

COMMENT AND DEBATE

Social class differences in early cognitive development: a response from Leon Feinstein

Leon Feinstein Early Intervention Foundation and London School of Economics
leon.feinstein@eif.org.uk

(Received September 2015 Revised October 2015)

<http://dx.doi.org/10.14301/llcs.v6i4.385>

Introduction

The July edition of this journal (Feinstein et al. 2015) included a special comment and debate section comprising five papers, including my own, on how the trajectories of cognitive skill through development of children vary with respect to family of origin and early scores at different times and places, how these trajectories might be modelled and what can be inferred from this about the impact of social structure. This debate impinges on what might be said about the opportunities for policy to address structural inequalities in child development. The focus of the debate was a graph in Feinstein (2003) that is recognised as influential in discussions about early intervention but has been criticised as being flawed in a number of ways (Tu and Law, 2010; Jerrim and Vignoles, 2013). I summarise the general background in Feinstein et al. (2015) so do not repeat that here. The annex to this paper includes figure 1 and figure 2 from Feinstein (2003) and the reader is directed there or to Feinstein et al. (2015) for a description of methods. I am particularly grateful to the authors of the comment and debate papers who have contributed further insight and reflection on the underlying questions and to the editors for letting me respond here to the thoughts they set out in their commentaries.

Since the 1970s structural inequalities of wealth and income have increased very substantially (Atkinson, 2015) as have public and private investment in cognitive development (Willetts, 2011). It would be interesting to know how these and other trends have changed the degree to which social circumstance influences the development and educational achievement of children so it is

important to have reliable, comparable developmental and structural measures. It would be useful to test whether cross-national differences in the structure of inequality in a nation are reflected in differences in structural inequalities in children's development, but this requires sound, agreed and comparable ways to measure and model trajectories. Until we have clarity on what differences in trajectories are due to measurement or to the way trajectories are modelled we cannot test what differences are socially structured by time and place, nor what might be thought local and what more universal. So I am grateful to the co-authors of Feinstein et al. (2015) for proposing methods that might allow more comparative study. They all offer exciting potential and I hope they are used widely. Between them they offer the possibility of triangulating one approach against another with the aim of achieving convergent conclusions. My remaining comments are intended mainly to highlight some of the difficulties and uncertainties in thinking about the implications of these methods to the question of the possibility of intergenerational change. I also indicate why I think that despite decades of policy failure there is room for optimism.

Numbers and words

I would like to bring out two themes from these papers. The first is that this is not a debate solely about statistical methods. If the graph published in my 2003 paper and reproduced in figure 2 had only ever been published in an economic journal and never been used in policy debates it would have inspired much less discussion. It is the use of the statistics as much as the statistics themselves that is at issue. This theme is emphasised by Lupton (2015)

who stresses the importance of inter-disciplinarity, highlighting that much of the important evidence on why there is a social class achievement gap comes from qualitative and sociological research. She points to the problems resulting from the dominance of variable-based statistics and econometrics that appear to have had a much more prominent profile and influence in the analytical work of government, a theme also picked up in Feinstein and Peck (2008). Lupton asks why qualitative and sociological approaches are so often ignored in policy debate. This is an important contribution because she asks us to consider not just which statistical method is best but what might be the limitations of a statistical understanding, a theme not picked up in the other papers. I would only add that in my experience there are few social policy questions on which practitioners and policy makers would consider information solely from narrow quantitative sources sufficient. However, it does appear to be true that quantitative analysts are paid a premium compared to qualitative researchers in government as in the private sector.

Lupton also highlights the importance of language, of how debates are framed. This is not just because of the differences between and within science and social science in how study is undertaken and what terms are used, but also because of the large gap between this and the public or policy understanding of these same terms.

Measuring, modelling and inference

The second theme and a main focus of the other three commentaries by Goldstein and French, Jerrim and Vignoles and Washbrook and Lee (Feinstein et al. 2015) is ways to measure and interpret the degree to which pathways or trajectories are shaped by early scores and family of origin. Three different modelling frameworks are discussed, all of which offer advances in estimation of a more specific set of questions than addressed by figure 2 of Feinstein (2003).

Some points of consensus

Jerrim and Vignoles (2015) provide a useful summary of five points of consensus. Three relate to what is known statistically about the emergence of gaps between social groups in cognitive development through childhood, although the scope of this assessment in terms of time and place is not defined. First, there are large average gaps by socioeconomic group in cognitive skill that can be

observed from a very early age. Second, these do not decline through schooling in absolute or, third, relative terms.

