



**Leading education
and social research**
Institute of Education
University of London

Department of Quantitative Social Science

The socio-economic gradient in children's reading skills and the role of genetics

**John Jerrim
Anna Vignoles
Raghu Lingam
Angela Friend**

**DoQSS Working Paper No. 13-10
June 2013**

Disclaimer

Any opinions expressed here are those of the author(s) and not those of the Institute of Education. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

DoQSS Workings Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

The socio-economic gradient in children's reading skills and the role of genetics

John Jerrim¹, Anna Vignoles², Raghu Lingam³ and Angela Friend⁴

Abstract

By the time children leave primary school there is a large socio-economic gap in their reading proficiency. There are a number of potential explanations for this socio-economic gap and in this paper we investigate the role of three particular genes and gene-environment interactions in determining children's reading skills, using the Avon Longitudinal Study of Parents and Children (ALSPAC) dataset. We find that whilst these genes are indeed correlated with reading outcomes, effect sizes are small and sensitive to the choice of test used and the sample selected. Our results suggest that the leading candidate genes can jointly explain just 2% of the socio-economic gap in children's reading test scores. We conclude that the influence of these three genes on children's reading ability is limited, and their role in producing socio-economic gaps in reading ability is even more limited still.

JEL classification: I20, I24

Keywords: Reading, Genes, Gene-environment interaction, ALSPAC

¹ Corresponding author: Department of Quantitative Social Science, Institute of Education, University of London, 20 Bedford Way London, WC1H 0AL (J.Jerrim@ioe.ac.uk)

² Faculty of Education, University of Cambridge (av404@cam.ac.uk)

³ Department of Population Health, London School of Hygiene and Tropical Medicine, University of London (raghu.lingam@lshtm.ac.uk)

⁴ Institute of Cognitive Science, University of Colorado Boulder (angela.friend@colorado.edu)

Acknowledgements:

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team. This project has been funded by the ESRC ALSPAC large grant programme. We would also like to thank Silvia Paracchini for her advice on genetic data.

Executive summary: the socio-economic gradient in children's reading skills and the role of genetics

(Full paper available on request)

Existing research suggests that family background has a pivotal role in child development⁵ with certain mediating factors, such as parental education and aspirations, being particularly important⁶. Yet there is much about the role of family background that we still do not understand. Although the large socio-economic gaps in children's cognitive skills at an early age are well documented⁷, little work has considered the extent to which this is due to *genetic* influences, or their interaction with the environment. By ignoring the possible role of genetics, potentially important variables may be missing from models of social stratification.

Of course one reason that the role of genetics has been somewhat ignored in the social stratification literature is because of the controversy surrounding the notion of heritability. For instance, in "The Bell Curve" by Herrnstein and Murray (1994) it was claimed that intergenerational worklessness is (to a large extent) genetically determined. Writers such as Saunders (2010) have claimed that intergenerational transmission of social advantage may also be driven by inheritance of intelligence. Clearly such views are unpopular but they have been a recurring theme within social science research for over 50 years⁸. It is only now, with the growth in genetic data being collected, that it is possible to consider properly the role of genetic factors in creating differences between socio-economic groups. Indeed there is a growing expectation that genes and particularly their interaction with the environment, will shed new light on this important issue.

The social science literature on heritability to date has had to rely heavily on research comparing outcomes for identical (monozygotic – MZ) and non-identical (dizygotic – DZ) twins. The idea is that since the former share all their genes, whilst the latter only share 50%, one can use this to determine the extent to which an outcome is heritable. This literature has found reading heritability to be high. Indeed some researchers have found 'genetics' to be the dominant factor predicting reading achievement using this approach⁹. However, there are limitations to twin studies which we discuss in our paper, not least the problematic assumption that twins really do share the same environment and by implication that parents

⁵ Coleman 1966, Haveman and Wolfe 1995, Todd and Wolpin 2003.

⁶ Behrman and Rosenzweig 2002, Black et al 2005, Cunha and Heckman 2008, Goodman et al 2009.

⁷ Feinstein 2003, Jerrim and Vignoles 2013, Goodman et al 2009.

⁸ See, for instance, Jensen 1968.

⁹ Harlarr et al (2005) for example report that genetic factors account for 75% of the variance in reading skills.

do not treat twins differently. Further, the simplistic nature versus nurture argument underlying many twin studies cannot be supported by recent advances in genetic research, which have suggested that how a particular gene impacts on a particular outcome is actually influenced by the environment to which it is exposed. In other words, the two interact¹⁰.

