The Texas Twin Project

K. Paige Harden,¹ Elliot M. Tucker-Drob,¹ and Jennifer L. Tackett²
¹Department of Psychology, The University of Texas at Austin, Austin, TX 78712, USA
²Department of Psychology, The University of Houston, Houston, TX 77004, USA

Socioeconomic position, racial/ethnic minority status, and other characteristics of the macro-environment may be important moderators of genetic influence on a wide array of psychosocial outcomes. Designed to maximize representation of low socioeconomic status families and racial/ethnic minorities, the Texas Twin Project is an ongoing study of school-age twins (preschool through 12th grade) enrolled in public schools in the Austin, Texas and Houston, Texas metropolitan areas. School rosters are used to identify twin families from a target population with sizable populations of African American (18%), Hispanic/Latino (48%), and non-Hispanic White (27%) children and adolescents, over half of whom meet US guidelines for classification as economically disadvantaged. Initial efforts have focused on a large-scale, family-based survey study involving both parent and child reports of personality, psychopathology, physical health, academic interests, parent–child relationships, and aspects of the home environment. In addition, the Texas Twin Project is the basis for an in-laboratory study of adolescent decision-making, delinquency, and substance use. Future directions include geographic expansion of the sample to the entire state of Texas (with a population of over 25 million) and genotyping of participating twins.

Keywords: twins, socioeconomic status, gene–environment interaction, neighborhoods, schools

In the United States, public policies to increase student achievement and improve public health frequently target low-income and ethnic minority children. Socioeconomic status (SES) and ethnic minority status are strong predictors of low academic achievement and educational attainment; these disparities are evident in early childhood and progressively widen over the course of schooling (Heckman, 2006; Sirin, 2005; Tucker-Drob et al., 2011). In addition, low-SES is associated with poorer physical health and increased risks of mental disorders (Adler et al., 1994; American Psychological Association, 2006; Bourdon et al., 1994). Finally, some behavioral genetic research suggests that macro-environmental contexts, including SES and neighborhood characteristics, may moderate the influence of genetic factors on phenotypes such as intelligence (Harden et al., 2007; Rhentulla & Tucker-Drob, 2012; Tucker-Drob et al., 2011; Turkheimer et al., 2003), externalizing behavior (Dick et al., 2009; Legrand et al., 2008; Sadeh et al., 2010; Tuvblad et al., 2006), internalizing psychopathology (South & Krueger, 2011), and sexual behavior (Carlson et al., 2012).

Nevertheless, many existing twin registries are poorly situated to address research questions relating to how economic and social advantage interacts with genetic influences on behavioral phenotypes, because they focus on middle- to upper-class families or are racially and ethnically homogenous. The under-representation of low-income and minority children in twin research overall may stem from several sources, including the demographic characteristics of the larger population (e.g., there are relatively few African-Americans in Minnesota), the lower twinning rate among Hispanic American and young mothers (US Department of Health and Human Services, 2011), and the well-known obstacles to engaging low-SES populations in psychological research (Schnirer & Stack-Cutler, 2012). In addition, several of the largest national twin registries are located in northern European nations, which have comprehensive social welfare nets that guarantee access to medical care, adequate housing, parental leave, and other social programs that improve the standard of living for low-income children, and thus may yield results that do not generalize to low-income children in the United States.

The overarching goal of the Texas Twin Project is to build a new, US-based twin registry that allows us to...
address questions about how broad, macro-environmental contexts interact with genetic influences on child development. The state of Texas shares a large border with Mexico and has a substantial Hispanic population. In addition, as we describe below, the Texas Government Code includes the Public Information Act (formerly known as the Open Records Act), which ensures citizens’ access to government records, including records compiled by public schools. The Public Information Act, in combination with the demographic profile of the state, provides us with a unique opportunity to build a new twin registry that represents large numbers of low-income and ethnic minority children.

