel of Figure 4, there appeared to be a notable difference in AUD symptoms between intervention and control participants at Time 1.

Figure 4.

Plotted Means and Standard Errors of Alcohol Consumption and AUD Symptoms across Time for Intervention and Control Participants.





To probe these variations further, a hierarchical multiple regression was conducted to test for an interaction between EXT PRS and intervention group on change in AUD symptoms from baseline (Time 0) to the first follow-up assessment (Time 1). Results of these analyses are displayed in Table 5. We observed a significant main effect of intervention on change in AUD symptoms after accounting for covariates (gender, LR to alcohol, propensity score, and ancestry PCs). There was also a significant GxI interaction between EXT PRS and intervention on change in AUD symptoms over and above the main effects and covariates, $\Delta R^2 = .04$, $\beta = .21$, $t(297) = 2.6$, $p = .01$. The nature of this interaction, displayed in Figure 5, was such that individuals with lower EXT PRS reported significantly greater reductions in AUD symptoms in the intervention group compared to those with similar EXT PRS scores in the control condition, $t(297) = -3.44$, p

< .001. Simple slopes analyses identified that the moderation effect was present for individuals with EXT PRS below values of .02. EXT PRS was mean centered at zero, accordingly intervention participants with low to approximately average polygenic risk associated with externalizing behaviors demonstrated significantly greater reduction in AUD symptoms from Time 0 to Time 1 than control participants with similar EXT PRS scores. These results suggest that the intervention may have been differentially effective for individuals based on their genetic risk; however, the direction of this effect was opposite from that which was hypothesized. Prior literature has indicated that individuals with greater genetic risk may respond better to intervention (Bakermans-Kranenburg & van IJzendoorn, 2015; van Ijzendoorn & Bakermans-Kranenburg, 2015), whereas our results indicated that the intervention was more effective for individuals with lower genetic predisposition for alcohol use problems, as measured by EXT PRS. Implications of these findings are explored in the discussion section.

Table 5.

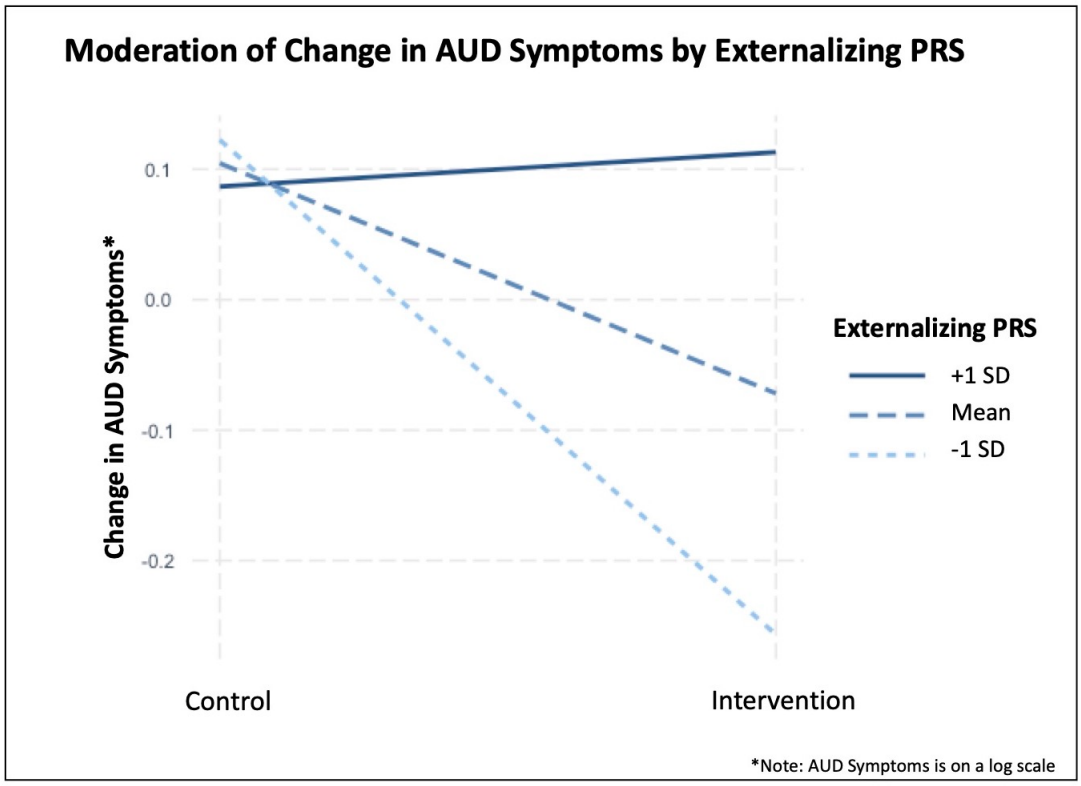
Results of hierarchical multiple regression examining the interactions between Externalizing PRS and Intervention on change in log-transformed AUD symptoms from Time 0 to Time 1

Predictor	*b*	*b* 95% CI	sr^2	sr^2 95% CI
AUD Symptoms				
(Intercept)	0.38**	[0.10, 0.66]		
EXT PRS	-0.02	[-0.11, 0.07]	.00	[-.00, .01]
Intervention Group	-0.17*	[-0.33, -0.01]	.01	[-.01, .04]
Gender (female)	-0.09	[-0.30, 0.13]	.00	[-.01, .01]
LR to alcohol	-0.02	[-0.05, 0.01]	.00	[-.01, .02]
Propensity Score	-0.46	[-1.50, 0.58]	.00	[-.01, .01]
Ancestry PC1	0.27	[-4.10, 4.65]	.00	[-.00, .00]
Ancestry PC2	-1.02	[-5.39, 3.35]	.00	[-.00, .01]
Ancestry PC3	0.39	[-3.48, 4.27]	.00	[-.00, .00]
Ancestry PC4	-0.33	[-4.16, 3.50]	.00	[-.00, .00]
Ancestry PC5	2.16	[-1.64, 5.95]	.00	[-.01, .02]
Ancestry PC6	-0.85	[-5.24, 3.54]	.00	[-.00, .00]
Ancestry PC7	-3.80	[-7.83, 0.23]	.01	[-.01, .03]
Ancestry PC8	3.52	[-0.23, 7.27]	.01	[-.01, .03]
Ancestry PC9	-2.02	[-5.69, 1.66]	.00	[-.01, .02]
Ancestry PC10	0.10	[-4.04, 4.24]	.00	[-.00, .00]
EXT PRS * Intervention	0.21**	[0.05, 0.36]	.02	[-.01, .05]
Model Fit: Multiple R^2 = .087*, 95% CI[.00, .10], ΔR^2 = .038				

Note. A significant *b*-weight indicates the semi-partial correlation is also significant. *b* represents unstandardized regression weights. sr^2 represents the semi-partial correlation squared. * indicates $p < .05$. ** indicates $p < .01$.

Figure 5.

Interaction between externalizing polygenic risk score and intervention group on change in Alcohol Use Disorder symptoms



Aim 2: Results of examining peer deviance and drinking motives as mediators of gene-by-intervention effects on AUD symptoms.

In preparation for Aim 2 analyses, we calculated correlations between all independent, dependent, mediating, and moderating variables, displayed in Table 6. We observed several significant correlations between covariates and outcome variables. In addition, drinking to enhance was significantly correlated with both drinking to cope and peer deviance. We proceeded with the mediated moderation analyses using separate models to examine each mediator in succession, while controlling for significantly correlated mediators in each respective

model. Results of the mediation analyses are displayed in Table 7. Analyses covaried for the effects of gender, LR to alcohol, propensity score, and the first two ancestry PCs. We observed no evidence that changes in peer deviance, drinking to cope, or drinking to enhance mediated the effect of the intervention on AUD symptoms at Time 1 for those at varying levels of genetic risk. However, we did observe significant direct effects of peer deviance and drinking to cope on AUD symptoms among both the intervention and control groups for individuals with lower EXT PRS (1 SD below the mean). The total effect, which is the sum of the indirect and direct effects, was also significant for both peer deviance and drinking to cope; however, in the absence of a significant indirect effect, this result was likely driven by the strength of the direct effect. For drinking to enhance, there was evidence of a significant direct effect on AUD symptoms in intervention participants, but not in control participants. To further examine these variations, we conducted post-hoc moderation analyses to determine if there was an interaction between drinking to enhance and the intervention. Results suggested that while there is a significant positive association between drinking to enhance and Time 1 AUD symptoms [$t(299) = 2.61$, $p<.001$], there was no evidence of an interaction between drinking to enhance and the intervention [$t(299) = -1.35$, $p = .18$]. In summary, drinking to cope, drinking to enhance, and peer deviance significantly influenced change in AUD symptoms; however, they did not explain the variation in intervention effects on AUD symptoms for individuals with high versus low EXT PRS.

Table 6

Means, standard deviations, and correlations of variables included in Mediated Moderation models

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11
1. Δ Peer Deviance	0	4.1											
2. Δ Drinking to Enhance	0	0.56	.16*										
3. Δ Drinking to Cope	0	0.89	.05	.15*									
4. EXT PRS	0	1	.12*	-.03	.03								
5. Intervention	0.33	0.47	-.02	.00	-.07	-.01							
6. Propensity Score	0.22	0.08	-.02	.07	-.03	-.04	.00						
7. Gender (female=1)	0.71	0.46	-.06	-.01	-.00	.04	-.00	.37**					
8. LR to Alcohol	6.06	2.67	-.05	-.01	-.02	-.03	.00	.25**	-.45**				
9. AUD Symptoms (T0)	1.04	0.66	-.01	-.18**	-.07	-.00	-.00	.04	-.03	.12**			
10. AUD Symptoms (T1)	1.07	0.65	.17**	-.01	.08	.08	.11*	-.08	-.07	.05	.50**		
11. Alcohol Consumption (T0)	4.63	1.44	-.00	-.15**	-.06	.01	-.02	-.23**	-.24**	.40**	.40**	.32**	
12. Alcohol Consumption (T1)	4.94	1.47	.26	-.12	-.19	-.09	-.17	-.43	-.59*	.45	.46	.62*	.76**

Note. M and *SD* are used to represent mean and standard deviation, respectively. T0 = Time 0 (baseline assessment). T1 = Time 1 (first follow-up assessment). * indicates $p < .05$. ** indicates $p < .01$.

Table 7

Results of Mediated Moderation Analyses when the Moderator is Set to Low PRS

	Δ Peer Deviance *n =301*			Δ Drinking to Enhance *n = 313*			Δ Drinking to Cope *n = 312*		
Variable	*Estimate*	*95% CI*	*p*	*Estimate*	*95% CI*	*p*	*Estimate*	*95% CI*	*p*
ACME (control)	-.01	[-.11, .04]	.99	-.00	[-.02, .01]	.93	-.00	[-.03, .01]	.99
ACME (intervention)	.00	[-.05, .25]	.98	-.01	[-.05, .00]	.19	.00	[-.03, .05]	.98
ADE (control)	-.16	[-.38, .01]	.03*	-.13	[-.26, .01]	.08	-.16	[-.30, -.02]	.03*
ADE (intervention)	-.16	[-.33, .04]	.03*	-.14	[-.27, .00]	.05*	-.16	[-.30, -.02]	.03*
Total Effect	-0.16	[-.33, .02]	.03*	-.14	[-.27, .00]	.05*	-.16	[-.30, -.02]	.03*
Proportion Mediated (control)	.00	[-1.86, .37]	.99	.00	[-.01, 1.73]	.92	.00	[-.02, -.07]	.99
Proportion Mediated (intervention)	-.00	[-6.50, .72]	.98	.10	[-.04, 5.57]	.22	-.00	[.16, 293.74]	.98
ACME (average)	.00	[-.02, .13]	.98	-.01	[-.03, .00]	.23	.00	[-.02, .02]	.98
ADE (average)	-.16	[-.32, .00]	.03*	-.14	[-.26, .00]	.06	-.16	[-.30, -.02]	.03*
Proportion Mediated (average)	-.00	[-4.61,.11]	.99	.05	[.03, 5.48]	.25	-.00	[-.02, 32.00]	.99

Note. ACME is the Average Causal Mediated Effect (i.e., the indirect effect of mediator on the outcome), ADE is the Average Direct Effect (i.e., the direct effect of mediator on the outcome), and Total Effect represents the sum of the ACME and ADE. The moderator (EXT PRS) was set to 1 SD below the mean for the analyses reported in the table. Estimates and CI were calculated using a bias-corrected and accelerated bootstrap resample of 2,000 simulations. Analyses covaried for the effects of gender, LR to alcohol, propensity score, and the first two ancestry PCs * indicates $p < .05$. ** indicates $p < .01$.

Discussion: Study One

The present study examined the moderating effect of polygenic risk associated with externalizing problems on the effect of a brief alcohol prevention program for college students, with the goal of understanding whether underlying genetic predispositions contribute to differential response to prevention and intervention. In addition, change in peer deviance and drinking motives were explored as potential mediators of gene-by-intervention effects, with the goal of understanding potential pathways through which genetic risk may contribute to differential prevention and intervention effects. The following discussion section will provide a brief review of study results and discussion of findings within the context of the broader literature.

The first aim of this study was to evaluate the underlying genetic risk for externalizing behaviors as a moderator of a brief, web-based alcohol intervention program for college students. Preliminary analyses examined the degree to which the EXT PRS was associated with relevant outcomes in our sample of interest. The EXT PRS performed adequately in the larger Spit for Science EA sample (Ns ranging from 837 to 2967), significantly predicting alcohol consumption, AUD symptoms, peer deviance, and antisocial behavior (Table 2, Figure 3). After controlling for ancestry PCs, the EXT PRS accounted for approximately 1-1.5% of the variance in baseline measures of all phenotypes examined. The variances accounted for in the Spit for Science EA sample are similar to those observed in the Karlsson Linnér et al. (2020) validation of the EXT PRS when applied to target samples. For example, Karlsson Linnér et al. (2020) found that the EXT PRS accounted for 2.28% of the variance in AUD symptoms and 2.52% of the variance in antisocial personality disorder symptoms in the Collaborative Studies on the Genetics of

Alcoholism (COGA) Study sample. Although the values are higher in the Karlsson Linnér et al. (2020) paper compared to the present study, the COGA study sample is both larger and includes families densely affected by AUD, thus it is expected that the variance accounted for by EXT PRS may be slightly higher than those observed in a population-based study like Spit for Science. Indeed, the estimate of variance in AUD symptoms accounted for by EXT PRS in Spit for Science exceeded the effect observed in the population based sample Add Health (ΔR^2 = 0.66%) in the Karlsson Linnér et al. (2020) paper.

Despite the significant effect of the EXT PRS in the broader Spit for Science sample, the primary growth curve analyses in the S4S-LR intervention sample showed no differences in AUD symptoms or alcohol consumption based on intervention group, EXT PRS or their interaction on trajectories of these outcomes across time. The second aim, which tested peer deviance and drinking motives as mediators of GxI effects, also produced non-significant results. The only significant finding in the Spit for Science sample resulted from post-hoc analyses focused on short-term GxI effects on AUD symptoms. We observed a significant interaction between EXT PRS and the intervention, such that individuals with lower PRS in the intervention group reported greater reduction in AUD symptoms at the first follow-up compared to individuals in the control group and those with higher PRS. These results suggest that short-term effects of a web-based alcohol intervention for college students may vary for individuals with different levels of underlying genetic risk for externalizing problems; however, the majority of the analyses, including those focused on longitudinal GxI effects and mediators of GxI effects, were not supported by the study findings. The implications of this single significant finding, in the context of predominantly null results, are discussed in further detail in the sections to follow.

It is notable that the direction of the significant moderation effect (individuals with lower PRS had improved intervention response) was opposite from that which was hypothesized (individuals with higher PRS will respond better to the intervention). The initial hypothesis was based on the differential susceptibility hypothesis, which suggests that some individuals are more sensitive to both promotive and harmful environments based on their genotype (Belsky & Pluess, 2009). The theory emerged from Belsky and colleagues as a response to the limitations of diathesis-stress, which stipulates that certain biological factors place an individual at elevated risk for disorder in the presence of environmental stressors. Differential susceptibility extends diathesis-stress to include enhanced response to positive environments, which Belsky and Pluess posit provides evolutionary advantages that allow these variants to persist in the population across time (Belsky & Pluess, 2009; Pluess, 2017; Pluess & Belsky, 2013). Prior GxI studies that examined the differential susceptibility hypothesis were predicated on the belief that certain genetic factors influence susceptibility to environmental exposure, otherwise referred to as "plasticity" genes. Studies that sought to test the differential susceptibility hypothesis using GxI focused on certain genetic variants, typically candidate genes, which were believed to play a role in the sensitivity to the environment as well as the outcome of interest. For example, Cleveland et al. (2015) focused their analyses on the dopamine receptor D4 gene (*DRD4*) due to literature indicating that *DRD4* may not only be associated with increased risk for negative outcomes (Daurio et al., 2020; McGeary, 2009b; Ptáček et al., 2011), but also increased receptiveness to environmental changes due to the role of dopamine neurons in transmitting signals related to both rewarding and aversive events (Bromberg-Martin et al., 2010). Cleveland et al. (2015), as well as a number of other alcohol-related GxI studies (Beach et al., 2010; Brody et al., 2013,

2014, 2015; Ewing et al., 2009), observed findings that suggested the 7-repeat allele predicted poorer outcomes under control conditions and improved outcomes under intervention conditions (Bakermans-Kranenburg & van IJzendoorn, 2015; van Ijzendoorn & Bakermans-Kranenburg, 2015). In contrast, the present study findings suggest that individuals with greater polygenic risk for externalizing behaviors are less likely to benefit from a brief alcohol intervention for college students. The shape of this interaction aligns more with the vantage resistance model, or a diminished ability to benefit from promotive environmental conditions (i.e., intervention) than those at lower risk (Pluess & Belsky, 2013).

There are a few possible reasons why the present study's findings did not align with previous GxI research supporting the differential susceptibility hypothesis. First, most prior GxI studies used candidate gene methods for the measurement of genotype, which involve the study of a marker or markers located in a single gene selected a priori for its hypothesized role in a given phenotype. Subsequent well-powered GWAS have not supported the hypothesized role of most candidate genes, suggesting that many significant candidate gene findings may have been the result of false positives or spurious effects (Auwera et al., 2018; Border et al., 2019; Johnson et al., 2017). In the case of *DRD4*, a recent meta-analysis of the role of *DRD4* in alcohol-related outcomes identified mixed results, with some evidence that the 7-repeat allele was associated with increased drinking days, binge drinking days, and AUD symptom severity but no differences in typical drinks per day and maximum drinks per occasion. However, the reliability of these findings was limited by small sample sizes and the inability to account for population stratification in most studies due to the lack of ancestry data, both of which increase the likelihood of Type I error. Thus, it is likely that candidate GxI studies encountered similar

problems observed in the broader candidate gene study literature, including publication bias, false positives, and limitations of power due to the very small effect sizes of single variants examined in small samples (Border et al., 2019; Duncan & Keller, 2011; Neale et al., 2020).

In contrast to previous candidate gene work, the present study used a polygenic score to index underlying genetic risk, which provides an estimate of the accumulation of thousands of very small genetic effects that contribute to complex behaviors. Polygenic scores serve as indices of aggregate genetic vulnerability, in this case indexing risk for externalizing behaviors across the genome. Although the polygenic score for externalizing behavior likely includes some markers that may play a role in sensitivity to environment, the impact of those individual variants cannot be discerned due to the cumulative nature of polygenic scores. Furthermore, single variants explain very little variance in any given complex behavior, and extremely large samples would be needed to detect the effect of individual variants (Timpson et al., 2018). In order to test for the impact of sensitivity to the environment using a polygenic score, one would need to measure sensitivity to the environment and use that phenotype to create a polygenic score. Keers et al. (2016) pursued this approach by creating a polygenic score for environmental sensitivity based on a measure of within twin-pair differences in emotional problems. They found that polygenic risk for environmental sensitivity moderated the effect of a cognitive behavioral therapy (CBT) intervention on childhood anxiety symptoms, such that individuals with higher polygenic scores for environmental sensitivity responded significantly better to individual CBT compared to parent-led or group CBT. The environmental sensitivity polygenic score has also been used to assess gene-by-intervention effects of the Family Check-Up on childhood internalizing psychopathology (Lemery-Chalfant et al., 2018). Lemery-Chalfant et al. (2018)

found that children with higher polygenic scores assigned to the intervention reported approximately 5-10 fewer total internalizing symptoms than those with similar genetic risk assigned to the control condition. The effect of polygenic risk for environmental sensitivity has yet to be tested in studies of alcohol and substance use outcomes, this future research is warranted to explore this area of research further. However, it is important to note that research studying the genetics of environmental sensitivity asks a different question than studies of underlying genetic risk for externalizing behaviors, as examined in the present study. Accordingly, patterns of findings for studies of polygenetic risk associated with given outcome (i.e., alcohol consumption, externalizing behaviors) may differ from those which study environmental sensitivity.