In our work we make a contribution to this emerging field by investigating the direct association between family background and reading achievement, using genetic data for a large sample of children. We focus on three suspected genetic risk factors (DCDC2, KIAA0319 and CMIP) and children's reading test scores. These are the most promising candidate genes (DCDC2, KIAA0319 and CMIP) available within the Avon Longitudinal Study of Parents and Children (ALSPAC) dataset, which is a data set consisting of women in the Avon district of England who were pregnant between April 1991 and December 1992 and who agreed to participate in the study¹¹. 14,541 mothers were recruited into the study. They and their children have been re-interviewed at regular intervals and a large number of participating children were invited to special clinic sessions, where a series of diagnostic tests were performed (including reading skills assessments). Genotypic information has also been collected. This paper is one of the first to use such bio-molecular data encoded from the human genome to explain differences across socio-economic groups in one particular cognitive skill, namely children's reading ability.

Using these data, we find that whilst these genes are indeed correlated with reading outcomes, effect sizes are small and sensitive to the choice of test used and the sample selected. We find no evidence of gene-environment interactions. Rather our results suggest that the leading candidate genes can jointly explain just 2% of the socio-economic gap in children's reading test scores. We conclude that the influence of these three genes on children's reading ability is limited, and their role in producing socio-economic gaps even more limited still. However, we have yet only investigated the impact of a small set of quite specific genes. As the literature on missing heritability suggests, many more genes may be implicated in the reading process (possibly hundreds, each with small, independent effects). Hence this work is in its infancy and whilst the availability of genotypic data is unlikely to have a significant impact on our understanding of the major factors influencing socio-

¹⁰ Epigenomics refers to the study of genomic modifications that alter gene expression, loosely meaning the study of gene-environment interactions. See Carey (2012) for the origins of epigenomics.

¹¹ Please note that the study website contains details of all the data that is available through fully searchable data dictionary (<http://www.bristol.ac.uk/alspac/researchers/data-access/data-dictionary/>)

economic differences in children's outcomes in the short term, much more research is needed on this issue before we can dismiss the role of genetics in influencing children's outcomes.

References

- Behrman, J. and Rosenzweig, M. (2002) 'Does increasing women's schooling raise the schooling of the next generation?', *American Economic Review* 92(1): 323 – 334.
- Black, S.; Devereux, P. and Salvanes, K. (2005) 'From the cradle of the labour market? The effect of birth weight on adult outcomes?', *The Quarterly Journal of Economics* 122(1): 409 – 439.
- Carey, N. (2012) *The Epigenetics Revolution: How Modern Biology is Rewriting Our Understanding of Genetics, Disease, and Inheritance*. New York: Columbia University Press
- Cunha, F.; Heckman, J. and Lochner, L. (2006) 'Interpreting the evidence on life cycle skill formation.' In E. Hanushek and F. Welch (Eds) *Handbook of the Economics of Education* pp. 697-812. Amsterdam: Holland North.
- Feinstein, L. (2003) 'Inequality in the early cognitive development of British children in the 1970 cohort', *Economica* 70(1): 73-97.
- Goodman, A.; Sibietta, L. and Washbook, E. (2009) *Inequalities in Educational Outcomes Among Children Aged 3 to 16*, Final report for the National Equality Panel, Institute for Fiscal Studies, London. Accessed 30th November 2012 from <http://sta.geo.useconnect.co.uk/pdf/Inequalities%20in%20education%20outcomes%20among%20children.pdf>
- Harlaar, N.; Spinath, F.; Dale, P. and Plomin, R. (2005) 'Genetic influences on early word recognition abilities and disabilities: A study of 7-year-old twins', *Journal of Child Psychology and Psychiatry*, 46(4): 373 – 384.
- Haveman, R. and Wolfe, B. (1995) 'The determinants of children's attainments: A review of methods and findings', *Journal of Economic Literature* 33(4): 1829-78.
- Herrnstein, R. and Murray, C. (1994) *The Bell Curve: Intelligence and Class Structure in American Life*, New York: The Free Press.
- Hobcroft, J. (2012) 'Epigenetics and the social sciences: progress, prospects and pitfalls', paper presented at the ESRC International Grand Challenge Symposium, 25-26 June 2012. Accessed 30th November 2012 from http://www.esrc.ac.uk/_images/Epigenetics-symposium-background-paper_tcm8-21507.pdf.
- Jensen, A. (1968) 'Social class, race, and genetics: implications for education', *American Educational Research Journal*, 5(1): 1 – 42.
- Jerrim, J. and Vignoles, A. (2013) 'Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes', *Journal of the Royal Statistical Society Series A*, doi: 10.1111/j.1467-985X.2012.01072.x
- Saunders, P. (2010) *Social Mobility Myths*, London: Civitas.

Todd, P. and Wolpin, K. (2003) 'On the specification and estimation of the production function for cognitive achievement', *The Economic Journal* 113(485): F3-F33.