**Target Population**

The Texas Twin Project currently focuses on school-age children and adolescents in the greater metropolitan areas of Austin, Texas and Houston, Texas. The Austin metropolitan area has a population of nearly 1.7 million and is notable for its racial and ethnic diversity: Based on the data from the US Census, less than half the population (49.9%) is non-Hispanic White, 8.5% is Black, and 34.2% is Hispanic. Nearly 17% of Austin residents under age 18 years live below the federal poverty line. Houston is the fourth most populous city in the United States, with over 6 million people in the metropolitan area. According to the US Census, 37% of Houston’s population is Hispanic (predominantly Mexican-American), 31% non-Hispanic White, 25% Black, and 5% Asian. Over a quarter (26%) of Houston residents under the age of 18 years live below the federal poverty line.

The target population for the Texas Twin Project includes children enrolled in eight ‘independent’ school districts (ISDs) in Austin and in 24 ISDs in Houston. Together, these 32 school districts contain over 1 million students. Overall, nearly half of the students in the targeted districts are Hispanic/Latino (47.7%), 17.6% are African-American, and 26.9% are White. Students in targeted public schools are, as is often the case in the United States, more economically disadvantaged than the general population. Over half of the students in the target districts (54.2%) are classified as ‘economically disadvantaged’ by the state of Texas, which means the student is eligible for free or reduced-price lunch on the basis of family income at or below 185% of the federal poverty line (US Department of Agriculture, 2012), or that the student or his/her family is eligible for another form of public assistance (e.g., food stamps, public housing). School districts differ widely in their concentration of low-income students, ranging from 3% to 100% economically disadvantaged. The largest school district in each of the two cities (Austin ISD and Houston ISD) includes a substantial number of low-income students: 64% and 81% respectively. Finally, 20% of students in target districts are English-language learners.

To gain a conservative estimate of the number of twins in the target population, we used published rates of multiple births in young mothers (aged 20–24 years) from the US National Vital Statistics Report (US Department of Health and Human Services, 2011): 15.5, 23.0, and 13.2 per 1,000 births for non-Hispanic Whites, non-Hispanic Blacks, and Hispanics respectively. As illustrated in Figure 1, this yields an estimated 16,852 multiples. Thus, the target population of the Texas Twin Project has the potential to generate sufficiently large sample sizes for sophisticated multivariate behavioral genetic analyses.

**Twin Family Recruitment**

The Public Information Act, which can be found in Chapter 552 of the Texas Government Code, ensures that individual citizens have access to ‘any information collected, assembled, or maintained by or for a governmental body’, including public school districts. Once a request is made, the governmental body must provide the requested records within 10 days, or must respond to the request by providing a date and time when the information will be available. (Further information on the Public Information Act can be found on the Texas Attorney General website: https://www.texasattorneygeneral.gov/page?site:https://www.oag.state.tx.us/open/og FAQs.shtml). Public school ‘directory information’ is requested from each school district under the Public Information Act. Directory information typically contains the student’s name, grade, school, parent name(s), mailing address, and telephone number(s). Directory information may also contain the student’s date of birth, but school district policies differ regarding whether date of birth is protected from disclosure under the Family Educational Rights and Privacy Act (US Department of Education, 2012). At the beginning of each school year, parents may elect for their directory information to be private, which protects this information from being disclosed under the Public Information Act.

For school districts that provide date of birth, potential twins are identified by matching students with the same date of birth who reside at the same mailing address. Because Hispanic American families often give their children both maternal and paternal surnames, and these surnames are inconsistently recorded by school districts, we do not require that potential twins have the same last name. For school districts that do not provide date of birth, potential twins are identified by matching students who are enrolled in the same grade, and who also reside at the same mailing address. This approach likely misses a small number of twins who are in different grades and includes a small number of other types of sibling or sibling-like relationships (i.e., non-twin siblings born within 12 months of each other, step-siblings, or same-age cousins raised in the same household).