Another question that arises from the GxI analyses in the present study is the following: why were the GxI effects significant in the immediate post-intervention period (Time 0 to Time 1), but not in the primary growth curve analyses assessing trajectories of AUD symptoms from Time 1 to Time 4? One possible explanation for these findings is that the intervention effects were not sufficiently robust to generate discernible differences in the slope of alcohol consumption and AUD symptoms across the four-year follow-up period. The intervention involved only 4-weeks of one-hour videos delivered online, which may be insufficient to create lasting effects on college student drinking given the wide array of other influential individual and environmental factors contributing to drinking outcomes. Another possible explanation for the null results for the growth curve analyses involves statistical power. Power was calculated using an adapted version of the *mlm_test* function from the *paramtest* package in *R.* Simulated values of the dependent variable were generated using the model specified in *mlm_test*, expanded to

include fixed effects for all predictors of interest in this analysis using coefficients for each fixed effect drawn from the growth curve model. The values of the coefficients used in the simulation pipeline are displayed in Table 4, under the Interaction Model. After simulating the dependent variable as a function of the predictors, the model was fit in the simulated data and the p-value was saved. This pipeline was repeated 1000 times. Power is calculated as the proportion of p-values below .05. Results indicated that the model achieved approximately 74% power to detect an effect in a sample of 344, meaning 74% of the time, models with these indicators are sufficiently powered to detect an effect if there is one in the population. At 74% power, the growth curve models were slightly underpowered compared to the preferred standard of 80% power to detect an effect. By comparison, according to post-hoc power analysis conducted using G*Power, the hierarchical multiple regression examining GxI effects on change in AUD symptoms from Time 0 to Time 1 achieved 93% power to detect an interaction effect given the small effect size (f^2 = .038) observed in the model (Faul et al., 2007). Thus, it is possible that the limitations of analytical power in this sample played a role in the null effects observed in the growth curve analyses, whereas the hierarchical multiple regression was adequately powered to discern an effect.

It is also worth considering the possibility that the single significant GxI interaction effect observed in this study may be a statistical artifact, or a spurious finding resulting from measurement error in psychometric variables. Statistical artifacts have long been a topic of discussion in GxE research due to the psychometric complexity of measuring psychiatric phenomena in the absence of objective measures, and the preponderance of scale transformations (e.g., diagnostic cutoffs, sum scores, etc.) used to address distributional problems in psychiatric

measures (Eaves et al., 1977; Jinks & Fulker, 1970; Mather & Jinks, 1982). As Eaves and Verhulst (2014) stated, "you can generate almost any interaction you want by changing the scale of measurement." Scale transformations can emphasize or exclude certain points of the scale, and in turn alter the conclusions derived from analyses. Indeed, simulations of GxE interaction studies have shown that the dichotomization of continuous measures (Eaves, 2006) and the use of sum scores (Eaves, 2017; Schwabe & van den Berg, 2014) can result in biased estimates of GxE effects. In the present study, there was a significant GxI effect on AUD symptoms, a log-transformed symptom sum score. Although log-transformations can reduce bias in GxE analyses, problems may still persist (Eaves & Verhulst, 2014). Therefore, it is possible that the observed GxI effects may be a statistical artifact resulting from scale transformation, or a biased overestimate the effect size. Recommendations for differentiating true versus artifact interactions include 1) categorizing each symptom and fitting models with logistic regression, 2) designing and using improved psychometric measures, 3) integrating item response theory, 4) using non-transformed variables to probe effects in sensitivity analyses, and/or 5) using a variety of statistical models to explore potential differences as a function of genotype (Domingue et al., 2020; Eaves & Verhulst, 2014). Although the problems related to scale artifact in GxE research discussed in this section are not unique to this study, it is important to consider the potential implications in the context of the predominantly null pattern of findings across the dissertation analyses. Accordingly, it is recommended that the GxI results in this study be interpreted with substantial caution, given the challenges of discerning true versus artifact GxI results.

Results of the mediation analyses focused on peer deviance, drinking to cope, and drinking to enhance were also non-significant. These mediators were selected based on prior

evidence that they significantly influence alcohol-related outcomes in youth, have a genetic component to their etiology, and are modifiable risk factors that can be targeted by interventions. Twin studies have demonstrated that genetic risk for externalizing problems is amplified under conditions of high peer deviance (Harden et al., 2008). GxE research also indicates that genetic liability for substance use and peer group deviance may be correlated, such that individuals carrying genetic risk for substance use problems may self-select into higher risk peer groups (Gillespie et al., 2009). Similarly, drinking motives are heritable (specifically drinking to cope with negative emotions, drinking to enhance positive feelings) and there is evidence that they mediate the effect of family history on alcoholism (Agrawal et al., 2008; Beseler et al., 2008). Despite the theoretical and empirically supported role of these factors in drinking related outcomes, we observed no mediating effect of these mechanisms in our analyses. One possible explanation is that that analyses may have had insufficient power to detect a mediating effect. Very few previous GxI studies of alcohol and substance use outcomes have examined mediators, and those that did used candidate gene methods to measure genotype (Brody et al., 2014, 2015). The large mediation effect sizes observed in previous studies are likely related to overestimates of the main effect of the genotype, and thus a more conservative estimate would be to assume small to very small effects for mediation analyses using polygenic scores. Based on estimates of sample sizes required for mediation analyses, the analytic samples for the mediation analyses in the present study (n ranging from 301-313, see Table 7) were adequately powered to detect halfway (d = .26), medium (d = .39), and large (d = .59) effects, but insufficiently powered to detect small (d = .14) effects (Fritz & MacKinnon, 2007). In order to detect small mediation effects, a sample size of 462 individuals would be required to achieve 80% power according to

the recommendations from Fritz & MacKinnon (2007). Post-hoc analyses also indicated that there were no main effects of the intervention on change in peer deviance, drinking to cope, or drinking to enhance, suggesting that changes in these factors were not a result of the intervention.

Therefore, additional research is needed to obtain larger, more well-powered samples, and explore alternative factors that may have contributed to changes in AUD symptoms observed in individuals with low EXT PRS in the intervention group. Mechanisms of action for brief alcohol interventions remain unclear (Gaume et al., 2014; Magill et al., 2015), though there is some evidence that changes in perceptions of normative drinking (Carey et al., 2010), increased understanding of values-discrepant behavior (Barnett et al., 2010; McNally et al., 2005; Miller & Rollnick, 2002a), and increased use of protective behavioral strategies may contribute to the beneficial effects of alcohol interventions (Pearson, 2013; Prince et al., 2013; Walters et al., 2009). For example, perhaps the present study's intervention led to increased use of protective strategies among individuals with lower EXT PRS, such as alternating alcoholic beverages with water, setting a drink limit, eating before consuming alcohol, and drinking at a slower pace (Bravo et al., 2017; Pearson, 2013; Prince et al., 2013). Although these factors are important to consider as potential mediators of intervention effects, they were not assessed as part of the present intervention study or the Spit for Science data collection efforts. Future research would benefit from increased assessment of potential mediators of intervention effects, so that questions of how and why interventions are differentially effective can be explored.

The present study has a number of strengths that distinguish it from the extant GxI literature. First, most previous GxI studies of alcohol and other substance use outcomes relied

on candidate gene methods (Neale et al., 2020). Although the early reliance on candidate genes in GxI literature parallels the progression of genotypic methods in the broader field of genetics research, the adoption of newer methods, such as polygenic scores has been slower to progress in GxI research. The present study advances GxI research by incorporating polygenic scores into the analyses, joining a handful of prior studies that have done the same (Kuo et al., 2019; Musci et al., 2015, 2018). Second, in order to calculate an adequately-powered polygenic score this study used the largest GWAS of externalizing behaviors to date (Karlsson Linnér et al., 2020). Although there are a number of well-powered GWAS of alcohol use and AUD (Kranzler et al., 2019; Liu et al., 2019; Sanchez-Roige et al., 2019; Walters et al., 2018; Zhou et al., 2020), the externalizing GWAS discovery sample was selected due to evidence that genetic risk for externalizing behaviors is more likely to manifest earlier in development than alcohol-specific risk (Kendler & Myers, 2014; Meyers et al., 2014). Furthermore, the use of the EXT PRS is consistent with literature that suggests genetic risk for alcohol and substance use behaviors are better explained by shared genetic liability for externalizing behaviors (Hicks et al., 2004; Kendler et al., 2003; Krueger et al., 2002). Fourth, this study used propensity score matching to approximate the effects of random assignment and mitigate the effects of selection bias in a non-randomized sample (Austin, 2011). Although randomized-controlled trials are considered the gold standard study design for exploring causal effects of an intervention on outcomes, the application of propensity score matching in this study successfully resulted in the creation of a comparison group with no significant differences from the intervention group across a wide range of baseline characteristics. While causal interpretations of study findings should still be approached with caution, the use of propensity score matching represents an opportunity to

substantially expand the realm of available datasets in which to explore GxI effects. Finally, most prior GxI studies examined adolescent samples and very few attempted to explore mediators of GxI effects. Emerging adulthood is a critical period for the development of alcohol use behaviors (Skidmore et al., 2016; Sussman & Arnett, 2014), yet in the context of GxI research remains understudied (Ewing et al., 2009; Neale et al., 2020). The present study capitalized on a large study of genetic and environmental influences on substance use and emotional health in college students, expanding the scope of prior GxI literature into important populations of interest. Regarding the importance of mediators, although we observed no evidence of mediation via changes in peer deviance or drinking motives, the present study emphasized the need to better understand why interventions may be differentially effective for individuals with varying levels of genetic risk.

The results of this study should also be interpreted within the context of several limitations. First, the sample size for the current study (N=483) was relatively small and may have contributed limitations in statistical power to detect effects. As indicated previously, the sample was underpowered to detect a 3-way PRS * Intervention * Time interaction effect in the growth curve analyses, but adequately powered (93%) to detect the observed GxI effect of f^2 = .038 in the hierarchical multiple regression analyses. However, this significant GxI effects may also be a statistical artifact resulting from scale transformation, or a biased overestimate the effect size. Larger samples are needed to replicate the present study's findings and explore some of additional mediators of GxI effects. Recommendations for sample size requirements and strategies to attain larger samples are provided in the discussion of future directions. Second, the study was limited to individuals of European ancestry due to small numbers of individuals of

other racial/ethnic groups in the intervention sample. Current best practices for genetics research recommend analyzing ancestry groups separately due to differences in allele frequency across different populations that can bias results (Peterson et al., 2019). This is a significant limitation of the present study, as underrepresentation of diverse individuals inhibits the equitable application of GxI research findings (Martin et al., 2019; Popejoy & Fullerton, 2016). A combination of computational intricacies, such as those present in this dissertation study, and broader historical and systemic issues have negatively impacted representation of diverse ancestry in genetics research for decades (Bates et al., 2005; Dick et al., 2017; Furr, 2002; Tambor et al., 2002). Unfortunately, genetic findings observed in one ancestry population often do not replicate in other ancestry groups due to differences in genetic architecture, such as linkage disequilibrium and allele frequency (Sirugo et al., 2019). Without concerted effort to increase diverse representation in genomic research studies, disparities in the utility of genomic research findings will continue to expand (Martin et al., 2019). Finally, although polygenic scores offer many benefits over other methods for integrating genetics into prevention/intervention research, they also have some limitations. As measures of cumulative genetic risk, polygenic scores provide little information about individual SNP-level effects that may increase understanding of the biological processes that might explain gene-environment interactions. Polygenic scores also only capture genetic risk associated with common variants; as a result, rare variants may contribute additional variance not captured by polygenic scores (Crouch & Bodmer, 2020; Young, 2019). Although these limitations may inhibit the ability of GxI research to spur new ideas for molecular genetic research, polygenic scores remain the most accessible and cost-effective means to integrate genetics into prevention/intervention research.

Despite these limitations, there are a number of opportunities to advance the future directions of this work. First, it may be useful to conduct additional analyses using different polygenic scores. Although the externalizing polygenic score has both theoretical and empirical support for the association with alcohol-related behaviors, the degree to which genetic risk for externalizing problems influences alcohol-related outcomes may shift across environmental and developmental contexts. Twin studies have shown that under certain environmental conditions, such as high peer deviance and low parental monitoring, the influence of genetic risk is stronger than in more protective environments (Dick et al., 2007; Harden et al., 2008). The subjects included in the externalizing GWAS discovery sample were also older on average than the college students in the present study, which may impact the predictive validity of the polygenic scores. Accordingly, it will be important to examine the way that GWAS discovery sample characteristics may influence findings in the target sample, as well as explore the way polygenic scores for different outcomes (e.g., alcohol consumption, AUD diagnosis, etc.) may affect results. Second, as the present study identified no significant mediators of GxI effects, one viable future direction for this work is to identify and test new mediators. Active ingredients in brief interventions, such as increasing protective behavioral strategies and correcting perceptions of normative peer use, are both viable options for potential psychosocial mediators of GxI effects (Barnett et al., 2010; Lewis & Neighbors, 2006; Magill et al., 2015; Prince et al., 2013).

Given the limitations of statistical power and racial/ethnic diversity in the present study, exploring these research questions in larger, more diverse study samples would enhance the generalizability, applicability, and reliability of these findings. Larger sample sizes can increase the statistical power to detect small GxI effects; however, effect size estimates for GxI studies

using polygenic scores are sparse in the existing literature. In a phenomenon referred to as the "file drawer problem," significant findings are also more likely to be published than non-significant findings (Rosenthal, 1979); therefore, the effect sizes that are published may represent inflated estimates of effects. A systematic review of polygenic GxE studies of tobacco, alcohol, and cannabis indicated that effect sizes are rarely reported, further complicating the procedure for identifying appropriate effect size estimates from which to determine sample size needs (Pasman et al., 2019). The authors of the review paper recommended the use of conservative estimates for GxE effect sizes (e.g., $f^2 = .005$) to ensure studies are sufficiently powered. Assuming an effect size of $f^2 = .005$, with an average of 15 predictors in a hierarchical multiple regression, a sample size of 1572 would be required to achieve 80% power. The present study's sample (N=483) was adequately powered to discern GxI effects of $f^2 = .016$ or greater but was not sufficiently powered to detect more conservative estimates of GxI effect sizes.

Estimated sample size requirements for GxI analyses may exceed those typically seen in intervention research, thus researchers are encouraged to consider new and innovative strategies to attain larger, genetically informed intervention samples. Partnering with existing large-scale, diverse randomized-controlled trials to retrospectively collect DNA samples from study participants may be a useful, and potentially cost-effective approach to advancing this research. This approach may also have the added benefit of increasing representation of samples across the developmental spectrum, to include individuals from early prevention studies in youth to treatment efficacy trials in adulthood. All clinical trials in the United States are required to register on ClinicalTrials.gov, an online database of ongoing clinical trials in the United States, which includes recruitment information, sample sizes, and results (Anderson et al., 2015). The

ClinicalTrials.gov database provides an invaluable resource for identifying potential large, diverse, randomized-controlled trials that may be appropriate for collaborative GxI research. A brief search of ClinicalTrials.gov in April of 2021 returned 185 completed and 165 active, ongoing studies of alcohol use disorder. Filtering these studies by intervention type (i.e., behavioral versus pharmacological), the creation of rankings based on sample size and demographics may then help identify high-priority samples. After identifying prioritized studies, the next step would be to generate a proposal to partner with the intervention researchers to fund collection of DNA samples from their subjects, or access genotypic data if already available. If the intervention study sample sizes are still too small (i.e., fewer than 1500 participants), studies with similar types of interventions could be combined and analyzed together in an approach similar to the method used in collaborative genomics consortia (e.g., Psychiatric Genomics Consortium). Coordination of a collaborative effort such as the one described would likely benefit from the establishment of a consortium, or partnership with an existing genomic consortium, to create an organized and inclusive effort to increase statistical power in GxI studies and expand the applicability of GxI findings to individuals of all backgrounds.

Research Aims and Hypotheses: Study Two

Guided by the extant literature and relevant theoretical models, the proposed research has the following aims and hypotheses:

3. The first aim was to examine whether polygenic risk associated with externalizing behaviors moderates the effectiveness of a multi-component adolescent prevention program to reduce externalizing behavior in the Project Alliance sample.
    a. Informed by the differential susceptibility hypothesis, I hypothesize that individuals in the intervention with higher polygenic risk will show lower rates of alcohol use and problems in emerging adulthood than those with lower polygenic risk and controls.
4. The second aim is to examine whether peer deviance mediates changes in young adult alcohol use behaviors for those at greater genetic risk in the Project Alliance intervention sample.
    a. I expect that the intervention will lead to lower peer deviance among those with higher genetic risk, which will partially account for lower alcohol use and problems in intervention participants. Drinking motives were not measured in the Project Alliance sample, and thus cannot be examined as a mediator of GxI effects.

Methods: Study Two

Project Alliance Sample

Participants. Project Alliance (PAL) is a longitudinal study of students recruited in 6th grade and randomized to participate in an intervention aimed at preventing substance use and deviant behavior (Dishion et al., 2003). Participants (N=998; 47.3% female) and their families were recruited from three public middle schools in a metropolitan community in the Pacific Northwest region of the United States. The participation rate was 90%. The specific schools within the community were close in proximity and the neighborhoods had elevated rates of arrest. Of the students recruited for the study, one half of the participants were randomized to the intervention condition (n = 500), which took place in middle school. All participants were followed longitudinally with assessments approximately annually from grades 6-12 (ages 11-12 to 18-19) and in early adulthood at ages 22-23, 23-24 and 26-27. Retention rates across all waves were very high, ranging from 80.2% to 85.8%. At wave 10 (age 26-27), PAL participants were invited to provide a DNA sample in order to study gene-environment interplay and the influence of family-centered prevention on genetic and environmental risk. DNA saliva samples were collected using Oragene kits and extracted according to standard procedures. Participation in the DNA component was high among PAL participants who were still active at wave 9 (85% of 998). A total of 634 PAL participants provided DNA samples, of which 311 are intervention participants and 323 are controls. The racial/ethnic composition of this subset is 43.2% European American and 30.6% African American. Study procedures were approved by the Institutional Review Board at the University of Oregon.

Prevention program and procedure. The goal of the PAL intervention protocol was to improve family management, address youth adjustment problems, and reduce substance use.

The intervention protocol incorporated multiple components, ranging from universal, selected, and indicated levels of intervention. At the universal level, all intervention participants and their families were offered access to a Family Resource Center established in their school and staffed by parent consultants. The parent consultants offered brief in-person and telephone consultations, reports on their child's behavior in school, and access to a library of parenting materials. Intervention participants were assigned to the same homeroom classes in 7th grade, through which parent consultants led six lessons for the students based on the Life Skills Training program (Botvin et al., 1990) on topics such as positive peer groups and stress and anger.

At the selected level, families were offered the Family Check-Up (FCU), a strengths-based, family-centered intervention based on the principles of motivational interviewing, which is available for training and certification via Arizona State University's REACH Institute (Dishion et al., 2003; Dishion & Stormshak, 2007; Miller & Rollnick, 2002b). The FCU was available to all intervention participants, but families identified as high-risk based on teacher report were contacted directly and offered the FCU in 7th and 8th grade. Among families in the intervention condition, 23% agreed to take part in the FCU at least once. The FCU is comprised of an initial interview in the first session, followed by a family assessment in the second session. In the third session, participants are provided with feedback and guided through a discussion about appropriate follow-up intervention services using a motivational interviewing style. The follow-up intervention services comprised the indicated component of the PAL intervention protocol. These services included evidence-based interventions such as family behavior therapy, multisystemic family therapy, and parent group interventions. Assessments were collected at

each wave of the study from intervention and control participants. Youth were provided with $20 in compensation for each assessment they completed.

Measures. Processed and cleaned genotypic data as well as longitudinal phenotypic data from the PAL surveys were used for analyses.

Alcohol use and dependence. Alcohol use was assessed across all waves using items developed for the PAL study. Participants were asked to report lifetime alcohol use, age of initiation, frequency and quantity of typical use for several different types of alcohol (beer, wine, hard liquor). These items were recoded into semi-continuous measures of monthly frequency and quantity of alcohol use, then multiplied (frequency*quantity) to create a measure of drinks consumed per month. This approach has been validated for use in other studies with the PAL sample (Connell et al., 2007, 2012). Lifetime alcohol dependence (AD) was measured at wave 7 (age 22-23) using the Composite International Diagnostic Interview (Kessler & Ustün, 2004; World Health Organization, 1997). Items were based on DSM-IV criteria for alcohol dependence. An AD symptom sum score was created for each participant by summing the AD items endorsed.

Peer deviance and antisocial behavior. PAL participants completed the Peer Network Deviant Behavior Scale, a 22-item scale adapted for use with young adults from portions of the Community Action for Successful Youth Survey (Metzler et al., 1998). The measure assessed whether a young adult's peer network engaged in problem behaviors (e.g., stealing, fighting, substance use). The items were used to establish measures of deviant peer association at waves 2-4. At wave 6, participants were also invited to take part in an observational study of peer interactions with a self-nominated friend of the same sex. The methods for peer interaction task and coding are described in detail elsewhere (Piehler & Dishion, 2015). Briefly, trained research

assistants coded videotapes of participant behaviors during as series of peer interaction tasks. Derived measures included the duration of "deviant" talk and frequency of rule breaking. Participants also answered questionnaires assessing the frequency with which their peers engaged in antisocial behavior.

Covariates. Covariates include age, sex, teacher report of child risk behavior, and ancestry principal components. Teacher report of risk behavior (TRISK) is a 16-item measure in which teachers rated each child on their 6th Grade roster on the frequency with which they engaged in youth problem behaviors, such as aggression, oppositionality, and problems with peers (Soberman, 1994). Items were averaged and standardized within classroom. The TRISK score is a baseline measure of risk for externalizing problems in the PAL sample.

Genotyping. DNA samples were sent to the Rutgers University Cell and DNA Repository where they were genotyped on the Affymetrix Axiom BioBank Array Version 2. Genotypes were imputed using 1000 Genomes (Phase 3 reference panel; 1000 Genomes Project Consortium, 2015) using SHAPEIT2 and IMPUTE2. Quality control procedures included removal of: 1) palindromic SNPs, i.e. those with ambiguous directions (A/T or C/G), 2) SNPs with genotyping rate <0.95, 3) SNPs that failed tests of Hardy-Weinberg equilibrium (HWE; $p<10^{-6}$), and 4) SNPs with minor allele frequency less than .01. A total of 2,067,148 SNPs passed quality control and data cleaning thresholds and were included in the analyses.