It is clear from their paper that there are some important starting points but also many questions on which we have at best cursory and preliminary understanding and substantial disagreement. We cannot say categorically whether clear gaps in the attainment of children living in the UK currently broaden through childhood. I also note that there is not a clear specification of what is meant by a large gap. For the wider public debate it is also important to emphasise that these assertions are only true about averages; in public debate an average difference between groups is often interpreted as a universal difference between all members. Important also to recognise that these are historically contingent statements, true for specific times and places and not general truths.

The final two summary points made by Jerrim and Vignoles relate to understanding of the graph, thus fourth: "The striking decline between 22 and 42 months should *not* be used by academics or policymakers to stress the importance of the early years, that we are failing 'bright' young people from disadvantaged backgrounds, or to highlight the lack of social mobility in the UK." The point was made best by an official in the Scottish Government at a seminar in Edinburgh in 2012, who remarked that if she had known it was so complicated she would never have used the graph. I agree the shift between 22 and 42 months says nothing about the degree of social mobility and does not support the claim that anyone is being failed. As I pointed out in Feinstein (2015) this was never my claim on the basis of the 22 to 42 month shifts nor in my view is it indicative of how the graph was used in general, not least because the shifts between 22 and 42 months are so large for both high and low socioeconomic status (SES) groups. However, I do recognise there has been substantial confusion about this.

The final point of consensus noted by Jerrim and Vignoles is that there remains no robust and consistent evidence that initially high achieving young people from poor backgrounds are overtaken by low achieving children from affluent backgrounds in terms of their cognitive skills. True, although as Washbrook and Lee (2015) point out, whether this crossover happens or not depends on arbitrary assumptions about the cut-points by

which children are allocated to groups at the first age of measurement. In many ways figure 2 is best thought of as the corollary of figure 1, which concerns the change through development in the average gap by social class. If the SES gap broadens through development, more and more children from low SES families who score well early on will fall back relative to high SES children.

Washbrook and Lee (2015) offer a useful general approach that takes us away from the discrete approach of the figure 2 graphic and the debate about a crossover. The real question of interest for them is “whether low SES children systematically underperform relative to higher SES children with identical initial capacities,” and, as they go on to show, how this differs by age during development.

Washbrook and Lee treat the difficulty of measuring initial capacities as a measurement error problem and therefore responsive to econometric and statistical methods for dealing with measurement error, in particular the technique of instrumental variables (IV) by which an auxiliary measure unrelated to later scores is used to predict the initial cognitive performance score. Under assumptions set out by Washbrook and Lee the use of IV can address the measurement error aspect of the problem of accurately measuring initial abilities. A particular technical difficulty is the challenge of establishing that the instrumental variable does not contain predictive power for later scores, independently of the other factors such as SES included in the statistical model. They offer alternative approaches based on correcting for the reliability of test scores.

Using a prior cognitive score taken just after kindergarten entry as an instrumental variable for a large US sample of children who entered kindergarten in 1998 they find that a good part of the widening in scores between high and low SES groups with equivalent initial capabilities happens between the ages of seven and 14. This analysis does not claim to solve the problem of explaining the divergence by family background but offers a useful general approach to estimating and describing the degree and timing of such divergence given early abilities, having accounted for measurement error. Others will challenge them on their choice of instrumental variable but their approach is promising. I discuss below the question of defining what is meant by initial ability in the

context of a debate about true ability and the potential for social change.

Goldstein and French (2015) make two explicit challenges, asserting “flaws” in both my original analysis and my recent response. They also treat the issue as one of measurement error and propose a general framework based on Bayesian modelling not yet published. In my view these flaws indicate the difficulty of framing a clear, common language for understanding the issues involved in modelling the data presented in figure 2, rather than flaws of analysis, but it is useful to set out their comments and, given their use of language, I am compelled to respond.

