Heritability, family, school and academic achievement in adolescence

Artur Pokropek [a,*], Joanna Sikora [b,1]

[a] Institute of Philosophy and Sociology, Polish Academy of Sciences, Nowy Świat 72, 00-330 Warsaw, Poland
[b] School of Sociology, Research School of Social Sciences, College of Arts and Social Sciences, Australian National University, Canberra, ACT 2601, Australia

ARTICLE INFO

ABSTRACT

We demonstrate how genetically informed designs can be applied to administrative exam data to study academic achievement. ACE mixture latent class models have been used with Year 6 and 9 exam data for seven cohorts of Polish students which include 24,285 pairs of twins. Depending on a learning domain and classroom environment history, from 58% to 88% of variance in exam results is attributable to heritability, up to 34% to shared environment and from 8% to 15% depends on unique events in students' lives. Moreover, between 54% and 66% of variance in students' learning gains made between Years 6 and 9 is explained by heritability. The unique environment accounts for between 34% and 46% of that variance. However, we find no classroom effects on student progress made between Years 6 and 9. We situate this finding against the view that classroom peer groups and teachers matter for adolescent learning.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

While educational research on twins has a long history, it remains a fairly narrow and specialized niche based mostly on twin surveys or twin registers data. The ongoing fascination with twin data methodologies arises from their potential to tease out the role of heritability in the effects of family background on educational outcomes of youth. Heritability is usually off the radar in mainstream social science research and yet it is implausible to assume that it has no role in parental influence on offspring. This is why the last two decades have seen a reinvigorated interest in educational studies of twins which are often described as 'genetically informed designs' (Freese, 2008). In this paper we show how methodologies developed for twin data can be applied to other data, retaining full benefits of genetically sensitive research. This may have important implications for educational research in countries where twin data are not readily available. As such, our methodology is likely to facilitate an expansion of genetically informed educational research beyond the studies from the USA, the UK and Australia, which currently dominate the literature in this area.

This study showcases a unique approach because our empirical evidence comes from the entire cohorts of 13- and 16-year-old Poles who sat compulsory placement exams in mathematics and humanities between 2002 and 2011. We show how such longitudinal data can be used to generate reliable quasi-genetic designs. Compulsory exams are already a norm in

* Corresponding author. Fax: +48 228267823.
    E-mail addresses: artur.pokropek@gmail.com (A. Pokropek), joanna.sikora@anu.edu.au (J. Sikora).
  [1] Fax: +61 261252222.

many countries and are likely to be introduced in others, in line with the worldwide standardization and benchmarking advocated by intergovernmental organizations such as the UNESCO or the OECD (Kamens and McNeely, 2010). We make a contribution to genetically sensitive educational research not only by applying twin study methodologies to high-stakes exam data but also by considering learning gains made by youth in early adolescence.

We begin with a description of the history and logic of genetically informed designs. Then, we review the literature on twins and educational achievement. Next, we describe the Polish education system, our research questions, data and methods. Our results follow and we conclude with a discussion of their compatibility with prior literature.

1.1. Genetically informed designs involving twin data

Worldwide, one to four percent of live births are twins. The proportions of twins vary by country and region. For instance in Poland in 2010, 2.5% of newborns were twins (CEB, 2002–2011) while in the United States the comparable figure stood at 4% (Plomin et al., 2013). Although many twins are born smaller than single babies, various studies show that twins do not differ consistently from other children in physical, psychological or social characteristics (Plomin et al., 2013). If differences exist, they peter out by the time children turn five (Evans and Martin, 2008).

Twins are either identical (monozygotic or MZ) or fraternal (dizygotic or DZ). MZ twins come from one egg and one sperm. DZ twins come from two eggs and two sperms. Identical twins are very similar and are always of the same sex. Genetically, the status of fraternal twins resembles that of other siblings; they grow up in the same environment and share on average about 50% of their segregating genes (Plomin et al., 2013). As DZ twinning is heritable and MZ twinning is random, the proportion of fraternal twins varies by population while the proportion of identical twins is relatively constant in different populations across the world (Bortolus et al., 1999).

The information summarized above is used in behavioral genetics to build models which partition variance in a range of outcomes, including education, into components which depict environmental and genetic influences. The next section describes the typical logic of such models and reasons for their popularity.

1.2. Classical twin study designs

Behavioral geneticists routinely employ three types of comparisons (Plomin et al., 2013). First, they compare identical twins that grow up in the same family with fraternal twins who share as much environment but only 50% of their segregating genes. Second, dizygotic twins are compared with adopted children who share environment but no genotype. Third, the most informative type of comparison involves identical twins who were separated at birth and thus differ entirely in the effects of environment but not genotype. Comparisons of correlations between particular traits of such twins enable teasing out all environmental influences, including the part attributable to shared environment.

A typical behavioral genetics study assesses the variance in behavior, or phenotype, and then partitions this variance into three components: heritability (h^2), shared (common) environment (c^2) and unique environment (e^2). To obtain the estimates of these components behavioral genetics uses information about within-pair correlations for MZ and DZ twins (for details see: Falconer, 1981; Jinks and Fulker, 1970). In the simplest decomposition scenario it is assumed that similarities between MZ twins arise due to fully shared genotypes and environments, while dissimilarities are caused by unique environment. The correlation for DZ twins is half of the correlation for MZ twins, as the former share on average about 50% of segregating genes. Therefore, in a simple model, heritability is equivalent to twice the difference between the MZ and the DZ correlations for the trait of interest. The contribution of shared environment is then found by subtracting the value of heritability from the correlation for identical twins. Finally, the remaining portion of the variance is attributed to unique environment (for details see Plomin et al., 2013).

In practice these computations are conducted with advanced techniques which may take into account the interactions between all variance components as well as the measurement error that affects them. Sophisticated models can also utilize information about multiple family members or simultaneously consider a number of traits (Posthuma et al., 2003: 361). Regardless of estimation methods, all such models are conceptually based on within-pair trait correlation comparisons between the MZ and the DZ twins.

In the interpretation of these models heritability refers to "the contribution of genetic differences to observed differences among individuals in a particular population at a particular time" (Plomin et al., 2013: 93). Shared environment involves all non-genetic influences that affect children from the same family in the same way (Asbury and Plomin, 2014). Nielsen (2006: 197) describes it as "background characteristics that stratification researchers presumably have in mind when they conceptualize mechanisms of social reproduction". They may include cultural possessions and tastes of parents, parental education, characteristics of within-family relationships, family wealth or income and much more. In short, all influences experienced in the same way by the children make up shared environment.

In contrast, unique environment involves all experiences that affect each twin in a unique manner, as well as measurement error (Nielsen, 2006). They may involve illness, injuries, peer relationships or anything that influences one and not the other sibling. This includes perceptions and emotional reactions specific to each child.