Data Analysis Plan: Sample Two

All analyses were conducted using R, a flexible statistical computing program with several available methods for handling missing data (Kabacoff, 2011). R's functionality is expanded through various packages built to run advanced statistical techniques described below.

Data preparation. All phenotypic variables were examined for normality. Log transformations after adding one were computed when appropriate to reduce the effects of non-normality. Due to the longitudinal nature of the study, some missing data was expected. However, retention rates across all waves were very high, ranging from 80.2% to 85.8%. For the growth curve models, individuals missing greater than one time point were excluded list-wise. Complete data was required for mediation analyses.

Creation of polygenic risk scores (PRS). For the EA subsample in PAL, the method for polygenic score calculation was consistent with the method described for the S4S sample. Using summary statistics from a GWAS of externalizing behaviors in about 1.5 million subjects (Karlsson Linnér et al., 2020), polygenic scores were derived using PRS-CS and the linkage disequilibrium (LD) patterns observed in the 1000 Genome Phase 3 European Ancestry reference panel (The 1000 Genomes Project Consortium, 2015). Please see the section on creation of polygenic risk scores in Study 1 Methods for additional details on externalizing polygenic risk scores (EXT PRS) calculation for EA participants for PAL.

For the AA subsample, additional steps were taken to minimize the potential of bias due to population stratification. Population stratification refers to systematic variations in allele frequency among different geographical ancestry groups. When population stratification is not accounted for in genetic analyses there is an increased risk of false-positive results (Hellwege et al., 2017). The multiethnic polygenic risk score approach (MultiPRS) combines data from large European samples with data from a smaller, ancestry-matched sample to improve risk prediction accuracy in non-European populations (Márquez-Luna et al., 2017). To calculate the MultiPRS, we conducted a GWAS of externalizing behaviors in the African American subsample of the Collaborative Study of the Genetics of Alcoholism. Using summary statistics from this ancestry-

matched GWAS, we calculated polygenic scores in the target sample and combined them with PRS-CS scores derived using the EA Externalizing GWAS summary statistics (Karlsson Linnér et al., 2020). The MultiPRS is the saved predicted value of a linear combination of the COGA PRS and the EA Externalizing GWAS PRS for African ancestry participants in PAL. This method leverages the large sample size used in the Karlsson Linnér paper while also accommodating differences in allele frequency through the incorporation of ancestry-matched genotypic data in the COGA sample.

Principal Components Analysis (PCA). Ancestry principal components (PCs) account for variation in allele frequency across different population structures and are included in analyses to reduce confounding. Principal Components Analysis (PCA) was conducted using EIGENSOFT and SmartPCA with 1000 Genome Project phase 3 reference panel (Patterson et al., 2006; Price et al., 2006). Regions of high LD were excluded using PLINK 2.0, so as to ensure that all SNPs were relatively independent.

Evaluation of Externalizing PRS in PAL. To evaluate the association between EXT PRS and relevant outcomes (alcohol consumption, AD symptoms, peer deviance, and deviant behavior), we conducted hierarchical multiple regression estimating the effect of the EXT PRS on related phenotypes (alcohol, peer deviance, etc.) over and above the effect of the ancestry PCs and sex. The results of these analyses demonstrated the degree to which the EXT PRS is associated with expected phenotypes in the target sample. Analyses were conducted separately for the EA and AA participants and *p*-values were adjusted for multiple testing using the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).

Aim 1: Tests of GxI Effects on Alcohol Consumptions and AD Symptoms. The primary outcomes of interest were alcohol consumption and AD symptoms. Alcohol

consumption was measured at multiple time points; thus, it was appropriate to use growth curve analyses within a multilevel framework to estimate the effects of EXT PRS, intervention group, and their interaction on changes in alcohol consumption across time. However, AD symptoms were assessed at only one time point in emerging adulthood, wave 7 (age 18-19), and therefore, could not be modeled with growth curve analyses. As an alternative, hierarchical multiple regression was used to estimate GxI effects on AD symptoms in the PAL sample. Each of these analytic approaches is described in further detail below. All analyses were conducted separately for EA and AA participants.

Alcohol consumption. Alcohol consumption was measured across multiple time points in the PAL sample. As the focus of this study was on the developmental period of emerging adulthood, we modeled linear and quadratic growth curves within a multilevel framework using four timepoints in late adolescence to emerging adulthood (ages 16-17, 18-19, 22-23 and 23-24). The multilevel framework accommodated the nested structure of repeated measures within individuals, while also allowing for between-individual and between-group variation in intercept and slope of the dependent variable. We used a model building approach to compare the model fit between the unconditional (null) model, unconditional growth model, and conditional growth model with time-invariant covariates (ancestry PCs, sex, and TRISK). Analyses were conducted using the *lme4* package for R, with maximum likelihood estimation. Model fit was evaluated using AIC, intraclass correlation (ICC), and pseudo R^2.

Unconditional models. First, we constructed an unconditional null model to provide an estimate of the within-person (Level 1) and between-person (Level 2) variance components. The unconditional null model serves as a base model with no predictors to test whether model fit improves with the addition of Level-2 effects in subsequent models. Next, an unconditional

growth model was fit. The unconditional growth model estimates the trajectory of the outcome across four time points, with Time as a predictor at Level 1. As there were no significant differences between intervention and control groups at baseline, Time was centered at first follow-up (Time 1), with each successive follow-up coded to account for approximately equal time between follow-up assessments (Time 0 = 0, Time 1 = 2.83, Time 2 = 3.92, Time 3 = 8.58). Linear and curvilinear effects for time were tested, with the quadratic equation providing a better fit for the data. A comparison of model fit between the unconditional null model and the unconditional growth model confirmed that there was sufficient individual variability to warrant advancing to conditional models, which include predictors to estimate variation in intercept and/or slope.

Conditional models. The first conditional growth model included Time as the Level 1 variable, and intervention group and EXT PRS as Level 2 variables. Fixed and random effects for slope and intercept were tested, with random slope and intercept providing the best fit for the data. In the second conditional model, 2-way and 3-way interaction terms were added to examine the degree to which interactions between Time, EXT PRS, and intervention group contributed to variation in the outcomes. In the final model, ancestry PCs, TRISK, and sex were added as covariates at Level 2 to account for their potential impact on the resulting models. We applied within sample (EA and AA) corrections for multiple testing to the *p*-values for the Level 1 variables (intercept, Time), Level 2 variables (intervention group, EXT PRS), and Interaction components using the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).

Alcohol dependence symptoms. We conducted hierarchical multiple regression in *R* to evaluate the effect of EXT PRS, intervention group, and PRS*intervention group on AD symptoms at age 18-19. Tests of change in R^2 were used to evaluate the effect of the

PRS*intervention interaction over and above the main effects and covariates. All ten ancestry PCs, TRISK, and sex were included as covariates and *p*-values were adjusted using the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).

Aim 2: Examining peer deviance as a mediator of gene-by-intervention effects. Aim 2 analyses test whether peer deviance at wave 6 mediated an effect of the intervention on alcohol consumption and AD symptoms for those at varying levels of genetic risk. Correlations of all mediating, moderating, predictor, and outcome variables were computed. Analyses were conducted using the *mediation* package in *R*. In the first step, we regressed the effect of EXT PRS * intervention and covariates (gender, TRISK, and ancestry PCs) onto the centered mediating variable. Next, we estimated the effect of EXT PRS * peer deviance, EXT PRS * intervention, and covariates on the outcome variable. In the final step, we specified the levels of the moderator (EXT PRS) at which to calculate the mediation function, setting the values of EXT PRS at 1 standard deviation above and below mean. Finally, we tested for significant differences in the total, direct and indirect moderating effects of EXT PRS and intervention status on AD symptoms through peer deviance using a bias-corrected and accelerated bootstrap resample of 2,000 to calculate the 95% confidence interval for the indirect effects.

Results: Study Two

Descriptive Statistics and Preliminary Analyses

Descriptive Statistics. Descriptive statistics for the EA and AA participants are displayed in Table 8. Intervention and control participants were similar across most variables of interest in both EA and AA participants. There were no significant differences in TRISK scores between intervention and control participants in either ancestry group. Accordingly, TRISK was appropriate for use as a baseline measure of externalizing risk in the analyses. There was a significant difference in drinks per month at wave 8 for EA participants, with intervention participants reported significantly less alcohol consumption than controls. For AA participants, there was a significant difference in drinks per month at wave 10; however, control participants reported significantly fewer drinks per month than intervention participants. These variations are further examined in the growth curve analyses to follow. Please note, the drinks per month and AD symptoms were log-transformed to account for skewness and kurtosis.

Table 8.

Descriptive statistics for European Ancestry and African Ancestry PAL participants.

Variable	European Ancestry			African Ancestry		
	Intervention Mean (SD)	Control mean (SD)	*t* (p)	Intervention Mean (SD)	Control Mean (SD)	*t* (p)
Sample size	N = 138	N = 131		N=92	N = 99	
Gender (female=1)	n = 71	n = 65	χ^2(p) = .90 (.76)	n = 56	n = 45	χ^2(p) = 1.2 (.29)
Age in months at W6	228.31 (8.55)	228.15 (8.08)	-0.15 (0.65)	229.57 (0.93)	229.18 (7.96)	.75 (-.40)
TRISK	1.59 (0.73)	1.57 (0.71)	-0.21 (0.72)	2.24 (0.09)	2.04 (0.07)	-1.62 (.11)
EXT PRS	0.41 (0.91)	0.42 (0.97)	0.06 (0.9)	0.39 (0.01)	0.38 (0.01)	-.46 (.11)
Alcohol dependence symptoms (sum) W7	0.6 (0.66)	0.47 (0.6)	-0.8 (0.55)	0.45 (0.06)	0.38 (0.01)	-1.64 (.10)
Drinks per month W7	0.74 (0.69)	0.67 (0.68)	0.72 (0.72)	0.31 (0.06)	0.3 (0.06)	-.19 (.85)
Drinks per month W8	1.17 (0.65)	1.23 (0.68)	0.1 (0.02*)	0.88 (0.08)	0.83 (0.08)	-.46 (.65)
Drinks per month W9	1.14 (0.73)	1.15 (0.62)	0.88 (0.05)	0.88 (0.07)	0.87 (0.07)	-.09 (.93)
Drinks per month W10	1.06 (0.68)	1.13 (0.61)	0.09 (0.52)	1.09 (0.07)	0.81 (0.06)	-2.93 (.00**)

Deviant talk W6	1.91 (0.8)	1.92 (0.71)	0.72 (0.88)	1.94 (0.09)	1.98 (0.08)	.32 (.75)
Peer deviant talk W6	1.93 (0.87)	2.01 (0.77)	0.37 (0.29)	1.97 (0.1)	2.04 (0.08)	.50 (.62)
Rule breaking W6	0.06 (0.07)	0.07 (0.1)	0.16 (0.64)	0.07 (0.01)	0.08 (0.01)	.88 (.38)
Peer Antisocial Behavior W6	0.82 (0.69)	0.83 (0.71)	-0.15 (0.65)	0.61 (0.07)	0.66 (0.07)	.47 (.64)

Note: Waves are denoted by "W." For example, W6=Wave 6. EXT PRS for EA sample was derived using summary statistics from the Externalizing Consortium GWAS. EXT PRS for the AA sample was derived with a MultiPRS approach, using a linear combination of weights from the Externalizing Consortium GWAS and an ancestry-match discovery sample from the Collaborative Study of the Genetics of Alcoholism. *$p < .05$. **$p < .01$.

Evaluation of Externalizing PRS in PAL. The results of the hierarchical multiple regression analyses testing the associations between EXT PRS and phenotypes of interest in EA and AA samples are displayed in Table 9. The EXT PRS for the EA sample was derived from the ancestry-matched Externalizing Consortium discovery sample, while the analyses in the AA subsample were calculated with the MultiPRS method. Analyses controlled for ancestry PCs and sex. After adjusting for multiple testing using the Benjamini-Hochberg correction, there were no significant associations between the EXT PRS and relevant outcomes in the PAL EA or AA samples.

Table 9.

Variance in relevant phenotypes accounted for by externalizing PRS in the PAL European Ancestry, and African Ancestry samples

Outcome	**European Ancestry**				**African Ancestry**			
	N	Change in R^2	*p*	p^a	N	Change in R^2	*p*	p^a
TRISK	269	.009	.112	.499	191	.019	.042*	.382
Alcohol dependence symptoms (sum) W7	241	.002	.545	.663	168	.002	.531	.791
Drinks per month (log) W6	168	.000	.848	.848	90	.003	.877	.917
Drinks per month (log) W7	202	.020	.038*	.499	104	.001	.878	.917
Drinks per month (log) W8	227	.008	.188	.499	134	.004	.615	.791
Drinks per month (log) W9	225	.006	.233	.499	148	.000	.789	.917
Drinks per month (log) W10	266	.006	.199	.499	188	.005	.352	.791
Drinks per month (log) W11	244	.001	.575	.663	173	.010	.182	.791

Deviant Talk W6	226	.004	.336	.630	162	.003	.449	.791
Peer Deviant Talk W6	226	.003	.421	.663	162	.004	.389	.791
Rule Breaking W6	223	.001	.735	.788	160	.004	.441	.791
Deviant Peer Association W2	223	.002	.523	.663	160	.005	.385	.917
Deviant Peer Association W3	244	.009	.145	.499	161	.003	.519	.791
Deviant Peer Association W4	239	.006	.218	.499	163	.028	.026*	.381
Peer Antisocial Behavior W6	243	.002	.493	.663	167	.007	.268	.791

Note: Change in R^2 represents the variance accounted for by EXT PRS over and above the effect of Ancestry PCs and sex Waves are denoted by "W." For example, W6=Wave 6. EXT PRS for EA sample was derived using summary statistics from the Externalizing Consortium GWAS. EXT PRS for the AA sample was derived with a MultiPRS approach, using a linear combination of weights from the Externalizing Consortium GWAS and an ancestry-matched discovery sample from the Collaborative Study of the Genetics of Alcoholism. *$p < .05$. **$p < .01$.

[a]*p*-values have been adjusted for multiple testing using the Benjamini–Hochberg procedure.

Aim 1: Results of Analyses Examining GxI Effects on Alcohol Consumption and AD Symptoms in the PAL Sample.

Alcohol Consumption in the EA Sample. Results of the MLM for alcohol consumption are displayed in Table 10. The MLM was first constructed as an unconditional means model (Null Model; Table 10), with an estimated Intra-Class Correlation (ICC) coefficient of .45, indicating that slightly less than half (45%) of the variance in alcohol consumption was due to differences between individuals. In the unconditional growth model (Unconditional Growth Model; Table 10), the significant effects for intercept, time, and time^2 indicated that there was sufficient within-person variation in trajectories of alcohol consumption to warrant a multilevel framework. The intercept value for the Unconditional Growth Model indicates that the average value for log-transformed alcohol consumption at Wave 7 (age 16-17) was 1.03. The positive linear slope, represented by Time (Beta = 4.05), indicated that on average, alcohol consumption increased over time; however, the quadratic equation provided a significantly better fit for the data. The negative quadratic slope, represented by Time^2 of -5.15, indicated that the slope of alcohol consumption was curvilinear in nature with alcohol consumption increasing over time for

before dropping slightly at the final time point. Results of a likelihood ratio chi-square difference test comparing the Unconditional Growth Model with random intercept and slope (AIC=1749.7, ICC=.61) to the Null Model (AIC=1915.8, ICC=.45) showed significantly improved model fit with the Unconditional Growth Model [$X^2(4, 269) = 174.02$, $p<.001$].

Next, we constructed the Conditional Growth Model with fixed effects for time, intervention group, and EXT PRS, and random effects for slope and intercept at Level 1. We compared conditional models with random intercept only, random slopes only, and a combined model with random intercept and slopes, with the latter providing the overall best fit for the data structure (AIC=1751.1, ICC=.61). In the Conditional Model, the intercept, linear slope, and quadratic slope remained significant; however, there was no evidence of significant main effects for intervention group or EXT PRS. Next, an Interaction Model was constructed to evaluate whether there were variations in the slope and intercept of alcohol consumption across time as a function of the interaction between intervention group, EXT PRS, and time. In the Interaction Model, there was a significant three-way interaction between intervention group, EXT PRS, and $Time^2$, such that the trajectory of alcohol consumption varies based on EXT PRS and intervention group. We proceeded with the addition of covariates to fully evaluate the final model. There were significant effects for gender and TRISK score suggesting that these factors significantly influenced trajectories of alcohol consumption among EA in PAL. The significant negative associations between gender and alcohol consumption suggests that being a female was associated with lower rates of alcohol consumption over time. Similarly, higher TRISK score was associated with lower rates of alcohol consumption over time, which is a surprising finding given the measure is associated with higher risk of externalizing problems. There were no significant main effects for intervention group or EXT PRS observed, and the significant

interaction between EXT PRS, intervention group, and Time^2 did not survive Benjamini-Hochberg corrections for multiple testing in the final model.

Table 10.

Multilevel Quadratic Growth Curve Analysis of Alcohol Consumption in PAL EA Sample

Level and Variable	Model				
	Null	Unconditional Growth	Conditional Growth	Interaction Model	Interaction Model with Covariates[1]
Level 1					
Intercept	1.03 (.03)**	1.03 (.03)**	1.03(.05)**	1.03 (.05)**	1.17 (.05)**
Time		4.05 (.53)**	4.05 (.53)**	4.86 (.76)**	4.86 (.76)**
Time^2		-5.15 (.50)**	-5.15(.45)**	-5.50 (.64)**	-5.49 (.64)**
Level 2					
Intervention (control=0)			-.02 (.07)	-.01 (.07)	.00 (.06)
EXT PRS			.06 (.03)	.03 (.05)	.03 (.04)
Interaction components					
Time * EXT PRS				-.11 (.79)	-.11 (.79)
Time * Intervention				.06 (.07)	-1.55 (1.06)
Time^2 * EXT PRS				.46 (.69)	1.00 (.65)
Time^2 * Intervention				.71 (.89)	.68 (.89)
EXT PRS * Intervention				.06 (.07)	.06 (.07)
Time*EXT PRS*Intervention				.46 (1.14)	.39 (1.14)
Time^2*EXT PRS*Intervention				-2.03 (.95)*	-2.01 (.95)
Covariates					
Gender (female=1)					-.29 (.06)**
TRISK					-.10 (.05)*
Additional Information					
ICC	.45	.61	.61	.61	.59
-2 log likelihood (FIML)	-954.9	-	-866.6	-862.6	-847.7
AIC	1915.8	1749.73	1751.1	1757.1	1751.5
Pseudo R^2 (fixed effects)	.00	-	-	-	-
Pseudo R^2 (total)	.45	-	-	-	-
Number of individuals	269	269	269	269	269
Observations	1016	1016	1016	1016	1016

Note: Estimates of unstandardized coefficients are presented for fixed effects. Pseudo R^2 cannot be calculated for quadratic growth. The Unconditional Growth Model failed to converge, thus the -2 log likelihood could not be calculated. Values in parentheses are standard errors. *$p < .05$, **$p < .01$.

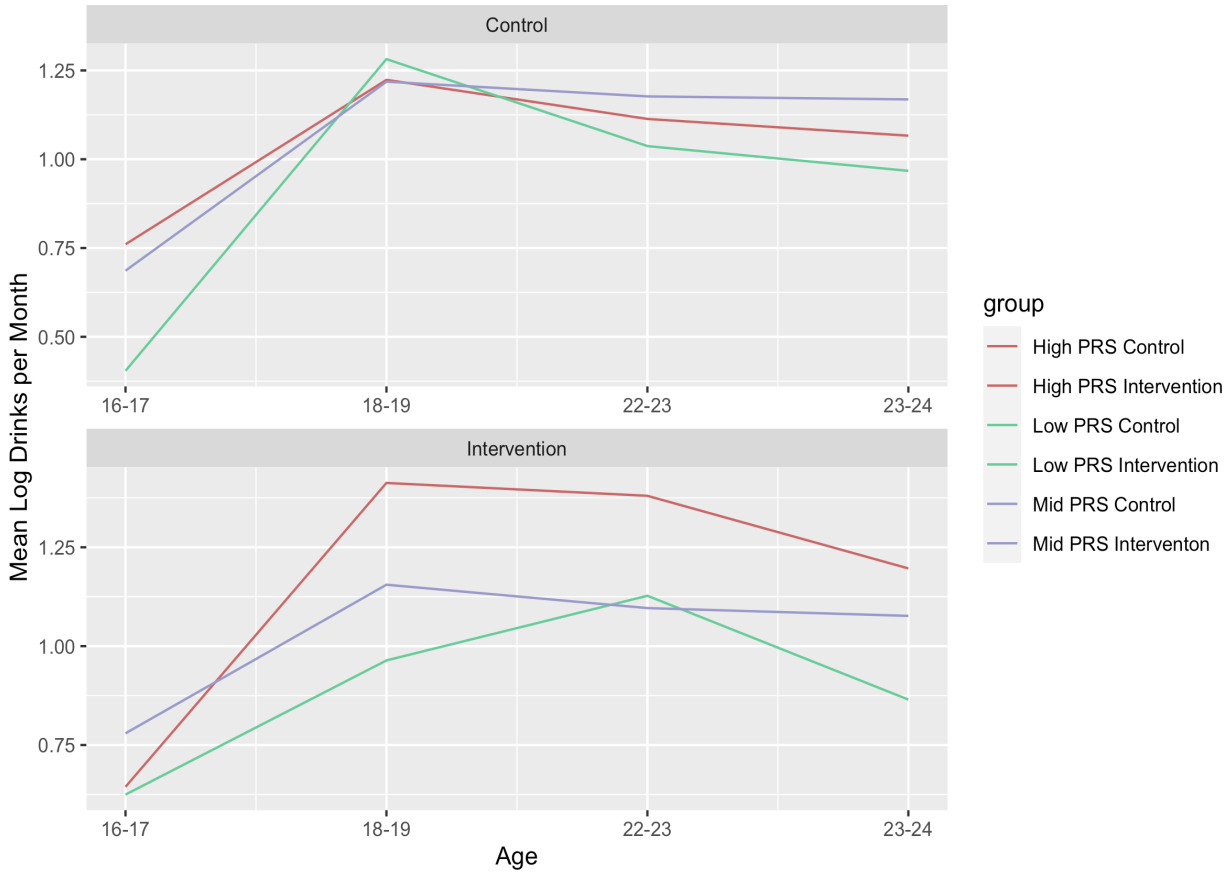
[1] p-values for Level 1, Level 2, and Interaction components were adjusted for multiple testing using the Benjamini–Hochberg procedure in the Interaction Model with Covariates.