The first of these alleged flaws is my statement that I was offering in figure 2 “a descriptive analysis,” - as I put it “the aim was to present the actual data rather than... corrected trajectories based on modelling assumptions.” Goldstein and French argue that there is no difference between an uncorrected and a corrected presentation of data, “each is meant to convey an inference about the underlying social process.” They are right of course that the actual data presented in figure 2 are not simple facts about a social process and inferences about what influenced it were implicit. The dependent variable itself is the rank in the first principal component of a set of age appropriate measures of cognitive development taken at four ages in development based on different underlying measures. I only included in the chart the average scores of children characterised as low and high SES, excluding the larger middle SES group and including only those scoring in the first or fourth quartile in the 22 months tests. My original paper (Feinstein, 2003) presented tables of the full set of transition matrices that showed how the children who scored in each quartile at 22 months scored at subsequent ages. This selection of specific cells for the graph has been called “extreme” (see Feinstein, 2015). I make no claim of pure objectivity in a transcendent world of pure fact.

However, I disagree if Goldstein and French wish to assert that a strong inference about an underlying causal process is a necessary element in all social science. Shame if a social scientist cannot explore without prejudice in an attempt to measure and describe the world, only able to test dimly understood hypotheses. In my introduction I made reference to a number of important hypotheses for policy making that do need testing in an attempt to

get closer to understanding the underlying processes. However, scientific progress depends, as Lupton indicates, on a multiplicity of views and approaches and there is scope for descriptive and explorative social science, even provocative social science. Nor do I agree that the act of correction for error based on modelling assumptions, whether Bayesian, based on simulations constructed from general linear modelling or instrumental variables techniques, is a trivial addition to the challenge of measurement. The difficulty of specifying even what is meant by true score indicates the difficulty of framing a clear basis for appropriate inference.

This brings us to the second alleged flaw of my argument, which is in my understanding of regression to the mean and true scores. The former, they assert, “simply occurs when the correlation between two measurements over time is less than one, as is the case with heights of fathers and sons.” The reference to an intergenerational transfer of associations is a mere detail in the modelling of correlation between two measures. They continue, “the notion of measurement error is entirely separate.” Washbrook and Lee make a similar assertion, although explicitly recognising that measurement error is one component, amongst others, of the problem of regression to the mean as described by Jerrim and Vignoles. Indeed, much of the Jerrim and Vignoles reanalysis focused on the difference between high and low SES groups in the degree of change between 22 and 42 months. They emphasise that in the classification of children as high ability at 22 months, more low SES children are misclassified than are high SES children resulting in a higher degree of what they call regression to the mean for the low SES children in the 22 and 42 month scores. They emphasise the role of measurement error (and luck) in this misclassification.

Goldstein and French also point to a lack of clarity in my discussion of the statistical notion of ‘true score’ which in their specification “proceeds from the common observation that the actual score that a child obtains on a test will depend on the actual items chosen, plus other factors that might be considered ‘transient’ (their quotation marks) such as time of day, test environment etc.”

This may be a fair if imprecise and general specification of the statistical notion of measurement error but ignores the distinction I make between this imprecise notion of a true level

of capability that a social scientist might seek to measure and the potential capability of the child, a distinction elided in the simple use of the phrase ‘true ability’, and in application of these data to the question of the possibility of social change. This double measurement issue is not addressed by these statistical models, which is why I remain cautious about the use of corrected data presented as facts about children of different ability groupings. However, Goldstein and French are right to admonish me with their regret that in my handling of the trends by sub-group in the 1958 cohort study (Feinstein, 2004) I did not reference Goldstein’s important work on this topic. I do not argue that corrections for measurement error are inappropriate or misleading, I merely note that they are not trivial nor are their own limitations always as clear as they might be.

Goldstein and French argue that as this debate has been “a difficult one for policy makers... a more cautious long-term attitude should be taken to research findings.” They suggest that policymakers should “promote a wide debate about any findings that appear important, where technical and interpretational issues are debated in terms which are widely accessible.” I look forward to this and hope that future contributions to the debate from statisticians offer more aid to accessibility than hitherto.

There is more focus in these papers on how to measure the extent to which cognitive development is moderated by early scores and presumed underlying abilities than on the problem of defining social groups or measuring cognitive skill – weaknesses also of Feinstein (2003). There is no discussion of problems of aggregation in interpreting data on profiles of averages without adequate specification of a coherent multilevel framework, yet many in the public debate struggle to appreciate the difference between aggregate findings reflective of general social averages and the likely experience and outcomes of individual children. The ecological fallacy (see e.g. Diez-Roux, 1998) in statements such as that “it is all over by age 5” based on charts of averages such as figure 2 is widespread and misleading.