Our initial efforts to establish the Texas Twin Project have focused on ascertaining response rates and collecting survey data on individual twin families. Once potential twins are identified using the strategies described above, contact packets are mailed out in English and Spanish. Parents can indicate interest in participation by returning their
FIGURE 1
Flow chart illustrating the race/ethnic composition of the target population and the estimated number of twins to be targeted by the Texas Twin Project. Note: Rate of multiple births is based on the US Vital Statistics Report, which can be accessed at http://www.cdc.gov/nchs/data/nvsr/nvsr60/nvsr60_01.pdf.

contact information by mail or by registering online, in which case family-level surveys are sent back through the mail or online, whichever format the parents prefer. Families who do not respond to the initial mailing are then contacted by phone. If successfully contacted by phone, families can agree to participate in the survey data collection, decline to participate but agree to be enrolled in the Texas Twin Project database for future studies, or decline both participation in the survey and enrollment in the registry. Families receive financial compensation, with increased compensation offered as an incentive for families who are slow to respond to initial recruitment efforts. All questionnaires and related study materials have been translated into Spanish by a native Spanish-speaking doctoral student, and families indicate their preferred language for survey administration at the time of recruitment. Spanish-speaking research assistants on our staff communicate with Spanish-speaking families by both email and telephone.

Recruitment for the Texas Twin Project is ongoing. Thus far, we have recruited 630 sets of multiples – 599 pairs of twins (194 female–female, 189 male–male, and 216 mixed-sex), 30 sets of triplets (4 female, 7 male, 19 mixed-sex), and one set of quadruplets (mixed-sex) – from 619 families (11 families have more than one set of multiples). We have completed data collection for 335 families. Zygosity was determined using a 10-item questionnaire, completed by the parent, which assesses co-twin similarity in appearance (e.g., hair color, hair structure, eye color, difficulty telling the twins apart). Based on this questionnaire, 33% of twins have been classified as monozygotic and the remaining 67% as dizygotic.

Measures
Parent and Child Surveys
Parents complete a variety of questions, including demographic information about themselves and their spouse (e.g., education, occupation), as well as information about their household composition, their religion, personality, symptoms of depression, anxiety, and substance use, and aspects of their relationships with their spouse and children. Parents also complete questionnaire items regarding both twins, including the twins’ personality, emotional and behavioral problems, health functioning, and school and after-school activities. Adolescents (enrolled in grades 7–12) complete a self-report questionnaire, which includes items assessing their personality, experiences in the classroom, religious beliefs and practices, peer relationships, family relationships, body image, behavioral and emotional problems, and pubertal development. Table 1 summarizes key constructs and example measures included in the parent and child surveys.

School Records of Academic Achievement
In addition to completing surveys, all participating parents give active signed consent for the school district to release their children’s cumulative educational records. These records contain scores on the state-mandated standardized achievement tests, the Texas Assessment of Knowledge and Skills (TAKS; soon to be renamed as State of Texas Assessments of Academic Readiness, or STAAR). Reading/English Language Arts and Math tests are administered annually in Grades 3–11. Science tests are administered in Grades 5, 8, 10, and 11; and Social Studies tests are administered in Grades 8, 10, and 11. For eligible English language learners, achievement tests are administered in Spanish-language versions, and linguistically accommodated testing procedures are available. Development and scoring of the TAKS and STAAR tests is a collaborative effort by professional test developers and psychometricians at the Educational Testing Service, Pearson Corporation, and the Texas Educational Agency. Test content aligns closely with the Texas statewide curriculum, and is updated yearly and field-tested before

TWIN RESEARCH AND HUMAN GENETICS FEBRUARY 2013
School Characteristics
The effect of school composition on student achievement has been an issue of strong practical interest, with relevance for school choice, tracking, and desegregation policies, and there has been a correspondingly voluminous research literature. Higher peer academic achievement, higher school-level SES, and greater school racial/ethnic diversity are all associated with improved academic outcomes (Borman & Overman, 2004; Crosooe, 2009; Hanushek et al., 2003; van Ewijk & Sleegers, 2010). As part of its Academic Excellence Indicator System (AEIS), the Texas Education Agency annually publishes a rich dataset regarding aspects of the school environment, including information on individual grade levels. These annual records are published soon after the end of every school year, and these are available at http://ritter.tea.state.tx.us/perfreport/aeis/ for every school year since 1993.