To better examine the trajectories of alcohol consumption across time, we plotted the mean of drinks per month over time for individuals who fall within different ranges of EXT PRS in the interaction and control group (Figure 6). Means were plotted separately for control and

intervention participants with High PRS (greater than 1 SD above the mean), Mid PRS (within 1 SD around the mean), and Low PRS (less than 1 SD below the mean). Based on visual examination of the data, the slope and curvature of the High PRS participants in the intervention condition appears steeper than Mid PRS and Low PRS participants who participated in the intervention. This suggests that individuals in the intervention condition with higher EXT PRS increased alcohol consumption more sharply from age 16-17 to age 18-19 and consumed alcohol at a higher rate compared to Mid PRS and Low PRS intervention participants. In the control condition, Low PRS participants appeared to report lower drinks per month on average at age 16-17 (wave 7) but caught up to Mid PRS and High PRS control group participants by age 18-19 before decreasing more sharply by age 23-24. In summary, the slope and shape of the change in alcohol consumption over time varies based on the interaction between EXT PRS and intervention condition, with High PRS participants in the intervention group appearing to be less responsive to the effects of the intervention and Low PRS participants in the control group reporting fewer drinks per month at age 16-17. However, the results of the growth curve analyses suggest that the observed variation in the trajectory of alcohol consumption across time for individuals with varying levels of EXT PRS in the intervention and control groups were not significantly different following correction for multiple testing.

Figure 6.

Plotted Log Means of Alcohol Consumption (drinks per month) across Time for Intervention and Control Participants at Different Levels of EXT PRS in the PAL EA Sample.



Note. Values of EXT PRS were as follows: High PRS is greater than 1 SD above the mean, Mid PRS is within 1 SD around the mean, and Low PRS is less than 1 SD below the mean.

Alcohol Consumption in the AA Sample. Results of the MLM for alcohol consumption are displayed in Table 11. The MLM was first constructed as an unconditional means model (Null Model; Table 11), with an estimated Intra-Class Correlation (ICC) coefficient of .35, indicating that approximately a third of the variance in alcohol consumption was due to differences between individuals. In the unconditional growth model (Unconditional Growth Model; Table 11), the significant effects for intercept, Time, and Time^2 indicated there was

sufficient within-person variation in trajectories of alcohol consumption to warrant a multilevel framework. The intercept value for the Unconditional Growth Model indicates that the average value for log-transformed alcohol consumption at Wave 7 (age 16-17) was 0.75. The positive linear slope, represented by Time (Beta = 5.59), indicated that on average, alcohol consumption increased over time; however, the quadratic equation provided a significantly better fit for the data. The negative quadratic slope, represented by $Time^2$ of -4.01, indicated that the slope of alcohol consumption was curvilinear in nature with alcohol consumption increasing over time before leveling off. Results of a likelihood ratio chi-square difference test comparing the Unconditional Growth Model with random intercept and slope (AIC=1299.0, ICC=.40) to the Null Model (AIC=1441.2, ICC=.35) showed significantly improved model fit with the Unconditional Growth Model [$X^2(4, 191) = 150.19$, $p<.001$].

Next, we constructed the Conditional Growth Model with fixed effects for time, intervention group, and EXT PRS, and random effects for slope and intercept at Level 1. We compared conditional models with random intercept only, random slopes only, and a combined model with random intercept and slopes. When comparing the random intercepts model (AIC=1299.7, ICC=.43) to the random intercepts and slopes model (AIC=1301.3, ICC=.41), the random intercepts model resulted in lower AIC and -loglikelihood values suggesting improved model fit. Accordingly, we proceeded with the Conditional Growth Model with random intercepts and fixed effects for time, intervention group, and EXT PRS. In the Conditional Model, the intercept, linear slope, and quadratic slope remained significant; however, there was no evidence of significant main effects for intervention group or EXT PRS. Next, an Interaction Model was constructed to evaluate whether there were variations in alcohol consumption across time as a function of the interaction between intervention group and EXT PRS. There was

slightly improved model fit, as indicated by lower AIC, but there was no evidence that trajectories of AUD symptoms varied as a function of two-way or three-way interactions between Time, Time^2, intervention group and EXT PRS. We proceeded with the addition of covariates to fully evaluate the final model. There were significant main effects for gender, such that being a female was associated with lower rates of alcohol consumption over time. There was also a significant interaction between Time and intervention group; however, there was no significant interaction with the quadratic term (Time^2). Given that the quadratic equation provided a better overall fit for the structure of the data, the interaction between Time and intervention group is not interpreted within the context of the null effects for Time^2. Furthermore, there was no evidence of three-way interactions between EXT PRS, intervention group, and time on alcohol consumption in the PAL AA sample.

Table 11.

Multilevel Quadratic Growth Curve Analysis of Alcohol Consumption in PAL AA Sample

Level and Variable	Model				
	Null	Unconditional Growth	Conditional Growth	Interaction Model	Interaction Model with Covariates[1]
Level 1					
Intercept	.76 (.04)**	.75 (.04)**	.70 (.05)**	.70 (.05)**	.82 (.07)**
Time		5.59 (.53)**	5.57 (.52)**	4.42 (.71)**	4.44 (.71)**
Time^2		-4.01 (.52)**	-4.00(.52)**	-4.84 (.72)**	-4.86 (.72)**
Level 2					
Intervention (control=0)			.11 (.07)	.10 (.07)	.10 (.07)
EXT PRS			.01 (.04)	.02 (.05)	-.05 (.06)
Interaction components					
Time * EXT PRS				-.28 (5.73)	-.04 (.73)
Time^2 * EXT PRS				-11.29 (5.85)	-1.44 (.74)
Time * Intervention				.18 (3.31)	2.37 (1.02)
Time^2 * Intervention				-3.58 (3.35)	1.64 (1.04)
EXT PRS * Intervention				-.09 (.58)	.02 (.07)
Time*EXT PRS*Intervention				5.71 (8.18)	.73 (1.04)
Time^2*EXT PRS*Intervention				13.45 (8.30)	1.72 (1.06)
Covariates					
Gender (female=1)					-.23 (.08)**
TRISK					-.00 (.05)
Additional Information					
ICC	.35	.40	.43	.44	.40
-2 log likelihood (FIML)	-717.6	-642.5	-642.8	-636.5	-626.7
AIC	1441.21	1299.02	1299.7	1301.0	1305.4
Pseudo R^2 (fixed effects)	.00	-	-	-	-
Pseudo R^2 (total)	.35	-	-	-	-
Number of individuals	191	191	191	191	191
Observations	711	711	711	711	711

Note: Values in parentheses are standard errors. Time was modeled with as both a linear and quadratic factor, with the quadratic equation providing a better fit for the data structure. Results of the quadratic growth models are presented in the table. For the Full Interaction Model, Ancestry PCs 2-10 were dropped from the model to improve statistical power. Pseudo R^2 cannot be calculated for non-linear growth curves. *$p < .05$, **$p < .01$.

[1] p-values for Level 1, Level 2, and Interaction components were adjusted for multiple testing using the Benjamini–Hochberg procedure in the Interaction Model with Covariates.

Alcohol Dependence Symptoms in the EA Sample. We conducted a hierarchical multiple regression testing for an interaction between EXT PRS and intervention condition on AD symptoms in the EA sample. Results are displayed in Table 12. We observed a significant main effect of TRISK score, such that higher scores on teacher perception of child behavioral risk predicted greater AD symptoms at age 18-19, $t(225)=2.34$, p = .02. However, there was no

evidence of significant main effects for intervention group or EXT PRS, and no evidence of an interaction between intervention group and EXT PRS on AD symptoms, $F(15, 255) = 1.42$, p = .14. Results show no evidence for differential effectiveness of the intervention of AD symptoms for individuals at varying levels of genetic risk.

Table 12.

Results of hierarchical multiple regression examining the interactions between Externalizing PRS and Intervention on AD symptoms in the PAL EA Sample

Predictor	b^1	b 95% CI	sr^2	sr^2 95% CI
(Intercept)	0.32**	[0.05, 0.58]		
EXT PRS	0.01	[-0.29, 0.31]	.00	[-.00, .00]
Intervention	0.37	[-0.07, 0.81]	.01	[-.01, .04]
TRISK	0.45**	[0.14, 0.76]	.03	[-.01, .08]
Gender (female)	-0.22	[-0.63, 0.19]	.00	[-.01, .02]
Ancestry PC1	0.25	[-0.04, 0.54]	.01	[-.01, .04]
Ancestry PC2	-0.29	[-0.81, 0.24]	.00	[-.01, .02]
Ancestry PC3	-0.16	[-0.52, 0.21]	.00	[-.01, .02]
Ancestry PC4	0.09	[-0.26, 0.44]	.00	[-.01, .01]
Ancestry PC5	0.01	[-0.42, 0.44]	.00	[-.00, .00]
Ancestry PC6	-0.15	[-0.63, 0.34]	.00	[-.01, .01]
Ancestry PC7	-0.21	[-0.62, 0.20]	.00	[-.01, .02]
Ancestry PC8	-0.19	[-0.51, 0.13]	.01	[-.01, .02]
Ancestry PC9	0.07	[-0.16, 0.30]	.00	[-.01, .01]
Ancestry PC10	0.05	[-0.17, 0.27]	.00	[-.01, .01]
EXT PRS * Intervention	-0.12	[-0.55, 0.32]	.00	[-.01, .01]
		Model Fit: R^2 = .087*, 95% CI [.00, .10]		

Note. A significant *b*-weight indicates the semi-partial correlation is also significant. *b* represents unstandardized regression weights. sr^2 represents the semi-partial correlation squared. AD symptoms were log-transformed to account for kurtosis. *p*-values were adjusted for multiple testing using the Benjamini–Hochberg procedure. * indicates $p < .05$, ** indicates $p < .01$ after correction.

Alcohol Dependence Symptoms in the AA Sample. We conducted a hierarchical multiple regression testing for an interaction between EXT PRS and intervention condition on

AD symptoms in the EA sample. Results are displayed in Table 13. We observed a significant main effect of intervention group, such that being in the intervention group predicted greater AD symptoms at age 18-19. This result suggests that the intervention may have been less effective at preventing AD symptoms among the African American sample. There was no evidence of a significant main effect for EXT PRS, and there was no evidence of an interaction between intervention group and EXT PRS on AD symptoms. Results show no evidence for differential effectiveness of the intervention of AD symptoms for individuals at varying levels of genetic risk in the PAL AA sample.

Table 13.

Results of hierarchical multiple regression examining the interactions between Externalizing PRS and Intervention on AD symptoms in the PAL AA Sample

Predictor	b^1	*b* 95% CI	sr^2	sr^2 95% CI
(Intercept)	0.35**	[0.19, 0.52]		
EXT PRS	-0.12	[-0.27, 0.02]	.02	[-.02, .05]
Intervention	0.19	[0.01, 0.38]	.03	[-.02, .07]
TRISK	-0.10	[-0.22, 0.02]	.02	[-.02, .05]
Gender (female)	-0.12	[-0.32, 0.08]	.01	[-.02, .04]
Ancestry PC1	0.00	[-0.16, 0.16]	.00	[-.00, .00]
Ancestry PC2	-0.03	[-0.17, 0.12]	.00	[-.01, .01]
Ancestry PC3	0.02	[-0.08, 0.11]	.00	[-.01, .01]
Ancestry PC4	0.01	[-0.10, 0.12]	.00	[-.01, .01]
Ancestry PC5	0.05	[-0.06, 0.16]	.01	[-.02, .03]
Ancestry PC6	0.01	[-0.10, 0.12]	.00	[-.00, .00]
Ancestry PC7	-0.02	[-0.13, 0.08]	.00	[-.01, .01]
Ancestry PC8	-0.07	[-0.17, 0.03]	.01	[-.02, .04]
Ancestry PC9	0.03	[-0.08, 0.14]	.00	[-.01, .02]
Ancestry PC10	-0.02	[-0.12, 0.09]	.00	[-.01, .01]
EXT PRS * Intervention	0.15	[-0.05, 0.34]	.01	[-.02, .05]
			Model Fit: R^2 = .08, 95% CI [.00, .07]	

Note. A significant *b*-weight indicates the semi-partial correlation is also significant. *b* represents unstandardized regression weights. sr^2 represents the semi-partial correlation squared. AD symptoms were log-transformed to account for kurtosis. *p*-values were adjusted for multiple testing using the Benjamini–Hochberg procedure. * indicates $p < .05$, ** indicates $p < .01$ after correction.

Aim 2: Results of examining peer deviance as a mediator of gene-by-intervention effects on alcohol consumption and AD symptoms.

Peer Deviance as a Mediator in the EA Sample. Mediation analyses were conducted with outcome variables (drinks per month and AD symptoms) at a single time point. Descriptive statistics for the EA sample (Table 8) indicated that control participants reported significantly more drinks per month at wave 8 (age 22-23) compared to intervention participants. Accordingly, wave 8 drinks per month was selected to test for mediation of GxI effects via peer deviance. AD symptoms were measured only at wave 7. Peer deviance was represented by a measure of peer antisocial behavior measures at wave 6. In preparation for analyses, we calculated correlations between all independent, dependent, mediating, and moderating variables, displayed in Table 14. We observed several significant correlations between variables. Gender as female was associated with significantly lower peer antisocial behavior, drinks per month, AD symptoms, and TRISK scores. Peer antisocial behavior was positively correlated with EXT PRS, which aligns well with the theoretical relationship between genetic risk associated with externalizing behavior and observed measures of externalizing in the PAL EA sample. We proceeded with the mediated moderation analyses using separate models to examine each dependent variable in succession, while controlling for significantly correlated variables in each respective model.

Table 14.

Means, standard deviations, and correlations for variables included in the EA PAL Mediated Moderation models

Variable	*M*	*SD*	1	2	3	4	5	6
1. EXT PRS	0.45	0.95						
2. Intervention group	0.50	0.50	.07					
3. Gender (female=1)	0.50	0.50	.09	-.03				
4. Peer Antisocial Behavior W6	0.88	0.67	.14*	-.01	-.17*			
5. Drinks per Month W8	1.25	0.66	.02	-.01	-.17*	.19**		
6. AD Symptoms (log)	0.57	0.65	-.07	.11	-.16*	.14	.31**	
7. TRISK	1.56	0.71	-.12	-.01	-.22**	.07	-.02	.19**

Note. M and *SD* are used to represent mean and standard deviation, respectively. W represents Wave, for example W8 = Wave 8. * indicates $p < .05$. ** indicates $p < .01$.

Results of the mediation analyses are displayed in Table 15. Analyses covaried for the effects of gender and the first two ancestry PCs. We observed no evidence that peer deviance at wave 6 mediated the effect of the intervention on drinks per month at wave 8 or AD symptoms for those at varying levels of genetic risk. There was also no evidence of direct effects of peer deviance on drinks per month or AD symptoms. In the absence of indirect and direct effects, the total effect, which is the sum of the indirect and direct effects, also showed no evidence of a significant effect on the outcomes. In summary, peer deviance was significantly correlated with EXT PRS, but it did not explain the variation in intervention effects on alcohol-related behaviors for individuals with high versus low levels of EXT PRS.

Table 15.

Results of Mediated Moderation Analyses for EA in PAL when the Moderator is Set to High PRS

	Drinks per month (W8) *n =204*			AD Symptoms (W7) *n = 204*		
Variable	*Estimate*	*95% CI*	*p*	*Estimate*	*95% CI*	*p*
ACME (control)	.00	[-.01, .04]	.67	-.00	[-.02, .02]	.97
ACME (intervention)	-.01	[-.04, .01]	.67	-.00	[-.01, .06]	.70
ADE (control)	.04	[-.16, .22]	.72	.11	[-.12, .28]	.34
ADE (intervention)	.03	[-.18, .22]	.78	.12	[-.10, .28]	.30
Total Effect	.03	[-.17, .22]	.75	.12	[-.11, .28]	.31
Proportion Mediated (control)	.14	[-.19, 3.84]	.87	-.00	[-.13, 2.82]	.95
Proportion Mediated (intervention)	-.17	[-5.25, .41]	.91	.04	[-.11, 4.98]	.81
ACME (average)	-.00	[-.02, .01]	.93	.00	[-.01, .04]	.78
ADE (average)	.03	[-.17, .22]	.74	.11	[-.11, .28]	.32
Proportion Mediated (average)	-.02	[-.22, 1.74]	.98	.02	[-.05, 2.09]	.84

Note. ACME is the Average Causal Mediated Effect (i.e., the indirect effect of mediator on the outcome), ADE is the Average Direct Effect (i.e., the direct effect of mediator on the outcome), and Total Effect represents the sum of the ACME and ADE. The moderator (EXT PRS) was set to 1 SD above the mean for the analyses reported in the table. Estimates and CI were calculated using a bias-corrected and accelerated bootstrap resample of 2,000 simulations. * indicates $p < .05$, ** indicates $p < .01$.

Peer Deviance as a Mediator in the AA Sample. Mediation analyses were conducted with outcome variables (drinks per month and AD symptoms) at a single time point. Descriptive statistics for the EA sample (Table 8) indicated that control participants reported significantly more drinks per month at wave 10 (age 26-27) compared to intervention participants. Accordingly, wave 10 drinks per month was selected to test for mediation of GxI effects via peer deviance. AD symptoms were measured only at wave 7. Peer deviance was represented by a measure of peer antisocial behavior measures at wave 6. In preparation for analyses, we calculated correlations between all independent, dependent, mediating, and moderating variables, displayed Table 16. We observed several significant correlations between variables. Gender as female was associated with significantly lower EXT PRS, drinks per month, and TRISK scores. Peer antisocial behavior was positively correlated with drinks per month, and AD symptoms. We proceeded with the mediated moderation analyses using separate models to examine each dependent variable in succession, while controlling for significantly correlated variables in each respective model.

Table 16.

Means, standard deviations, and correlations for variables included in the AA PAL Mediated Moderation models

Variable	*M*	*SD*	1	2	3	4	5	6
1. EXT PRS	0.38	0.12						
2. Intervention group	0.45	0.50	.08					
3. Gender (female=1)	0.56	0.50	-.17*	-.05				
4. Peer Antisocial Behavior W6	0.63	0.58	.05	-.06	-.21*			
5. Drinks per month W8 (log)	0.83	0.71	.02	-.01	-.19*	.25**		
6. AD Symptoms (log)	0.39	0.57	-.04	.09	-.07	.21*	.46**	
7. TRISK	2.10	0.79	.09	.10	-.28**	.05	.03	-.13

Note. M and *SD* are used to represent mean and standard deviation, respectively. W represents Wave, for example W8 = Wave 8. * indicates $p < .05$, ** indicates $p < .01$.

Results of the mediation analyses are displayed in Table 17. Analyses covaried for the effects of gender and the first two ancestry PCs. We observed no evidence that peer deviance at wave 6 mediated the effect of the intervention on drinks per month at wave 10 for those at varying levels of genetic risk. However, we did observe significant direct effects of peer deviance on drinks per month among both the intervention and control groups for individuals at high and low levels of EXT PRS. The total effect, which is the sum of the indirect and direct effects, was also significant for both peer deviance; however, in the absence of a significant indirect effect, this result was likely driven by the strength of the direct effect. For AD symptoms, we observed no evidence that peer deviance at wave 6 mediated the effect of the intervention on AD symptoms. There was also no evidence of direct effects or total effects of peer deviance on AD symptoms. In summary, peer deviance significantly influenced drinks per month at wave 10, but it did not explain the variation in intervention effects on drinks per months or AD symptoms for individuals with high versus low levels of EXT PRS.

Table 17.

Results of Mediated Moderation Analyses for AA in PAL when the Moderator is Set to High PRS

	Drinks per month W10 *n =151*			AD Symptoms (W7) *n = 148*		
Variable	*Estimate*	*95% CI*	*p*	*Estimate*	*95% CI*	*p*
ACME (control)	-.02	[-.10, .02]	.50	-.01	[-.08, .01]	.64
ACME (intervention)	.01	[-.01, .10]	.56	.00	[-.03, .04]	.96
ADE (control)	.39	[.14, .63]	.00**	.18	[-.03, .37]	.10
ADE (intervention)	.42	[.15, .70]	.00**	.19	[-.04, .40]	.11
Total Effect	.40	[.15, .66]	.00**	.18	[-.04, .38]	.12
Proportion Mediated (control)	-.04	[-.34, .04]	.50	-.03	[-1.33, .06]	.67
Proportion Mediated (intervention)	.02	[-.04, .20]	.56	.01	[-.68, .26]	.92
ACME (average)	-.00	[-.04, .02]	.90	-.00	[-.05, .01]	.79
ADE (average)	.40	[.14, .65]	.00**	.18	[-.04, .39]	.11
Proportion Mediated (average)	-.01	[-.24, .03]	.86	-.01	[-5.63, .05]	.84

Note. ACME is the Average Causal Mediated Effect (i.e., the indirect effect of mediator on the outcome), ADE is the Average Direct Effect (i.e., the direct effect of mediator on the outcome), and Total Effect represents the sum of the ACME and ADE. The moderator (EXT PRS) was set to 1 SD above the mean for the analyses reported in the table. Estimates and CI were calculated using a bias-corrected and accelerated bootstrap resample of 2,000 simulations. * indicates $p < .05$. ** indicates $p < .01$.

Discussion: Study Two

The present study integrates polygenic scores into prevention and intervention research by examining whether polygenic risk associated with externalizing problems influenced the effect of a randomized, family-centered intervention for adolescents on alcohol-related behaviors in emerging adulthood. We conducted analyses in European American and African Americans separately, and discuss the findings for each group separately.

In the European American sample, there was some preliminary evidence that the intervention significantly moderated the effect of polygenic risk associated with externalizing behaviors on growth in alcohol consumption across emerging adulthood; however, this effect did not survive corrections for multiple testing, suggesting no reliable differences in alcohol consumption due to the interaction between intervention group and polygenic risk for externalizing behaviors. Second, we tested for GxI effects on AD symptoms and found no evidence to suggest that EXT PRS moderated the effect of the intervention on AD symptoms at age 23-24 in the European American sample. Third, we examined peer deviance as a mediator of GxI effects on alcohol consumption and AD symptoms and found no evidence of an indirect effect of peer deviance on intervention effects for individuals at varying levels of genetic risk in the European American sample.

We repeated the same three analyses in the African American sample: GxI effects on trajectories of alcohol consumption, GxI effects on AD symptoms at age 23-24, and mediation of GxI effects by peer deviance. In the African American sample, there was no evidence of GxI effects on alcohol consumption or AD symptoms. There was also no support for peer deviance as a mediator of GxI effects on alcohol consumption or AD symptoms in African Americans.

Accordingly, these findings suggest that intervention effects on alcohol consumption and AD symptoms do not differ in individuals with varying levels of polygenic risk in the PAL European American and African American sample. Our hypothesis was that individuals at greater genetic risk would report greater reductions in alcohol-related outcomes relative to individuals with lower genetic risk and controls. The initial hypothesis was based on existing GxI literature in line with the differential susceptibility hypothesis, which suggests that some individuals are more sensitive to both promotive (e.g., positive parenting, supportive peer groups) and harmful (e.g., trauma exposure, deviant peers) environments (Bakermans-Kranenburg & van IJzendoorn, 2015; Belsky & Pluess, 2009; van Ijzendoorn & Bakermans-Kranenburg, 2015). However, as explored in the discussion of the Spit for Science findings, there are a number of potential reasons why the existing body of research on differential susceptibility in GxI studies may not align well with the present study's methodology, and consequently, the null findings. The differential susceptibility hypothesis emphasizes the importance of plasticity genes, or genetic sensitivity to environmental exposure (Belsky & Pluess, 2009). Our polygenic score for externalizing behavior indexes the influence of genetic variants across the genome on the development of this cluster of related phenotypes. Although some environmental sensitivity is likely captured by these scores (Young et al., 2019), they are not an explicit measurement of how likely an individual is to be affected by promotive or harmful environments. Second, the existing GxI literature is predominantly comprised of candidate gene studies, which are prone to false positives and publication bias due to insufficient statistical power to detect the very small effects of candidate genes (Border et al., 2019; Dick et al., 2015; Johnson et al., 2017). Furthermore, candidate gene studies do not align with current understanding that complex

behaviors are polygenic in nature (Visscher et al., 2017). Taken together, these considerations may explain why the existing body of research on differential susceptibility may not align well with the present study's methodology, and partly explain the study's null findings.

Although we observed no support for the hypotheses in this study, there are a few relevant studies that have observed significant GxI effects using polygenic scores. First, using the same sample as the present study, Kuo et al. (2019) examined whether the Family Check-up intervention moderated the effect of polygenic risk for alcohol dependence on lifetime diagnosis of alcohol dependence at age 26-27. The study used summary statistics from a well-powered GWAS of DSM-IV alcohol dependence (Gelernter et al., 2014b) to derive polygenic scores. Kuo et al. found that the intervention moderated the effect of alcohol dependence polygenic scores, such that higher polygenic risk was associated with greater likelihood of alcohol dependence diagnosis in the control condition, but not in the intervention condition in the European American sample. The findings suggest that the effects of underlying genetic predispositions on age 26-27 alcohol dependence diagnosis were mitigated by the intervention. A second study, in which the Family Check-Up intervention was delivered to families of young children (age 2), showed a similar pattern of results indicating mitigation of genetic risk by the intervention. Elam et al. (2020) examined the influence of polygenic risk for aggression on aggressive behavior and peer rejection in early adolescence. Although it was not a formal test of GxI interaction, they found that the aggression polygenic score predicted peer rejection in control participants, but the same effect was not observed in intervention participants. Consistent with the Kuo et al. study, these findings provide an indication that the Family Check-Up intervention may blunt the effect of certain genetic predispositions on the relevant outcomes.

There are a number of possible reasons why the present study's null findings are not consistent with the pattern of findings from the Kuo et al. and Elam et al. studies. One possible factor that may have influenced the variable pattern of findings is the difference in alcohol outcomes reported. Kuo et. al examined AD diagnosis, whereas the GxI effects examined in this study focused on patterns of alcohol consumption across time and AD symptoms. In the present study, alcohol consumption and AD symptoms were moderately correlated in both European Americans [$r(202) = .31$, $p<.01$] and African Americans [$r(146) = .46$, $p<.01$], indicating that variations in findings across the outcomes are plausible. In addition, there are also unique and shared components to the underlying genetic architecture of the two phenotypes (Kranzler et al., 2019; Walters et al., 2018). Accordingly, although heavy alcohol consumption is a risk factor for the development of alcohol-related problems, not everyone who drinks heavily meet criteria for alcohol use disorder.

A second possible explanation of the discrepancy relates to the Kuo et al. study's use of an alcohol dependence polygenic score to measure effects on AD diagnosis in the target sample. The discovery sample phenotype and the target sample phenotype were very closely aligned. The polygenic score for the present study was derived from a multivariate GWAS of seven different externalizing phenotypes that was used to predict only two phenotypes in the target sample (i.e., alcohol consumption and AD symptoms). Although the externalizing PRS has been shown to significantly predict alcohol consumption and alcohol dependence in target samples (Karlsson Linnér et al., 2020), it likely accounts for much broader phenotypic variation than what is represented by the two alcohol-related outcomes in this study. Furthermore, the externalizing pathway of risk to alcohol-related problems is just one of several pathways through which

underlying genetics may influence alcohol use outcomes (Dick & Agrawal, 2008; Kendler et al., 2003; Saraceno et al., 2009; Schuckit, 2009). For example, the influence of the internalizing pathway of risk to substance use problems, characterized by patterns of self-medication or drinking to cope with distress, may not be well-represented by the externalizing PRS. Low level of response to alcohol is another genetically-informed risk factor for future alcohol problems, yet it too may not be indexed by the externalizing PRS (Morozova et al., 2014; Schuckit, 2018; Schuckit et al., 1997). The genetic correlation between alcohol consumption and externalizing behaviors is .50 (Karlsson Linnér et al., 2020), suggesting that a well-powered, polygenic score for alcohol consumption may capture additional risk pathways through which genetics influence alcohol-related behaviors, and thus result in different findings.

Another possible contributing factor to the present study's null findings is statistical power. The sample sizes for the Europeans American (n=269) and African American (n=191) PAL participants were relatively small, and thus underpowered to detect a GxI effect in the growth curve analyses examining trajectories of alcohol consumption across time. Hierarchical multiple regression has less stringent power requirements, but even these models, which examined GxI effects on AD symptoms in PAL, were limited in power. Given a model with 15 predictors, α =.05, and power (1-β) = .8, the European American sample was adequately powered to detect effect sizes of f^2 >.029 and the African American sample was adequately powered to detect effect sizes of f^2 >.042. The effect size observed in the Spit for Science sample was f^2 =.038, meaning that only the PAL European American was sufficiently powered to detect a similar effect. However, it is also possible that the effect size observed in the Spit for Science study is an overestimate of GxI effects, perhaps due to problems related to scale

transformation. Based on conservative recommendations from a systematic review of alcohol, tobacco and cannabis GxE studies using polygenic scores (Pasman et al., 2019), a sample size of 1572 would be required to achieve 80% power to detect an effect size of $f^2 = .005$ in a hierarchical multiple regression with 15 predictors. A larger GxI sample would improve the ability to detect very small interaction effect if they are present in the population or confirm that the absence of an effect is not due to Type II error.

One additional factor that contributed to diminished power in this study was the decision to analyze European American and African American samples separately due to differences in genetic architecture related to ancestry that can bias results (Peterson et al., 2019). In the African American sample, we used an empirically-supported method for improving PRS prediction power in non-European ancestry groups by combining PRS derived from the well-powered GWAS of externalizing behavior in European Americans with PRS from a smaller, ancestry-matched GWAS of externalizing behavior (Marquez-Luna et al., 2016). However, this method can also increase noise in the polygenic score, which, along with the smaller African American sample size, may have a negative impact on predictive power relative to using a well-powered ancestry-matched discovery sample. Unfortunately, the availability of well-powered GWAS for individuals of non-European descent is substantially limited (Popejoy & Fullerton, 2016; Sirugo et al., 2019), and in turn the predictive power and utility of polygenic scores is highly disparate across ancestry groups (Martin et al., 2019). However, analyzing ancestry groups separately is just one way to address the computational complexity of conducting genetic analyses in diverse ancestry groups (Peterson et al., 2019). An alternative approach is to analyze European Americans and African Americans in a combined model, covarying for race/ethnicity to account

for differences in genetic architecture due to ancestry. Given the similarity in the pattern of findings between European Americans and African Americans (see Tables 9-14), analyzing the samples together may prove to be a useful approach to enhance statistical power to detect an effect while accounting for the impact of variation in genetic architecture through the inclusion of covariates. The combined approach would also allow for the inclusion of other racial/ethnic groups, such as individuals of Asian and Hispanic background, which were too small to analyze separately. This method would improve the applicability of GxI research findings for individuals of all backgrounds.

The inclusion of diverse samples in GxI research is critically important. Non-Europeans have long been underrepresented in genetic research due to a combination of historical and systemic factors (Dick et al., 2017; Popejoy & Fullerton, 2016). The history of the eugenics movement in the U.S. has had a lasting impact on trust in the scientific community, and concerns about confidentiality and the ways in which genetic data might be used (or misused) remain a concern to many individuals (Bates et al., 2005; Dick et al., 2017; Furr, 2002; Tambor et al., 2002). In addition, increased computational complexity of using diverse samples in genomic research (Peterson et al., 2019), and smaller populations from which to recruit participants further limit the representation of diverse ancestry in research (Dick et al., 2017). Genetic findings for complex traits conducted in one ancestry group often do not replicate in other ancestry groups due to differences in genetic architecture, such as linkage disequilibrium and allele frequency, and improper application of these findings could be harmful (Sirugo et al., 2019). Accordingly, the degree to which GxI findings are useful and applicable to all ancestry groups remains limited by the underrepresentation of non-European individuals in genomic

research. Large-scale efforts to diversify representation in genomic research are underway. Projects such as the All of Us Research Program and collaborative consortia (e.g., the Psychiatric Genomics Consortium) research may improve the ability to advance precision medicine for individuals of all backgrounds. Diversity, equity, and inclusion remain important factors to prioritize in future studies examining the degree to which underlying genetics contributes to differential intervention effects.

The present study has a number of strengths that fill critical gaps in the GxI literature. First, this study is one of a few GxI studies to integrate polygenic scores into prevention and intervention research. The use of polygenic scores rather than candidate genes helps to propel the literature toward the adoption of methods now understood to better represent genomic risk for complex traits/behaviors, and reduce risk of spurious findings through underpowered studies of single gene effects (Dick, 2018; Latendresse et al., 2018b; Musci & Schlomer, 2018b; Neale et al., 2020). Second, we examined long-term effects of an adolescent preventive intervention on young adult alcohol-related outcomes. Few prior GxI studies have focused on emerging adulthood, despite knowledge that it is a critical period for the development of alcohol use behaviors that set the stage for patterns of behavior in adulthood (Schulenberg & Maggs, 2002; Sussman & Arnett, 2014). Finally, this study explored potential mechanisms through which genetics may influence differential response to prevention and intervention. Based on estimates of sample sizes required for mediation analyses, the present study was adequately powered to detect halfway (.26), medium (.39), and large (.59) effect, but insufficiently powered to detect small (.14) effects (Fritz & MacKinnon, 2007). The few prior GxI studies that examined mediators observed large mediation effects (Brody et al., 2014, 2015); however, given the use of

candidate genes and likelihood of overestimating single variant main effect sizes, it is likely that small to very small effects are a more appropriate estimate of mediation effect sizes in GxI studies using polygenic scores. Despite the limitations of statistical power and null findings, the present study advances efforts to better understand why interventions may be more or less effective for individuals with varying levels of genetic risk.

In addition to the aforementioned strengths, the present study findings should be interpreted within the context of the following limitations. First, the present study examined GxI effects in European Americans and African Americans, but other racial/ethnic groups were not included in the analyses due to small numbers of individuals in those groups. As a result, the generalizability of these study findings to individuals from other genetic ancestry groups remains unknown. Second, although polygenic risk scores offer a number of benefits, they also have limitations that are important to consider. Current polygenic scores calculated using large, well-powered GWAS still account for only a small percent of the variance in complex psychiatric outcomes, typically about 5% of the variance in substance use behaviors (Clarke et al., 2017; Karlsson Linnér et al., 2019; Kranzler et al., 2019; Liu et al., 2019). Although the amount of variance is likely to increase as GWAS discovery sample sizes increase, it may take quite some times before substance use research can amass samples needed to produce polygenic risk scores comparable to estimates observed for others phenotypes such as educational attainment and schizophrenia (Barr et al., 2020; Dudbridge, 2013; Evangelou et al., 2018; Lee et al., 2018; Pardiñas et al., 2018). Current polygenic scores may improve our understanding of relative risk (i.e., comparing those with high PRS to low PRS), but they are less informative about absolute risk until they can account for a greater proportion of variance in the overall outcome. Findings

from other more well-powered health-related outcomes, such as cardiovascular disease, suggest modest but meaningful improvement in prediction of coronary artery disease with the addition of polygenic scores (Elliott et al., 2020). In the future, such models may help to identify individuals at elevated need of early prevention of alcohol and substance use problems, or assist patients and providers with selection of the optimal treatment approach. However, at present, the clinical impact of polygenic scores in GxI research is negligible and the utility of polygenic scores for clinical decision-making remains to be explored (Lewis & Green, 2021). Finally, as a measure of aggregate genetic risk, polygenic scores provide little insight into the biological processes that may lead to differential intervention effects. Therefore, if polygenic scores prove useful in predicting differential treatment effects, mediators will be critically important to increasing our understanding of how and why genetics may contribute to differential intervention effects.

Although the limitations of the present study may inhibit generalizability and clinical relevance of the current research, these limitations also spur insight into exciting future directions for further research. As noted previously, mediators are necessary to better understand the mechanisms through which genetics influence differential intervention effects. Both psychosocial and biological mediators are worthy of examination. Psychosocial mediators, such as use of protective behavioral strategies or self-regulation, may help to identify the active ingredients in intervention that can be harnessed to improve intervention effects for those at elevated risk. Biological mediators, such as DNA methylation and metabolic changes in the brain measured by functional magnetic resonance imaging (fMRI), may also help to explain the pathways through which polygenetic risk may lead to variations in intervention effects. For example, cognitive behavioral therapy has been shown to change functional neural activity in

patients who complete treatment, and these effects may differ for individuals with varying levels of genetic risk (Beauregard, 2014; Porto et al., 2009). Although the feasibility of such studies might be challenging due to current costs associated with collecting genotypic and neuroimaging data, opportunities may become increasingly possible as methods advance, and collaborative research consortium become the "norm" in the scientific community.

Another important future direction for this research is to test the study hypotheses in larger studies, to determine the degree to which GxI research remains a fruitful topic of study. Conducting GxI research in larger samples may help to clarify whether the present study findings are the result of Type II error (i.e., failure to detect an effect when there is one in the population), or consistent with the absence of an effect in the population. Most evidence-based psychosocial intervention eventually advance to large-scale randomized trials, and by partnering with principal investigators of these studies, it may be possible to retrospectively collect DNA samples from participants already enrolled in randomized-controlled trials. Using this approach, it may also be possible to prioritize studies involving diverse samples, which would help to expand the generalizability for this research to individuals from different ancestry backgrounds. The ClinicalTrials.gov database is a useful tool for identifying potential study samples to approach with collaborative research ideas to integrate genetics into prevention and intervention. Another possible way to amass larger GxI samples is through collaboration with biobanks with electronic medical records. For example, the Million Veteran Program (MVP) is large-scale, coordinated research effort to understand how genetics, lifestyle factors, and military experiences influence health and disease (Gaziano et al., 2016). Since it began, MVP has enrolled over 800,000 veterans and collected DNA samples through collaboration with Veterans Affairs (VA) Medical

Centers. The VA is also well-known for its leadership in the development, evaluation, and implementation of evidence-based psychological interventions for mental health problems. Countless randomized-controlled trials (RCTs) have been conducted in VA Medical Centers, and effective treatments are then disseminated to VA mental health clinics around the country. Accordingly, there is likely a wealth of both efficacy (via RCT studies) and effectiveness data (via real-world implementation of interventions) already available in the MVP biobank dataset. Although there are additional complexities to examining GxI effects in non-randomized samples, analytical methods, such as propensity score matching, may be able to mitigate the effects of selection bias in these studies (Austin, 2011; Guo & Fraser, 2010, 2014). By implementing these approaches, it may be possible to attain samples greater than N=1500, thereby increasing the statistical power to detect GxI effects and improving the generalizability of findings to diverse samples. If GxI effects are observed in larger samples, it will also be important to consider the clinical significance of those results. In order words, future studies are encouraged to consider whether the observed effects account for meaningful (rather than just statistically significant) differences between individuals with varying levels of genetic risk. With clinically meaningful effects, GxI research findings may enhance our ability to improve the effectiveness of prevention and treatment, and in turn, reduce the burden of alcohol and substance use problems.

Global Summary and Conclusions

The goal of this dissertation was to understand whether alcohol intervention outcomes varied as a function of genetic risk for alcohol problems, as indexed by polygenic risk associated with externalizing behaviors. The study also sought to test whether peer deviance and drinking motives mediated gene-by-intervention effects, which would shed light on mechanisms through which genetic risk might contribute to differential intervention effects. In order to understand whether findings generalized across different samples and interventions, we examined two existing datasets: Spit for Science (S4S), a longitudinal study of genetic and environmental influences on substance use and emotional health in college students, in which a subset of participants took part in online alcohol intervention, and Project Alliance (PAL), a community-based study in which participants took part in Family Check-Up, a strengths-based, family-centered intervention administered in middle school to promote family management and address child and adolescent adjustment problems.

In this dissertation study, all primary GxI analyses and mediation analyses across both the S4S and PAL samples resulted in null findings. In the S4S sample, a post-hoc analysis was conducted to explore more proximal, short-term effects of the intervention on AUD symptoms approximately five months post-intervention. Findings indicated that individuals with higher PRS in the intervention condition reported less reduction in AUD symptoms compared to individuals with lower PRS in the intervention condition. In the control group, there were no differences in AUD symptoms at varying levels of polygenic risk. The potential clinical implications of these findings are important, as this indicates a possible need to dedicate

additional resources to develop interventions that better address alcohol-related problems for individuals with higher polygenic risk for externalizing problems. For example, the intervention modality (web-based, group or individual), intensity, and content, as well as developmental timing of the intervention delivery are all potential avenues to explore with the goal of improving outcomes for individuals across levels of genetic risk. However, it is also important to note the possibility that this single significant GxI finding may be a statistical artifact resulting from scale transformations, which is a well-documented phenomenon in the broader GxE literature (Domingue et al., 2020; Eaves, 2006, 2017; Eaves et al., 1977; Eaves & Verhulst, 2014; Schwabe & van den Berg, 2014). Accordingly, we recommend that the significant findings from the S4S sample be interpreted with considerable caution.

The results presented in this dissertation study warrant a discussion of the merits of pursuing further GxI research given that the findings provided minimal evidence that response to intervention varies based on genetic risk. Although the S4S (n = 483) and PAL (EA: n = 269; AA: n = 191) samples used in this study are some of the largest to integrate polygenic scores into GxI analyses, statistical power remains a significant concern. Polygenic scores account for only a small percentage of the variance alcohol use behaviors (Liu et al., 2019; Sanchez-Roige et al., 2019; Walters et al., 2018; Zhou et al., 2020); coupled with the modest effects of alcohol prevention intervention programs (Huh et al., 2015), much larger samples are needed to confidently draw conclusions about the presence or absence of GxI effects. The samples in the present study are not sufficiently sized to conclude that there is no effect in the population. Future research would benefit from examining GxI effects on alcohol and substance related outcomes in larger, more diverse samples. This would both allow opportunity to explore whether

the present study findings are observed in other populations, and also address the need to increase diverse representation in genomic studies (Popejoy & Fullerton, 2016; Sirugo et al., 2019). In the event that significant GxI effects are detected in larger samples, it will also be important for researchers to consider possible mediating mechanisms of these effects, and the clinical significance of interaction findings. Large samples may increase the likelihood of detecting an interaction effect if there is one in the population; however, a statistically significant interaction may not translate to meaningful variation at a clinical level. This may present an important choice point for future researchers to consider, as the allocation of limited resources may have a larger clinical impact in other areas of research.