Neighborhood Characteristics
Leventhal & Brooks-Gunn (2000), in their comprehensive review of the literature, concluded that neighborhood poverty, racial/ethnic diversity, and residential instability are ‘the structural dimensions of most theoretical importance’ for child development. The Texas Twin Project uses census tract data to construct measures of these neighborhood characteristics. Census tracts are small geographical subdivisions that usually have between 2,500 and 8,000 residents and are delineated by social or physical boundaries. Census tract designations are designed to produce areas that are relatively stable and homogeneous regarding population characteristics, economic advantage, and physical conditions (US Census Bureau, 2010). The mailing address of each family can be geocoded and matched to a census tract, which allows us to link twins to time-varying data from the American Community Survey (ACS) and to data from the 2010 US Census on neighborhood poverty, population density, residential instability, and racial/ethnic diversity, in addition to other important contextual features. ACS data are actually released by the US Census Bureau (2010) every year and represent a ‘rolling average’ of population characteristics from the past 60 months, and can thus be used to construct measures of neighborhood change.

Plans for Future Research
In future research with Texas Twin Project families, we plan to take an integrative approach that spans multiple levels of analysis, including genes, hormones, personality, cognition, relationships, neighborhood factors, and school characteristics. In particular, we aim to better understand how these levels of analysis interact in predicting individual-level outcomes, and the extents to which such interactions are moderated by ethnic and economic background. To meet these goals, we plan to extend the Texas Twin Project in three major ways. First, we hope to expand the Texas Twin Project to the entire state of Texas. Texas is the second most populous state in the United States, with a population of more than 25 million people. In particular, geographic expansion will allow us to examine differences between urban versus rural populations. Second, families from the Texas Twin Project will be recruited for smaller, in-laboratory studies that allow for ‘deep phenotyping’ on constructs of particular interest. For example, we are currently conducting an in-laboratory study of adolescent (grades 9–12) twins. This project focuses specifically on alcohol use and antisocial behavior, and it involves cognitive testing, a battery of computerized decision-making tasks, and hormonal samples. Finally, we plan to collect DNA samples, both to validate our zygosity diagnoses and to allow for molecular genetic analyses.

Conclusions
The Texas Twin Project, while still in the early stages of development, has the potential to be a distinctive asset to implementation. Tests are scored using Item Response Theory methods and placed on vertical scales, which are appropriate for assessing change over time. Internal consistencies of the TAKS tests are very high, ranging from 0.87 to 0.90.
behavioral genetic research. In particular, the Texas Open Records Act, which ensures access to school rosters and provides a means to identify twin families, has provided us with a unique opportunity to build a twin registry that maximizes representation of economically disadvantaged and ethnic minority children. We are confident that this new registry will prove to be a valuable scientific resource in the years ahead.

Acknowledgments

K. Paige Harden and Elliot Tucker-Drob are members of the Population Research Center at the University of Texas at Austin, which is supported by National Institutes of Health (NIH) center grant R24-HD042849, and provided a seed grant for the establishment of the Texas Twin Project. Ongoing work on the Texas Twin Project is supported by NIH grant R21-AA020588. This manuscript was prepared while Elliot Tucker-Drob was supported by NIH grant R21-HD069772. We thank Daniel Briley, Natalie Kretsch, Marie Carlson, Amanda Cheung, Cassandra Moreno, and Davin Duval for their assistance on the registry. Texas Twin Project data are collected and managed using REDCap electronic data capture tools hosted at the Population Research Center at the University of Texas at Austin (Harris et al., 2009). Special acknowledgment is due to Dr. Sean Banks for his assistance with REDCap.

References