Finally, the ethics of integrating genetics into prevention and intervention research are worthy of discussion. Discoveries about conditional effects of genetic risk on prevention and intervention programs raises concerns about how this information could be used or misused in the context of healthcare. Some may worry that individuals will be declined treatment on the basis of their genetic risk profiles, while others may have concerns about the security of this data (Neale et al., 2020). The National Institutes of Health is invested in understanding these considerations, with a dedicated extramural funding program specifically for research on the ethical, legal, and social implications (ELSI) of genomics research. It is also very important to note that at present, the utility of using PRS to improve identification of individuals at increased risk for alcohol use problems is extremely limited, and thus these scores are not currently recommended for use in a clinical setting (Barr et al., 2020). However, as interest in genetics continues to rise, it is likely that polygenic scores will become integrated into our lives in new ways. Thus, it is imperative for researchers to understand whether genetics may interact with

intervention effects, and ensure that effective prevention and treatment approaches are available for all individuals.

In conclusion, the analyses in this dissertation represent an effort to fill a number of gaps in the GxI literature. By incorporating a novel polygenic score, studying emerging adult populations, including European American and African American samples, using propensity score matching to approximate random assignment, and examining longitudinal effects, this study contributes to our understanding of the role of genetic predispositions in differential intervention effects. Although larger, more diverse samples are needed to more comprehensively explore the research questions and detect very small GxI effect sizes, this study provides evidence of the feasibility of this research and presents a number of opportunities to explore the way that genetic risk may influence intervention effects in future studies.

References

Agrawal, A., Dick, D. M., Bucholz, K. K., Madden, P. A. F., Cooper, M. L., Sher, K. J., & Heath, A. C. (2008). Drinking expectancies and motives: A genetic study of young adult women. *Addiction (Abingdon, England)*, *103*(2), 194–204. https://doi.org/10.1111/j.1360-0443.2007.02074.x

Anderson, M. L., Chiswell, K., Peterson, E. D., Tasneem, A., Topping, J., & Califf, R. M. (2015). Compliance with Results Reporting at ClinicalTrials.gov. *The New England Journal of Medicine*, *372*(11), 1031–1039. https://doi.org/10.1056/NEJMsa1409364

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469–480. https://doi.org/10.1037/0003-066X.55.5.469

Arria, A. M., Garnier-Dykstra, L. M., Caldeira, K. M., Vincent, K. B., Winick, E. R., & O'Grady, K. E. (2013). Drug Use Patterns and Continuous Enrollment in College:Results From a Longitudinal Study. *Journal of Studies on Alcohol and Drugs*, *74*(1), 71–83.

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Auwera, S. V. der, Peyrot, W. J., Milaneschi, Y., Hertel, J., Baune, B., Breen, G., Byrne, E., Dunn, E. C., Fisher, H., Homuth, G., Levinson, D., Lewis, C., Mills, N., Mullins, N., Nauck, M., Pistis, G., Preisig, M., Rietschel, M., Ripke, S., … Grabe, H. (2018, January 1). *Genome-wide gene-environment interaction in depression: A systematic evaluation of*

candidate genes. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. https://doi.org/10.1002/ajmg.b.32593

Baer, J. S., Marlatt, G. A., Kivlahan, D. R., Fromme, K., Larimer, M. E., & Williams, E. (1992). An experimental test of three methods of alcohol risk reduction with young adults. *Journal of Consulting and Clinical Psychology*, *60*(6), 974–979.

Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2015). The Hidden Efficacy of Interventions: Gene×Environment Experiments from a Differential Susceptibility Perspective. *Annual Review of Psychology*, *66*(1), 381–409. https://doi.org/10.1146/annurev-psych-010814-015407

Barnett, N. P., Apodaca, T. R., Magill, M., Colby, S. M., Gwaltney, C., Rohsenow, D. J., & Monti, P. M. (2010). Moderators and Mediators of Two Brief Interventions for Alcohol in the Emergency Department. *Addiction (Abingdon, England)*, *105*(3), 452–465. https://doi.org/10.1111/j.1360-0443.2009.02814.x

Barr, P. B., Ksinan, A., Su, J., Johnson, E. C., Meyers, J. L., Wetherill, L., Latvala, A., Aliev, F., Chan, G., Kuperman, S., Nurnberger, J., Kamarajan, C., Anokhin, A., Agrawal, A., Rose, R. J., Edenberg, H. J., Schuckit, M., Kaprio, J., & Dick, D. M. (2020). Using polygenic scores for identifying individuals at increased risk of substance use disorders in clinical and population samples. *Translational Psychiatry*, *10*(1), 1–9. https://doi.org/10.1038/s41398-020-00865-8

Bates, B. R., Lynch, J. A., Bevan, J. L., & Condit, C. M. (2005). Warranted concerns, warranted outlooks: A focus group study of public understandings of genetic research. *Social*

Science & Medicine (1982), *60*(2), 331–344. https://doi.org/10.1016/j.socscimed.2004.05.012

Bauer, L. O., Covault, J., Harel, O., Das, S., Gelernter, J., Anton, R., & Kranzler, H. R. (2007). Variation in GABRA2 predicts drinking behavior in project MATCH subjects. *Alcoholism, Clinical and Experimental Research*, *31*(11), 1780–1787. https://doi.org/10.1111/j.1530-0277.2007.00517.x

Beach, S. R. H., Brody, G. H., Lei, M.-K., & Philibert, R. A. (2010). Differential susceptibility to parenting among African American youths: Testing the DRD4 hypothesis. *Journal of Family Psychology : JFP : Journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, *24*(5), 513–521. https://doi.org/10.1037/a0020835

Beauregard, M. (2014). Functional neuroimaging studies of the effects of psychotherapy. *Dialogues in Clinical Neuroscience*, *16*(1), 75–81.

Belsky, J., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). For Better and For Worse: Differential Susceptibility to Environmental Influences. *Current Directions in Psychological Science*, *16*(6), 300–304. https://doi.org/10.1111/j.1467-8721.2007.00525.x

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, *135*(6), 885–908. https://doi.org/10.1037/a0017376

Belsky, J., & van Ijzendoorn, M. H. (2015). What works for whom? Genetic moderation of intervention efficacy. *Development and Psychopathology*, *27*(1), 1–6. https://doi.org/10.1017/S0954579414001254

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Beseler, C. L., Aharonovich, E., Keyes, K. M., & Hasin, D. S. (2008). Adult Transition From At-Risk Drinking to Alcohol Dependence: The Relationship of Family History and Drinking Motives. *Alcoholism, Clinical and Experimental Research*, *32*(4). https://doi.org/10.1111/j.1530-0277.2008.00619.x

Border, R., Johnson, E. C., Evans, L. M., Smolen, A., Berley, N., Sullivan, P. F., & Keller, M. C. (2019). No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples. *American Journal of Psychiatry*, *176*(5), 376–387. https://doi.org/10.1176/appi.ajp.2018.18070881

Botvin, G. J., Baker, E., Dusenbury, L., Tortu, S., & Botvin, E. M. (1990). Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: Results of a 3-year study. *Journal of Consulting and Clinical Psychology*, *58*(4), 437–446.

Bravo, A. J., Prince, M. A., & Pearson, M. R. (2017). College-Related Alcohol Beliefs and Problematic Alcohol Consumption: Alcohol Protective Behavioral Strategies as a Mediator. *Substance Use & Misuse*, *52*(8), 1059–1068. https://doi.org/10.1080/10826084.2016.1271985

Brody, G. H., Beach, S. R. H., Hill, K. G., Howe, G. W., Prado, G., & Fullerton, S. M. (2013). Using genetically informed, randomized prevention trials to test etiological hypotheses about child and adolescent drug use and psychopathology. *American Journal of Public Health*, *103 Suppl 1*, S19-24. https://doi.org/10.2105/AJPH.2012.301080

Brody, G. H., Beach, S. R. H., Philibert, R. A., Chen, Y., & Murry, V. M. (2009). Prevention effects moderate the association of 5-HTTLPR and youth risk behavior initiation: Gene x environment hypotheses tested via a randomized prevention design. *Child Development*, *80*(3), 645–661. https://doi.org/10.1111/j.1467-8624.2009.01288.x

Brody, G. H., Chen, Y.-F., Beach, S. R. H., Kogan, S. M., Yu, T., Diclemente, R. J., Wingood, G. M., Windle, M., & Philibert, R. A. (2014). Differential sensitivity to prevention programming: A dopaminergic polymorphism-enhanced prevention effect on protective parenting and adolescent substance use. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *33*(2), 182–191. https://doi.org/10.1037/a0031253

Brody, G. H., Chen, Y.-F., Beach, S. R. H., Philibert, R. A., & Kogan, S. M. (2009). Participation in a family-centered prevention program decreases genetic risk for adolescents' risky behaviors. *Pediatrics*, *124*(3), 911–917. https://doi.org/10.1542/peds.2008-3464

Brody, G. H., Yu, T., & Beach, S. R. H. (2015). A differential susceptibility analysis reveals the "who and how" about adolescents' responses to preventive interventions: Tests of first- and second-generation Gene x Intervention hypotheses. *Development and Psychopathology*, *27*(1), 37–49. https://doi.org/10.1017/S095457941400128X

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834. https://doi.org/10.1016/j.neuron.2010.11.022

Bucholz, K. K., Cadoret, R., Cloninger, C. R., Dinwiddie, S. H., Hesselbrock, V. M., Nurnberger, J. I., Reich, T., Schmidt, I., & Schuckit, M. A. (1994). A new, semi-structured psychiatric interview for use in genetic linkage studies: A report on the reliability of the SSAGA. *Journal of Studies on Alcohol*, *55*(2), 149–158.

Carey, K. B., Henson, J. M., Carey, M. P., & Maisto, S. A. (2010). Perceived Norms Mediate Effects of a Brief Motivational Intervention for Sanctioned College Drinkers. *Clinical Psychology : A Publication of the Division of Clinical Psychology of the American Psychological Association*, *17*(1), 58–71. https://doi.org/10.1111/j.1468-2850.2009.01194.x

Carey, K. B., Scott-Sheldon, L. A. J., Carey, M. P., & DeMartini, K. S. (2007). Individual-Level Interventions to Reduce College Student Drinking: A Meta-Analytic Review. *Addictive Behaviors*, *32*(11), 2469–2494. https://doi.org/10.1016/j.addbeh.2007.05.004

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7. https://doi.org/10.1186/s13742-015-0047-8

Clarke, T.-K., Adams, M. J., Davies, G., Howard, D. M., Hall, L. S., Padmanabhan, S., Murray, A. D., Smith, B. H., Campbell, A., Hayward, C., Porteous, D. J., Deary, I. J., & McIntosh, A. M. (2017). Genome-wide association study of alcohol consumption and

genetic overlap with other health-related traits in UK Biobank (*N*=112 117). *Molecular Psychiatry*, *22*(10), 1376–1384. https://doi.org/10.1038/mp.2017.153

Cleveland, H. H., Schlomer, G. L., Vandenbergh, D. J., Feinberg, M., Greenberg, M., Spoth, R., Redmond, C., Shriver, M. D., Zaidi, A. A., & Hair, K. L. (2015). The conditioning of intervention effects on early adolescent alcohol use by maternal involvement and dopamine receptor D4 (DRD4) and serotonin transporter linked polymorphic region (5-HTTLPR) genetic variants. *Development and Psychopathology*, *27*(1), 51–67. https://doi.org/10.1017/S0954579414001291

Connell, A. M., Dishion, T. J., Yasui, M., & Kavanagh, K. (2007). An adaptive approach to family intervention: Linking engagement in family-centered intervention to reductions in adolescent problem behavior. *Journal of Consulting and Clinical Psychology*, *75*(4), 568–579. https://doi.org/10.1037/0022-006X.75.4.568

Connell, A. M., Klostermann, S., & Dishion, T. J. (2012). Family Check Up Effects on Adolescent Arrest Trajectories: Variation by Developmental Subtype. *Journal of Research on Adolescence*, *22*(2), 367–380. https://doi.org/10.1111/j.1532-7795.2011.00765.x

Cooper, M. L. (1994). Motivations for alcohol use among adolescents: Development and validation of a four-factor model. *Psychological Assessment*, *6*(2), 117–128. https://doi.org/10.1037/1040-3590.6.2.117

Cronce, J. M., & Larimer, M. E. (2011). Individual-focused approaches to the prevention of college student drinking. *Alcohol Research & Health*, *34*(2), 210.

Crouch, D. J. M., & Bodmer, W. F. (2020). Polygenic inheritance, GWAS, polygenic risk scores, and the search for functional variants. *Proceedings of the National Academy of Sciences*, *117*(32), 18924–18933. https://doi.org/10.1073/pnas.2005634117

Daurio, A. M., Deschaine, S. L., Modabbernia, A., & Leggio, L. (2020). Parsing out the role of dopamine D4 receptor gene (DRD4) on alcohol-related phenotypes: A meta-analysis and systematic review. *Addiction Biology*, *25*(3), e12770. https://doi.org/10.1111/adb.12770

Davey Smith, G., & Ebrahim, S. (2003). 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease?*. *International Journal of Epidemiology*, *32*(1), 1–22. https://doi.org/10.1093/ije/dyg070

Delaneau, O., Zagury, J.-F., & Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nature Methods*, *10*(1), 5–6. https://doi.org/10.1038/nmeth.2307

DeMartini, K. S., & Carey, K. B. (2012). Optimizing the use of the AUDIT for alcohol screening in college students. *Psychological Assessment*, *24*(4), 954.

Dick, D. M. (2011). Gene-Environment Interaction in Psychological Traits and Disorders. *Annual Review of Clinical Psychology*, *7*, 383–409. https://doi.org/10.1146/annurev-clinpsy-032210-104518

Dick, D. M. (2018). Commentary for special issue of prevention science "using genetics in prevention: Science fiction or science fact?" *Prevention Science*, *19*(1), 101–108.

Dick, D. M., & Agrawal, A. (2008). The genetics of alcohol and other drug dependence. *Alcohol Research & Health*, *31*(2), 111.

Dick, D. M., Agrawal, A., Keller, M. C., Adkins, A., Aliev, F., Monroe, S., Hewitt, J. K., Kendler, K. S., & Sher, K. J. (2015). Candidate gene-environment interaction research: Reflections and recommendations. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *10*(1), 37–59. https://doi.org/10.1177/1745691614556682

Dick, D. M., Barr, P. B., Cho, S. B., Cooke, M. E., Kuo, S. I., Lewis, T. J., Neale, Z., Salvatore, J. E., Savage, J., & Su, J. (2018). Post-GWAS in psychiatric genetics: A developmental perspective on the "other" next steps. *Genes, Brain & Behavior*, *17*(3), No Pagination Specified-No Pagination Specified. https://doi.org/10.1111/gbb.12447

Dick, D. M., Barr, P., Guy, M., Nasim, A., & Scott, D. (2017). Review: Genetic research on alcohol use outcomes in African American populations: A review of the literature, associated challenges, and implications. *The American Journal on Addictions*, *26*(5), 486–493. https://doi.org/10.1111/ajad.12495

Dick, D. M., Cho, S. B., Latendresse, S. J., Aliev, F., Nurnberger, J. I., Edenberg, H. J., Schuckit, M., Hesselbrock, V. M., Porjesz, B., Bucholz, K., Wang, J.-C., Goate, A., Kramer, J. R., & Kuperman, S. (2014). Genetic Influences on Alcohol Use across Stages of Development: GABRA2 and Longitudinal Trajectories of Drunkenness from Adolescence to Young Adulthood. *Addiction Biology*, *19*(6), 1055–1064. https://doi.org/10.1111/adb.12066

Dick, D. M., & Hancock, L. C. (2015). Integrating basic research with prevention/intervention to reduce risky substance use among college students. *Frontiers in Psychology*, *6*, 544.

Dick, D. M., & Kendler, K. S. (2012). The Impact of Gene–Environment Interaction on Alcohol Use Disorders. *Alcohol Research : Current Reviews*, *34*(3), 318–324.

Dick, D. M., Nasim, A., Edwards, A. C., Salvatore, J., Cho, S. B., Adkins, A., Meyers, J., Yan, J., Cooke, M., & Clifford, J. (2014). Spit for Science: Launching a longitudinal study of genetic and environmental influences on substance use and emotional health at a large US university. *Frontiers in Genetics*, *5*, 47.

Dick, D. M., Pagan, J. L., Viken, R., Purcell, S., Kaprio, J., Pulkkinen, L., & Rose, R. J. (2007). Changing environmental influences on substance use across development. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies*, *10*(2), 315–326. https://doi.org/10.1375/twin.10.2.315

Dishion, T. J., Nelson, S. E., & Kavanagh, K. (2003). The family check-up with high-risk young adolescents: Preventing early-onset substance use by parent monitoring. *Behavior Therapy*, *34*(4), 553–571. https://doi.org/10.1016/S0005-7894(03)80035-7

Dishion, T. J., & Stormshak, E. A. (2007). *Intervening in children's lives: An ecological, family-centered approach to mental health care*. American Psychological Association. https://doi.org/10.1037/11485-000

Dodge, K. A., Dishion, T. J., & Lansford, J. E. (2006). Deviant Peer Influences in Intervention and Public Policy for Youth. *Social Policy Report*, *20*(1), 1–20. https://doi.org/10.1002/j.2379-3988.2006.tb00046.x

Domingue, B. W., Trejo, S., Armstrong-Carter, E., & Tucker-Drob, E. M. (2020). Interactions between Polygenic Scores and Environments: Methodological and Conceptual Challenges. *Sociological Science*, *7*, 465–486. https://doi.org/10.15195/v7.a19

Dudbridge, F. (2013). Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genetics*, *9*(3), Article 3. https://doi.org/10.1371/journal.pgen.1003348

Dudbridge, F., Pashayan, N., & Yang, J. (2018). Predictive accuracy of combined genetic and environmental risk scores. *Genetic Epidemiology*, *42*(1), 4–19. https://doi.org/10.1002/gepi.22092

Duncan, L. E., & Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *The American Journal of Psychiatry*, *168*(10), 1041–1049. https://doi.org/10.1176/appi.ajp.2011.11020191

Eaves, L. (2006). Genotype × Environment Interaction in Psychopathology: Fact or Artifact? *Twin Research and Human Genetics*, *9*(1), 1–8. https://doi.org/10.1375/twin.9.1.1

Eaves, L. (2017). Genotype × Environment Interaction in Psychiatric Genetics: Deep Truth or Thin Ice? *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies*, *20*(3), 187–196. https://doi.org/10.1017/thg.2017.19

Eaves, L., Last, K., Martin, N. G., & Jinks, J. L. (1977). A progressive approach to non-additivity and genotype-environmental covariance in the analysis of human differences. *British Journal of Mathematical and Statistical Psychology*, *30*(1), 1–42. https://doi.org/10.1111/j.2044-8317.1977.tb00722.x

Eaves, L., & Verhulst, B. (2014). Problems and Pit-Falls in Testing for G × E and Epistasis in Candidate Gene Studies of Human Behavior. *Behavior Genetics*, *44*(6), 578–590. https://doi.org/10.1007/s10519-014-9674-6

Edwards, A. C., & Kendler, K. S. (2013). Alcohol consumption in men is influenced by qualitatively different genetic factors in adolescence and adulthood. *Psychological Medicine*, *43*(9), 1857–1868. https://doi.org/10.1017/S0033291712002917

Elam, K. K., Clifford, S., Ruof, A., Shaw, D. S., Wilson, M. N., & Lemery-Chalfant, K. (2020). Genotype–environment correlation by intervention effects underlying middle childhood peer rejection and associations with adolescent marijuana use. *Development and Psychopathology*, 1–12. https://doi.org/10.1017/S0954579420001066

Elliott, J., Bodinier, B., Bond, T. A., Chadeau-Hyam, M., Evangelou, E., Moons, K. G. M., Dehghan, A., Muller, D. C., Elliott, P., & Tzoulaki, I. (2020). Predictive Accuracy of a Polygenic Risk Score-Enhanced Prediction Model vs a Clinical Risk Score for Coronary Artery Disease. *JAMA*, *323*(7), 636–645. https://doi.org/10.1001/jama.2019.22241

Evangelou, E., Warren, H. R., Mosen-Ansorena, D., Mifsud, B., Pazoki, R., Gao, H., Ntritsos, G., Dimou, N., Cabrera, C. P., Karaman, I., Ng, F. L., Evangelou, M., Witkowska, K., Tzanis, E., Hellwege, J. N., Giri, A., Velez Edwards, D. R., Sun, Y. V., Cho, K., … Caulfield, M. J. (2018). Genetic analysis of over one million people identifies 535 new loci associated with blood pressure traits. *Nature Genetics*, *50*(10), 1412–1425. https://doi.org/10.1038/s41588-018-0205-x

Ewing, S. W. F., LaChance, H. A., Bryan, A., & Hutchison, K. E. (2009). Do genetic and individual risk factors moderate the efficacy of motivational enhancement therapy? Drinking outcomes with an emerging adult sample. *Addiction Biology*, *14*(3), 356–365. https://doi.org/10.1111/j.1369-1600.2009.00149.x

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Feldstein Ewing, S. W., LaChance, H. A., Bryan, A., & Hutchison, K. E. (2009). Do genetic and individual risk factors moderate the efficacy of motivational enhancement therapy? Drinking outcomes with an emerging adult sample. *Addiction Biology*, *14*(3), 356–365. https://doi.org/10.1111/j.1369-1600.2009.00149.x

Foxcroft, D. R., & Tsertsvadze, A. (2011). Universal multi-component prevention programs for alcohol misuse in young people. *Cochrane Database of Systematic Reviews*, *9*. https://doi.org/10.1002/14651858.CD009307

Fritz, M. S., & MacKinnon, D. P. (2007). Required Sample Size to Detect the Mediated Effect. *Psychological Science*, *18*(3), 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x

Furr, L. A. (2002). Perceptions of genetics research as harmful to society: Differences among samples of African-Americans and European-Americans. *Genetic Testing*, *6*(1), 25–30. https://doi.org/10.1089/109065702760093889

Gaume, J., McCambridge, J., Bertholet, N., & Daeppen, J.-B. (2014). Mechanisms of Action of Brief Alcohol Interventions Remain Largely Unknown – A Narrative Review. *Frontiers in Psychiatry*, *5*. https://doi.org/10.3389/fpsyt.2014.00108

Gaziano, J. M., Concato, J., Brophy, M., Fiore, L., Pyarajan, S., Breeling, J., Whitbourne, S., Deen, J., Shannon, C., Humphries, D., Guarino, P., Aslan, M., Anderson, D., LaFleur, R., Hammond, T., Schaa, K., Moser, J., Huang, G., Muralidhar, S., … O'Leary, T. J. (2016).

Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *Journal of Clinical Epidemiology*, *70*, 214–223. https://doi.org/10.1016/j.jclinepi.2015.09.016

Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A., & Smoller, J. W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications*, *10*(1), 1776. https://doi.org/10.1038/s41467-019-09718-5

Gelernter, J., Kranzler, H. R., Sherva, R., Almasy, L., Koesterer, R., Smith, A. H., Anton, R., Preuss, U. W., Ridinger, M., Rujescu, D., Wodarz, N., Zill, P., Zhao, H., & Farrer, L. A. (2014a). Genome-wide association study of alcohol dependence:significant findings in African- and European-Americans including novel risk loci. *Molecular Psychiatry*, *19*(1), 41–49. https://doi.org/10.1038/mp.2013.145

Gelernter, J., Kranzler, H. R., Sherva, R., Almasy, L., Koesterer, R., Smith, A. H., Anton, R., Preuss, U. W., Ridinger, M., Rujescu, D., Wodarz, N., Zill, P., Zhao, H., & Farrer, L. A. (2014b). Genome-wide association study of alcohol dependence:significant findings in African- and European-Americans including novel risk loci. *Molecular Psychiatry*, *19*(1), 41–49. https://doi.org/10.1038/mp.2013.145

Gillespie, N. A., Neale, M. C., Jacobson, K., & Kendler, K. S. (2009). Modeling the genetic and environmental association between peer group deviance and cannabis use in male twins. *Addiction (Abingdon, England)*, *104*(3), 420–429. https://doi.org/10.1111/j.1360-0443.2008.02457.x

Gilligan, C., Wolfenden, L., Foxcroft, D. R., Williams, A. J., Kingsland, M., Hodder, R. K., Stockings, E., McFadyen, T.-R., Tindall, J., Sherker, S., Rae, J., & Wiggers, J. (2019).

Family-based prevention programmes for alcohol use in young people. *Cochrane Database of Systematic Reviews*, *3*. https://doi.org/10.1002/14651858.CD012287.pub2

Gotham, H. J., Sher, K. J., & Wood, P. K. (1997). Predicting stability and change in frequency of intoxication from the college years to beyond: Individual-difference and role transition variables. *Journal of Abnormal Psychology*, *106*(4), 619–629.

Grant, B. F., Chou, S. P., Saha, T. D., Pickering, R. P., Kerridge, B. T., Ruan, W. J., Huang, B., Jung, J., Zhang, H., Fan, A., & Hasin, D. S. (2017). Prevalence of 12-Month Alcohol Use, High-Risk Drinking, and DSM-IV Alcohol Use Disorder in the United States, 2001-2002 to 2012-2013. *JAMA Psychiatry*, *74*(9), 911–923. https://doi.org/10.1001/jamapsychiatry.2017.2161

Grant, B. F., Dawson, D. A., Stinson, F. S., Chou, S. P., Dufour, M. C., & Pickering, R. P. (2004). The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991-1992 and 2001-2002. *Drug and Alcohol Dependence*, *74*(3), 223–234. https://doi.org/10.1016/j.drugalcdep.2004.02.004

Grant, B. F., Goldstein, R. B., Saha, T. D., Chou, S. P., Jung, J., Zhang, H., Pickering, R. P., Ruan, W. J., Smith, S. M., Huang, B., & Hasin, D. S. (2015). Epidemiology of DSM-5 Alcohol Use Disorder: Results From the National Epidemiologic Survey on Alcohol and Related Conditions III. *JAMA Psychiatry*, *72*(8), 757–766. https://doi.org/10.1001/jamapsychiatry.2015.0584

Griffin, K. W., & Botvin, G. J. (2010). Evidence-Based Interventions for Preventing Substance Use Disorders in Adolescents. *Child and Adolescent Psychiatric Clinics of North America*, *19*(3), 505–526. https://doi.org/10.1016/j.chc.2010.03.005

Griswold, M. G., Fullman, N., Hawley, C., Arian, N., Zimsen, S. R. M., Tymeson, H. D., Venkateswaran, V., Tapp, A. D., Forouzanfar, M. H., Salama, J. S., Abate, K. H., Abate, D., Abay, S. M., Abbafati, C., Abdulkader, R. S., Abebe, Z., Aboyans, V., Abrar, M. M., Acharya, P., … Gakidou, E. (2018). Alcohol use and burden for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *392*(10152), 1015–1035. https://doi.org/10.1016/S0140-6736(18)31310-2

Group, C. P. P. R. (1992). A developmental and clinical model for the prevention of conduct disorder: The FAST Track Program. *Development and Psychopathology*, *4*(4), 509–527. https://doi.org/10.1017/S0954579400004855

Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. Sage Publications, Inc.

Guo, S., & Fraser, M. W. (2014). *Propensity Score Analysis: Statistical Methods and Applications*. SAGE Publications.

Hansen, W. B., & Graham, J. W. (1991). Preventing alcohol, marijuana, and cigarette use among adolescents: Peer pressure resistance training versus establishing conservative norms. *Preventive Medicine*, *20*(3), 414–430. https://doi.org/10.1016/0091-7435(91)90039-7

Harden, K. P., Hill, J. E., Turkheimer, E., & Emery, R. E. (2008). Gene-Environment Correlation and Interaction in Peer Effects on Adolescent Alcohol and Tobacco Use. *Behavior Genetics*, *38*(4), 339–347. https://doi.org/10.1007/s10519-008-9202-7

Hart, A. B., & Kranzler, H. R. (2015). Alcohol Dependence Genetics: Lessons Learned From Genome-Wide Association Studies (GWAS) and Post-GWAS Analyses. *Alcoholism,*

Clinical and Experimental Research, *39*(8), 1312–1327. https://doi.org/10.1111/acer.12792

Heils, A., Teufel, A., Petri, S., Stöber, G., Riederer, P., Bengel, D., & Lesch, K. P. (1996). Allelic Variation of Human Serotonin Transporter Gene Expression. *Journal of Neurochemistry*, *66*(6), 2621–2624. https://doi.org/10.1046/j.1471-4159.1996.66062621.x

Hicks, B. M., Krueger, R. F., Iacono, W. G., McGue, M., & Patrick, C. J. (2004). Family transmission and heritability of externalizing disorders: A twin-family study. *Archives of General Psychiatry*, *61*(9), 922–928. https://doi.org/10.1001/archpsyc.61.9.922

Hingson, R. W., Zha, W., & Weitzman, E. R. (2009). Magnitude of and Trends in Alcohol-Related Mortality and Morbidity Among U.S. College Students Ages 18-24, 1998-2005. *Journal of Studies on Alcohol and Drugs. Supplement*, *16*, 12–20.

Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, *42*(1), 1–28. https://doi.org/10.18637/jss.v042.i08

Ho, T., Krishna, D., Dick, D. M., Adkins, A. E., & Barr, P. B. (2016). *Risk factors associated with first-year college dropout* [Poster]. Virginia Commonwealth University Poster Symposium for Undergraduate Research and Creativity.

Homberg, J. R., & van den Hove, D. L. A. (2012). The serotonin transporter gene and functional and pathological adaptation to environmental variation across the life span. *Progress in Neurobiology*, *99*(2), 117–127. https://doi.org/10.1016/j.pneurobio.2012.08.003

Hong, E. P., & Park, J. W. (2012). Sample Size and Statistical Power Calculation in Genetic Association Studies. *Genomics & Informatics*, *10*(2), 117–122. https://doi.org/10.5808/GI.2012.10.2.117

Howie, B. N., Donnelly, P., & Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genetics*, *5*(6), e1000529. https://doi.org/10.1371/journal.pgen.1000529

Huh, D., Mun, E.-Y., Larimer, M. E., White, H. R., Ray, A. E., Rhew, I. C., Kim, S.-Y., Jiao, Y., & Atkins, D. C. (2015). Brief motivational interventions for college student drinking may not be as powerful as we think: An individual participant-level data meta-analysis. *Alcoholism, Clinical and Experimental Research*, *39*(5), 919–931. https://doi.org/10.1111/acer.12714

International HapMap Consortium. (2003). The International HapMap Project. *Nature*, *426*(6968), 789–796. https://doi.org/10.1038/nature02168

International Schizophrenia Consortium, Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., & Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*(7256), 748–752. https://doi.org/10.1038/nature08185

Jackson, K. M., Sher, K. J., Gotham, H. J., & Wood, P. K. (2001). Transitioning into and out of large-effect drinking in young adulthood. *Journal of Abnormal Psychology*, *110*(3), 378–391.

Jinks, J. L., & Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of the human behavior. *Psychological Bulletin*, *73*(5), 311–349. https://doi.org/10.1037/h0029135

Johnson, E. C., Border, R., Melroy-Greif, W. E., de Leeuw, C. A., Ehringer, M. A., & Keller, M. C. (2017). No Evidence That Schizophrenia Candidate Genes Are More Associated With Schizophrenia Than Noncandidate Genes. *Biological Psychiatry*, *82*(10), 702–708. https://doi.org/10.1016/j.biopsych.2017.06.033

Kabacoff, R. (2011). *R in action: Data analysis and graphics with R*. Manning.

Karlsson Linnér, R., Biroli, P., Kong, E., Meddens, S. F. W., Wedow, R., Fontana, M. A., Lebreton, M., Tino, S. P., Abdellaoui, A., Hammerschlag, A. R., Nivard, M. G., Okbay, A., Rietveld, C. A., Timshel, P. N., Trzaskowski, M., Vlaming, R. de, Zünd, C. L., Bao, Y., Buzdugan, L., … Beauchamp, J. P. (2019). Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nature Genetics*, *51*(2), 245. https://doi.org/10.1038/s41588-018-0309-3

Karlsson Linnér, R., Mallard, T. T., Barr, P. B., Sanchez-Roige, S., Madole, J. W., Driver, M. N., Poore, H. E., Grotzinger, A. D., Tielbeek, J. J., Johnson, E. C., Liu, M., Zhou, H., Kember, R. L., Pasman, J. A., Verweij, K. J. H., Liu, D. J., Vrieze, S., Collaborators, C., Kranzler, H. R., … Dick, D. M. (2020). Multivariate genomic analysis of 1.5 million people identifies genes related to addiction, antisocial behavior, and health. *BioRxiv*, 2020.10.16.342501. https://doi.org/10.1101/2020.10.16.342501

Keers, R., Coleman, J. R. I., Lester, K. J., Roberts, S., Breen, G., Thastum, M., Bögels, S., Schneider, S., Heiervang, E., Meiser-Stedman, R., Nauta, M., Creswell, C., Thirlwall, K., Rapee, R. M., Hudson, J. L., Lewis, C., Plomin, R., & Eley, T. C. (2016). A Genome-Wide Test of the Differential Susceptibility Hypothesis Reveals a Genetic Predictor of Differential Response to Psychological Treatments for Child Anxiety Disorders. *Psychotherapy and Psychosomatics*, *85*(3), 146–158. https://doi.org/10.1159/000444023

Kelpin, S. S., Ondersma, S. J., Weaver, M., & Svikis, D. S. (2018). Representativeness of patients enrolled in a primary care clinical trial for heavy/problem substance use. *Substance Abuse*, *39*(4), 469–475. https://doi.org/10.1080/08897077.2018.1526843

Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: A systematic review. *Psychological Medicine*, *37*(5), 615–626. https://doi.org/10.1017/S0033291706009524

Kendler, K. S., Gardner, C., & Dick, D. M. (2011). Predicting alcohol consumption in adolescence from alcohol-specific and general externalizing genetic risk factors, key environmental exposures and their interaction. *Psychological Medicine*, *41*(7), 1507–1516. https://doi.org/10.1017/S003329171000190X

Kendler, K. S., Jacobson, K. C., Gardner, C. O., Gillespie, N., Aggen, S. A., & Prescott, C. A. (2007). Creating a Social World. *Archives of General Psychiatry*, *64*(8), 958–965. https://doi.org/10.1001/archpsyc.64.8.958

Kendler, K. S., Jacobson, K., Myers, J. M., & Eaves, L. J. (2008). A genetically informative developmental study of the relationship between conduct disorder and peer deviance in

males. *Psychological Medicine*, *38*(7), 1001–1011. https://doi.org/10.1017/S0033291707001821

Kendler, K. S., Jaffee, S., & Romer, D. (Eds.). (2011). *The dynamic genome and mental health: The role of genes and environments in youth development*. Oxford University Press.

Kendler, K. S., & Myers, J. (2014). The boundaries of the internalizing and externalizing genetic spectra in men and women. *Psychological Medicine*, *44*(3). https://doi.org/10.1017/S0033291713000585

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The Structure of Genetic and Environmental Risk Factors for Common Psychiatric and Substance Use Disorders in Men and Women. *Archives of General Psychiatry*, *60*(9), 929–937. https://doi.org/10.1001/archpsyc.60.9.929

Kendler, K. S., Schmitt, E., Aggen, S. H., & Prescott, C. A. (2008). Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Archives of General Psychiatry*, *65*(6), 674–682. https://doi.org/10.1001/archpsyc.65.6.674

Kessler, R. C., & Ustün, T. B. (2004). The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 93–121.

Kivlahan, D. R., Marlatt, G. A., Fromme, K., Coppel, D. B., & Williams, E. (1990). Secondary prevention with college drinkers: Evaluation of an alcohol skills training program. *Journal of Consulting and Clinical Psychology*, *58*(6), 805–810.

Kranzler, H. R., Zhou, H., Kember, R. L., Smith, R. V., Justice, A. C., Damrauer, S., Tsao, P. S., Klarin, D., Baras, A., Reid, J., Overton, J., Rader, D. J., Cheng, Z., Tate, J. P., Becker, W. C., Concato, J., Xu, K., Polimanti, R., Zhao, H., & Gelernter, J. (2019). Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nature Communications*, *10*(1), 1499. https://doi.org/10.1038/s41467-019-09480-8

Krueger, R. F., Hicks, B. M., Patrick, C. J., Carlson, S. R., Iacono, W. G., & McGue, M. (2002). Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology*, *111*(3), 411–424.

Kuo, S. I.-C., Salvatore, J. E., Aliev, F., Ha, T., Dishion, T. J., & Dick, D. M. (2019). The Family Check-up Intervention Moderates Polygenic Influences on Long-Term Alcohol Outcomes: Results from a Randomized Intervention Trial. *Prevention Science: The Official Journal of the Society for Prevention Research*. https://doi.org/10.1007/s11121-019-01024-2

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., … International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. https://doi.org/10.1038/35057062

Larimer, M. E., & Cronce, J. M. (2007). Identification, prevention, and treatment revisited: Individual-focused college drinking prevention strategies 1999–2006. *Addictive Behaviors*, *32*(11), 2439–2468.

Latendresse, S. J., Musci, R., & Maher, B. S. (2018a). Critical issues in the inclusion of genetic and epigenetic information in prevention and intervention trials. *Prevention Science*, *19*(1), 58–67.

Latendresse, S. J., Musci, R., & Maher, B. S. (2018b). Critical issues in the inclusion of genetic and epigenetic information in prevention and intervention trials. *Prevention Science*, *19*(1), 58–67.

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Linnér, R. K., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., … Cesarini, D. (2018). Gene discovery and polygenic prediction from a 1.1-million-person GWAS of educational attainment. *Nature Genetics*, *50*(8), 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Lemery-Chalfant, K., Clifford, S., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2018). Genetic Moderation of the Effects of the Family Check-Up Intervention on Children's Internalizing Symptoms: A Longitudinal Study with a Racial/Ethnically Diverse Sample. *Development and Psychopathology*, *30*(5), 1729–1747. https://doi.org/10.1017/S095457941800127X

Leung, R. K., Toumbourou, J. W., & Hemphill, S. A. (2014). The effect of peer influence and selection processes on adolescent alcohol use: A systematic review of longitudinal studies. *Health Psychology Review*, *8*(4), 426–457. https://doi.org/10.1080/17437199.2011.587961

Lewis, A. C. F., & Green, R. C. (2021). Polygenic risk scores in the clinic: New perspectives needed on familiar ethical issues. *Genome Medicine*, *13*(1), 14. https://doi.org/10.1186/s13073-021-00829-7

Lewis, M. A., & Neighbors, C. (2006). Social Norms Approaches Using Descriptive Drinking Norms Education: A Review of the Research on Personalized Normative Feedback. *Journal of American College Health : J of ACH*, *54*(4), 213–218.

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., Datta, G., Davila-Velderrain, J., McGuire, D., Tian, C., Zhan, X., Choquet, H., Docherty, A. R., Faul, J. D., Foerster, J. R., Fritsche, L. G., Gabrielsen, M. E., Gordon, S. D., Haessler, J., … Vrieze, S. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature Genetics*, *51*(2), 237. https://doi.org/10.1038/s41588-018-0307-5

Lochman, J. E., & van den Steenhoven, A. (2002). Family-Based Approaches to Substance Abuse Prevention. *Journal of Primary Prevention*, *23*(1), 49–114. https://doi.org/10.1023/A:1016591216363

MacArthur, G., Caldwell, D. M., Redmore, J., Watkins, S. H., Kipping, R., White, J., Chittleborough, C., Langford, R., Er, V., Lingam, R., Pasch, K., Gunnell, D., Hickman, M., & Campbell, R. (2018). Individual-, family-, and school-level interventions targeting multiple risk behaviours in young people. *Cochrane Database of Systematic Reviews*, *10*. https://doi.org/10.1002/14651858.CD009927.pub2

Magill, M., Kiluk, B. D., McCrady, B. S., Tonigan, J. S., & Longabaugh, R. (2015). Active Ingredients of Treatment and Client Mechanisms of Change in Behavioral Treatments for

Alcohol Use Disorders: Progress 10 Years Later. *Alcoholism: Clinical and Experimental Research*, *39*(10), 1852–1862. https://doi.org/10.1111/acer.12848

Maher, B. S. (2015). Polygenic Scores in Epidemiology: Risk Prediction, Etiology, and Clinical Utility. *Current Epidemiology Reports*, *2*(4), 239–244. https://doi.org/10.1007/s40471-015-0055-3

Marquez-Luna, C., Consortium, T. S. T. 2 D., & Price, A. L. (2016). Multi-ethnic polygenic risk scores improve risk prediction in diverse populations. *BioRxiv*, 051458. https://doi.org/10.1101/051458

Márquez-Luna, C., Loh, P.-R., South Asian Type 2 Diabetes (SAT2D) Consortium, SIGMA Type 2 Diabetes Consortium, & Price, A. L. (2017). Multiethnic polygenic risk scores improve risk prediction in diverse populations. *Genetic Epidemiology*, *41*(8), 811–823. https://doi.org/10.1002/gepi.22083

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, *51*(4), 584. https://doi.org/10.1038/s41588-019-0379-x

Mather, K., & Jinks, J. L. (1982). *Biometrical Genetics: The Study of Continuous Variation* (3rd ed.). Springer US. https://doi.org/10.1007/978-1-4899-3406-2

McGeary, J. (2009a). The DRD4 exon 3 VNTR polymorphism and addiction-related phenotypes: A review. *Pharmacology Biochemistry and Behavior*, *93*(3), 222–229. https://doi.org/10.1016/j.pbb.2009.03.010

McGeary, J. (2009b). The DRD4 exon 3 VNTR polymorphism and addiction-related phenotypes: A review. *Pharmacology Biochemistry and Behavior*, *93*(3), 222–229. https://doi.org/10.1016/j.pbb.2009.03.010

McNally, A. M., Palfai, T. P., & Kahler, C. W. (2005). Motivational interventions for heavy drinking college students: Examining the role of discrepancy-related psychological processes. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors*, *19*(1), 79–87. https://doi.org/10.1037/0893-164X.19.1.79

Metzler, C., Biglan, A., Ary, D., & Li, F. (1998). The Stability and Validity of Early Adolescents' Reports of Parenting Constructs. *Journal of Family Psychology*, *12*(4), 600–619.

Meyers, J. L., Salvatore, J. E., Vuoksimaa, E., Korhonen, T., Pulkkinen, L., Rose, R. J., Kaprio, J., & Dick, D. M. (2014). Genetic influences on alcohol use behaviors have diverging developmental trajectories: A prospective study among male and female twins. *Alcoholism, Clinical and Experimental Research*, *38*(11), 2869–2877. https://doi.org/10.1111/acer.12560

Miller, W. R., & Rollnick, S. (2002a). *Motivational interviewing: Preparing people for change, 2nd ed*. The Guilford Press.

Miller, W. R., & Rollnick, S. (2002b). *Motivational interviewing: Preparing people for change, 2nd ed*. The Guilford Press.

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin*, *110*(3), 406–425.

Morozova, T. V., Mackay, T. F. C., & Anholt, R. R. H. (2014). Genetics and genomics of alcohol sensitivity. *Molecular Genetics and Genomics*, *289*(3), 253–269. https://doi.org/10.1007/s00438-013-0808-y

Musci, R. J., Fairman, B., Masyn, K. E., Uhl, G., Maher, B., Sisto, D. Y., Kellam, S. G., & Ialongo, N. S. (2018). Polygenic score × intervention moderation: An application of discrete-time survival analysis to model the timing of first marijuana use among Urban youth. *Prevention Science*, *19*(1), 6–14. https://doi.org/10.1007/s11121-016-0729-1

Musci, R. J., Masyn, K. E., Uhl, G., Maher, B., Kellam, S. G., & Ialongo, N. S. (2015). Polygenic score x intervention moderation: An application of discrete-time survival analysis to modeling the timing of first tobacco use among urban youth. *Development and Psychopathology*, *27*(1), 111–122. https://doi.org/10.1017/S0954579414001333

Musci, R. J., & Schlomer, G. (2018a). The Implications of Genetics for Prevention and Intervention Programming. *Prevention Science*, 1–5.

Musci, R. J., & Schlomer, G. (2018b). The Implications of Genetics for Prevention and Intervention Programming. *Prevention Science*, 1–5.

Neale, Z. E., Kuo, S. I.-C., & Dick, D. M. (2020). A systematic review of gene-by-intervention studies of alcohol and other substance use. *Development and Psychopathology*, 1–18. https://doi.org/10.1017/S0954579420000590

Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S. E., Bishop, S., Cameron, D., Hamshere, M. L., Han, J., Hubbard, L., Lynham, A., Mantripragada, K., Rees, E., MacCabe, J. H., McCarroll, S. A., Baune, B. T., Breen, G., … Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-

intolerant genes and in regions under strong background selection. *Nature Genetics*, *50*(3), 381–389. https://doi.org/10.1038/s41588-018-0059-2

Pasman, J. A., Verweij, K. J. H., & Vink, J. M. (2019). Systematic Review of Polygenic Gene–Environment Interaction in Tobacco, Alcohol, and Cannabis Use. *Behavior Genetics*, *49*(4), 349–365. https://doi.org/10.1007/s10519-019-09958-7

Patel, M. X., Doku, V., & Tennakoon, L. (2003). Challenges in recruitment of research participants. *Advances in Psychiatric Treatment*, *9*(3), 229–238. https://doi.org/10.1192/apt.9.3.229

Patterson, N., Price, A. L., & Reich, D. (2006). Population Structure and Eigenanalysis. *PLOS Genetics*, *2*(12), e190. https://doi.org/10.1371/journal.pgen.0020190

Pearson, M. R. (2013). Use of alcohol protective behavioral strategies among college students: A critical review. *Clinical Psychology Review*, *33*(8), 1025–1040. https://doi.org/10.1016/j.cpr.2013.08.006

Peterson, R. E., Edwards, A. C., Bacanu, S.-A., Dick, D. M., Kendler, K. S., & Webb, B. T. (2017). The utility of empirically assigning ancestry groups in cross-population genetic studies of addiction. *The American Journal on Addictions*, *26*(5), 494–501. https://doi.org/10.1111/ajad.12586

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C.-Y., Popejoy, A. B., Periyasamy, S., Lam, M., Iyegbe, C., Strawbridge, R. J., Brick, L., Carey, C. E., Martin, A. R., Meyers, J. L., Su, J., Chen, J., Edwards, A. C., Kalungi, A., Koen, N., Majara, L., … Duncan, L. E. (2019). Genome-wide Association Studies in Ancestrally Diverse Populations:

Opportunities, Methods, Pitfalls, and Recommendations. *Cell*. https://doi.org/10.1016/j.cell.2019.08.051

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*(2), 309–322.

Pluess, M. (2017). Vantage Sensitivity: Environmental Sensitivity to Positive Experiences as a Function of Genetic Differences. *Journal of Personality*, *85*(1), 38–50. https://doi.org/10.1111/jopy.12218

Pluess, M., & Belsky, J. (2013). Vantage sensitivity: Individual differences in response to positive experiences. *Psychological Bulletin*, *139*(4), 901–916. https://doi.org/10.1037/a0030196

Popejoy, A. B., & Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature*, *538*(7624), 161–164. https://doi.org/10.1038/538161a

Porto, P. R., Oliveira, L., Mari, J., Volchan, E., Figueira, I., & Ventura, P. (2009). Does cognitive behavioral therapy change the brain? A systematic review of neuroimaging in anxiety disorders. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *21*(2), 114–125.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Prince, M. A., Carey, K. B., & Maisto, S. A. (2013). Protective behavioral strategies for reducing alcohol involvement: A review of the methodological issues. *Addictive Behaviors*, *38*(7), 2343–2351. https://doi.org/10.1016/j.addbeh.2013.03.010

Project MATCH Research Group. (1993). Project MATCH (Matching Alcoholism Treatment to Client Heterogeneity): Rationale and methods for a multisite clinical trial matching patients to alcoholism treatment. *Alcoholism, Clinical and Experimental Research*, *17*(6), 1130–1145.

Project MATCH Research Group. (1997). Project MATCH secondary a priori hypotheses. *Addiction*, *92*(12), 1671–1698. https://doi.org/10.1111/j.1360-0443.1997.tb02889.x

Ptáček, R., Kuželová, H., & Stefano, G. B. (2011). Dopamine D4 receptor gene DRD4 and its association with psychiatric disorders. *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research*, *17*(9), RA215–RA220. https://doi.org/10.12659/MSM.881925

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Rassen, J. A., Shelat, A. A., Myers, J., Glynn, R. J., Rothman, K. J., & Schneeweiss, S. (2012). One-to-many propensity score matching in cohort studies. *Pharmacoepidemiology and Drug Safety*, *21*(S2), 69–80. https://doi.org/10.1002/pds.3263

Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*(1), 41–55. JSTOR. https://doi.org/10.2307/2335942

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Salvatore, J. E., Aliev, F., Edwards, A. C., Evans, D. M., Macleod, J., Hickman, M., Lewis, G., Kendler, K. S., Loukola, A., Korhonen, T., Latvala, A., Rose, R. J., Kaprio, J., & Dick, D. M. (2014). Polygenic Scores Predict Alcohol Problems in an Independent Sample and Show Moderation by the Environment. *Genes*, *5*(2), 330–346. https://doi.org/10.3390/genes5020330

Salvatore, J. E., Thomas, N. S., Cho, S. B., Adkins, A., Kendler, K. S., & Dick, D. M. (2016). The role of romantic relationship status in pathways of risk for emerging adult alcohol use. *Psychology of Addictive Behaviors : Journal of the Society of Psychologists in Addictive Behaviors*, *30*(3), 335–344. https://doi.org/10.1037/adb0000145

Samson, J. E., & Tanner-Smith, E. E. (2015). Single-Session Alcohol Interventions for Heavy Drinking College Students: A Systematic Review and Meta-Analysis. *Journal of Studies on Alcohol and Drugs*, *76*(4), 530–543.

Sanchez-Roige, S., Palmer, A. A., Fontanillas, P., Elson, S. L., 23andMe Research Team, the Substance Use Disorder Working Group of the Psychiatric Genomics Consortium, Adams, M. J., Howard, D. M., Edenberg, H. J., Davies, G., Crist, R. C., Deary, I. J., McIntosh, A. M., & Clarke, T.-K. (2019). Genome-Wide Association Study Meta-Analysis of the Alcohol Use Disorders Identification Test (AUDIT) in Two Population-Based Cohorts. *The American Journal of Psychiatry*, *176*(2), 107–118. https://doi.org/10.1176/appi.ajp.2018.18040369

Sandler, I., Wolchik, S. A., Cruden, G., Mahrer, N. E., Ahn, S., Brincks, A., & Brown, C. H. (2014). Overview of Meta-Analyses of the Prevention of Mental Health, Substance Use

and Conduct Problems. *Annual Review of Clinical Psychology*, *10*, 243–273. https://doi.org/10.1146/annurev-clinpsy-050212-185524

Saraceno, L., Munafo, M., Heron, J., Craddock, N., & van den Bree, M. B. M. (2009). Genetic and non-genetic influences on the development of co-occurring alcohol problem use and internalizing symptomatology in adolescence: A review. *Addiction (Abingdon, England)*, *104*(7), 1100–1121. https://doi.org/10.1111/j.1360-0443.2009.02571.x

Savage, J. E., Neale, Z., Cho, S. B., Hancock, L., Kalmijn, J. A., Smith, T. L., Schuckit, M. A., Donovan, K. K., & Dick, D. M. (2015). Level of response to alcohol as a factor for targeted prevention in college students. *Alcoholism: Clinical and Experimental Research*, *39*(11), 2215–2223.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype greater than environment effects. *Child Development*, *54*(2), 424–435.

Schlomer, G. L., Cleveland, H. H., Feinberg, M. E., Wolf, P. S. A., Greenberg, M. T., Spoth, R. L., Redmond, C., Tricou, E. P., & Vandenbergh, D. J. (2017). Extending Previous cGxI Findings on 5-HTTLPR's Moderation of Intervention Effects on Adolescent Substance Misuse Initiation. *Child Development*, *88*(6), 2001–2012. https://doi.org/10.1111/cdev.12666

Schuckit, M. A. (2009). An overview of genetic influences in alcoholism. *Journal of Substance Abuse Treatment*, *36*(1), S5-14.

Schuckit, M. A. (2018). A critical review of methods and results in the search for genetic contributors to alcohol sensitivity. *Alcoholism: Clinical and Experimental Research*, *42*(5), 822–835. https://doi.org/10.1111/acer.13628

Schuckit, M. A., Kalmijn, J. A., Smith, T. L., Saunders, G., & Fromme, K. (2012). Structuring a college alcohol prevention program on the low level of response to alcohol model: A pilot study. *Alcoholism: Clinical and Experimental Research*, *36*(7), 1244–1252.

Schuckit, M. A., Smith, T. L., Kalmijn, J., Skidmore, J., Clausen, P., Shafir, A., Saunders, G., Bystritsky, H., & Fromme, K. (2015). The impact of focusing a program to prevent heavier drinking on a pre-existing phenotype, the low level of response to alcohol. *Alcoholism, Clinical and Experimental Research*, *39*(2), 308–316. https://doi.org/10.1111/acer.12620

Schuckit, M. A., Smith, T. L., & Tipp, J. E. (1997). The Self-Rating of the Effects of Alcohol (SRE) form as a retrospective measure of the risk for alcoholism. *Addiction*, *92*(8), 979–988.

Schulenberg, J. E., Johnston, L. D., O'Malley, P. M., Bachman, J. G., Miech, R. A., & Patrick, M. E. (2017). *Monitoring the Future national survey results on drug use, 1975-2016: Volume II, college students and adults ages 19-55*.

Schulenberg, J. E., Johnston, L. D., O'Malley, P. M., Bachman, J. G., Miech, R. A., & Patrick, M. E. (2018). Monitoring the Future National Survey Results on Drug Use, 1975-2017. Volume II, College Students & Adults Ages 19-55. *Institute for Social Research*.

Schulenberg, J. E., & Maggs, J. L. (2002). A developmental perspective on alcohol use and heavy drinking during adolescence and the transition to young adulthood. *Journal of Studies on Alcohol, Supplement*, *s14*, 54–70. https://doi.org/10.15288/jsas.2002.s14.54

Schwabe, I., & van den Berg, S. M. (2014). Assessing Genotype by Environment Interaction in Case of Heterogeneous Measurement Error. *Behavior Genetics*, *44*(4), 394–406. https://doi.org/10.1007/s10519-014-9649-7

Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). The Missing Diversity in Human Genetic Studies. *Cell*, *177*(1), 26–31. https://doi.org/10.1016/j.cell.2019.02.048

Skidmore, C. R., Kaufman, E. A., & Crowell, S. E. (2016). Substance Use Among College Students. *Child and Adolescent Psychiatric Clinics of North America*, *25*(4), 735–753. https://doi.org/10.1016/j.chc.2016.06.004

Smit, E., Verdurmen, J., Monshouwer, K., & Smit, F. (2008). Family interventions and their effect on adolescent alcohol use in general populations; a meta-analysis of randomized controlled trials. *Drug and Alcohol Dependence*, *97*(3), 195–206. https://doi.org/10.1016/j.drugalcdep.2008.03.032

So, H.-C., & Sham, P. C. (2017). Improving polygenic risk prediction from summary statistics by an empirical Bayes approach. *Scientific Reports*, *7*. https://doi.org/10.1038/srep41262

Soberman, L. (1994). *Psychometric validation of a brief teacher screening instrument* [Unpublished doctoral dissertation]. University of Oregon.

Spoth, R., Greenberg, M., & Turrisi, R. (2008). Preventive interventions addressing underage drinking: State of the evidence and steps toward public health impact. *Pediatrics*, *121 Suppl 4*, S311-336. https://doi.org/10.1542/peds.2007-2243E

Stigler, M. H., Neusel, E., & Perry, C. L. (2011). School-based programs to prevent and reduce alcohol use among youth. *Alcohol Research & Health: The Journal of the National Institute on Alcohol Abuse and Alcoholism*, *34*(2), 157–162.

Stone, A. L., Becker, L. G., Huber, A. M., & Catalano, R. F. (2012). Review of risk and protective factors of substance use and problem use in emerging adulthood. *Addictive Behaviors*, *37*(7), 747–775. https://doi.org/10.1016/j.addbeh.2012.02.014

Strøm, H. K., Adolfsen, F., Fossum, S., Kaiser, S., & Martinussen, M. (2014). Effectiveness of school-based preventive interventions on adolescent alcohol use: A meta-analysis of randomized controlled trials. *Substance Abuse Treatment, Prevention, and Policy*, *9*. https://doi.org/10.1186/1747-597X-9-48

Stuart, G. L., McGeary, J., Shorey, R. C., & Knopik, V. S. (2016). Genetics moderate alcohol and intimate partner violence treatment outcomes in a randomized controlled trial of hazardous drinking men in batterer intervention programs: A preliminary investigation. *Journal of Consulting and Clinical Psychology*, *84*(7), 592–598. https://doi.org/10.1037/a0040219

Substance Abuse and Mental Health Services Administration. (2017). *Key substance use and mental health indicators in the United States: Results from the 2016 National Survey on Drug Use and Health (HHS Publication No. SMA 17-5044, NSDUH Series H-52)*. Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. https://www.samhsa.gov/data/

Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., Zhang, Y., Ye, K., Jun, G., Hsi-Yang Fritz, M., Konkel, M. K., Malhotra, A., Stütz, A. M., Shi, X., Paolo Casale, F., Chen, J., Hormozdiari, F., Dayama, G., Chen, K., … Korbel, J. O. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature*, *526*(7571), 75–81. https://doi.org/10.1038/nature15394

Sussman, S., & Arnett, J. J. (2014). Emerging Adulthood: Developmental Period Facilitative of the Addictions. *Evaluation & the Health Professions*, *37*(2), 147–155. https://doi.org/10.1177/0163278714521812

Tambor, E. S., Bernhardt, B. A., Rodgers, J., Holtzman, N. A., & Geller, G. (2002). Mapping the human genome: An assessment of media coverage and public reaction. *Genetics in Medicine*, *4*(1), 31–36. https://doi.org/10.1097/00125817-200201000-00006

Tanner-Smith, E. E., & Risser, M. D. (2016). A meta-analysis of brief alcohol interventions for adolescents and young adults: Variability in effects across alcohol measures. *The American Journal of Drug and Alcohol Abuse*, *42*(2), 140–151. https://doi.org/10.3109/00952990.2015.1136638

Tarantino, N., Tully, E. C., Garcia, S. E., South, S., Iacono, W. G., & McGue, M. (2014). Genetic and Environmental Influences on Affiliation with Deviant Peers during Adolescence and Early Adulthood. *Developmental Psychology*, *50*(3), 663–673. https://doi.org/10.1037/a0034345

Tawa, E. A., Hall, S. D., & Lohoff, F. W. (2016). Overview of the Genetics of Alcohol Use Disorder. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, *51*(5), 507–514. https://doi.org/10.1093/alcalc/agw046

The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Tian, C., Gregersen, P. K., & Seldin, M. F. (2008). Accounting for ancestry: Population substructure and genome-wide association studies. *Human Molecular Genetics*, *17*(R2), R143–R150. https://doi.org/10.1093/hmg/ddn268

Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: The shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics*, *19*(2), 110–124. https://doi.org/10.1038/nrg.2017.101

Uhl, G. R., Drgon, T., Johnson, C., Ramoni, M. F., Behm, F. M., & Rose, J. E. (2010). Genome-Wide Association for Smoking Cessation Success in a Trial of Precessation Nicotine Replacement. *Molecular Medicine*, *16*(11–12), 513–526. https://doi.org/10.2119/molmed.2010.00052

van Ijzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2015). Genetic differential susceptibility on trial: Meta-analytic support from randomized controlled experiments. *Development and Psychopathology*, *27*(01), 151–162. https://doi.org/10.1017/S0954579414001369

van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., Belsky, J., Beach, S., Brody, G., Dodge, K. A., Greenberg, M., Posner, M., & Scott, S. (2011). Gene-by-environment experiments: A new approach to finding the missing heritability. *Nature Reviews. Genetics*, *12*(12), 881; author reply 881. https://doi.org/10.1038/nrg2764-c1

Verhulst, B., Neale, M. C., & Kendler, K. S. (2015). The heritability of alcohol use disorders: A meta-analysis of twin and adoption studies. *Psychological Medicine*, *45*(5), 1061–1072. https://doi.org/10.1017/S0033291714002165

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *American Journal of Human Genetics*, *101*(1), 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E., Aliev, F., Bacanu, S.-A., Batzler, A., Bertelsen, S., Biernacka, J. M., Bigdeli, T. B., Chen, L.-S., Clarke, T.-K., Chou, Y.-L., Degenhardt, F., Docherty, A. R., Edwards, A. C., Fontanillas, P., … Agrawal, A. (2018). Trans-ancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nature Neuroscience*, *21*(12), 1656–1669. https://doi.org/10.1038/s41593-018-0275-1

Walters, S., Vader, A. M., Harris, T. R., Field, C. A., & Jouriles, E. N. (2009). Dismantling motivational interviewing and feedback for college drinkers: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *77*(1), 64–73. https://doi.org/10.1037/a0014472

Webb, B. T., Edwards, A. C., Wolen, A. R., Salvatore, J. E., Aliev, F., Riley, B. P., Sun, C., Williamson, V. S., Kitchens, J. N., Pedersen, K., Adkins, A., Cooke, M. E., Savage, J. E., Neale, Z., Cho, S. B., Dick, D. M., & Kendler, K. S. (2017). Molecular Genetic Influences on Normative and Problematic Alcohol Use in a Population-Based Sample of College Students. *Frontiers in Genetics*, *8*. https://doi.org/10.3389/fgene.2017.00030

Wechsler, H., Lee, J. E., Kuo, M., Seibring, M., Nelson, T. F., & Lee, H. (2002). Trends in college binge drinking during a period of increased prevention efforts. Findings from 4 Harvard School of Public Health College Alcohol Study surveys: 1993-2001. *Journal of American College Health: J of ACH*, *50*(5), 203–217. https://doi.org/10.1080/07448480209595713

White, A., & Hingson, R. (2013). The burden of alcohol use: Excessive alcohol consumption and related consequences among college students. *Alcohol Research: Current Reviews*, *35*(2), 201–218.

World Health Organization. (1997). *Composite International Diagnostic Interview (CIDI, ver. 2.1)*. World Health Organization.

Yanovitzky, I. (2006). Sensation seeking and alcohol use by college students: Examining multiple pathways of effects. *Journal of Health Communication*, *11*(3), 269–280. https://doi.org/10.1080/10810730600613856

Young, A. I. (2019). Solving the missing heritability problem. *PLoS Genetics*, *15*(6). https://doi.org/10.1371/journal.pgen.1008222

Young, A. I., Benonisdottir, S., Przeworski, M., & Kong, A. (2019). Deconstructing the sources of genotype-phenotype associations in humans. *Science (New York, N.Y.)*, *365*(6460), 1396–1400. https://doi.org/10.1126/science.aax3710

Zhou, H., Sealock, J. M., Sanchez-Roige, S., Clarke, T.-K., Levey, D. F., Cheng, Z., Li, B., Polimanti, R., Kember, R. L., Smith, R. V., Thygesen, J. H., Morgan, M. Y., Atkinson, S. R., Thursz, M. R., Nyegaard, M., Mattheisen, M., Børglum, A. D., Johnson, E. C., Justice, A. C., … Gelernter, J. (2020). Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nature Neuroscience*, *23*(7), 809–818. https://doi.org/10.1038/s41593-020-0643-5