As stated above, a focus of both figure 2 and the original paper from which it came is on the instability of scores. I presented transition matrices, which showed a great deal of movement by children in their test scores over time. It was never

my intention to imply that there exist fixed groupings of ability. As ever the fact that something can be quantified and measured is not proof of its existence and we should beware the over-determination of the meaning of statistics and estimated coefficients. This is a problem for estimates of all kinds, useful if handled carefully but for which the apparent precision of numbers suggests a certainty that is easily over interpreted, particularly in a policy debate in which balance of interpretation is hard to achieve.

My experience in using the graph with policy makers was that using the uncorrected data enabled me to point both to the issue of score instability and to the subsequent patterns at the later ages which under reasonable assumptions are not explained by measurement error (Feinstein, 2015). Presenting corrected data enables a policy maker to gloss over the difficulty of classifying children to groups of ability as though statistical science has resolved the underlying philosophical and biological issues, introducing further miscomprehensions and over-simplifications to the debate.

Genes

This brings me to an important aspect of explanation and causation raised by Jerrim and Vignoles who note the challenges highlighted in relation to understanding the role of genetics. They point to the important and fascinating work of Robert Plomin and the tradition of structural genetic research based on twins studies which indicates that for many observable features of human development there are variable but often substantial proportions of the difference in outcomes that are explained in a structural, statistical sense by genes. This has included analysis of outcomes like intelligence, social class, aspects of personality as well as more obviously physical phenotypes.

As Lupton describes, this set of findings has been linked by Saunders (2011) with the critique of figure 2 based on modelling of 'true ability' to imply that our current system of allocation of wealth and opportunity is both efficient and reflective of an underlying natural distribution of capability and hence fair as well. Clearly, as Jerrim and Vignoles, indicate this is a controversial topic. As Lupton (2015) explains this is said to have undermined the 2010 Coalition Government's commitment to increasing social mobility so some discussion is

necessary. I strongly agree with Jerrim and Vignoles that social science should not shy away from addressing the topic of genetics and biological science, not least to recognise how informative it is, how quickly it is changing and how broad is the opportunity it indicates for policy and practice. However, this does not negate the need for social scientific understanding of social questions.

It will be well known to readers of this journal, but is not yet always known in the world of politics, that there is a lot of dynamic complexity in the interactions between genes and environments. Epigenetics is finding that the interactions between environment and genome are so dynamic that it is falsely simplistic to think this a unilinear, biologically driven phenomenon (Carey, 2012). It is as wrong to overstate estimates of heritability as meaning destiny is fixed at birth as it is to ignore the evidence that genetics plays a role. The question of how much is to play for is not well established but we know, not least from genetically sensitive research designs (Weaver et al., 2004) that there is plenty of opportunity for practice and policy to play a substantial role in influencing intergenerational continuities of achievements and behaviour. There is no reason to look on the findings of structural genetics as implying any currently binding limit to the possibility of social change.

It is clear that epigenetic research is changing the nature of scientific understanding of the relationship between genes and environment. In a wide ranging review of the literature on the heritability of intelligence in the late 1990's Neisser et al. (1996) found a substantial role for environmental factors and scope for intervention to address social gaps in intelligence, as well as strong indications that intelligence is a multi-dimensional construct with only partial relationship to life outcomes. A recent update of that review (Nisbett et al., 2012) found even more scope for impact of the environment as more has been learnt about the interaction of genes and environment and about how environments impact on outcomes. The sequencing of the human genome has not led to the identification of specific genes that explain intelligence and, given all of this and other evidence, it should be very clear that outcomes such as intelligence, social class or income are not fixed, innate or immutable.

The issue of course is one of degree: How much difference can environments make? What is a high

degree of heritability? There are few models that go beyond assessment of the degree of heritability and ask what this means for intervention. A standard finding is that the structural genetic heritability of IQ is somewhere between .4 and .8 (Neisser op cit.). However as Bowles, Gintis and Osborne Groves (2008) discuss, these estimates do not measure persistence across generations, they measure the proportion of scores of intelligence statistically explained at population level by the genetic inheritance of children. To go from these estimates about differences between individuals at population level to the assertion that the average gap in scores between children from groups defined on the basis of distal characteristics of parents such as occupation or income is genetic to any fixed degree is stretching the science beyond its basis in fact – because the social class of parents is not equal to their intelligence, which is not equal to their children's intelligence, which is not equal to school achievement. To go beyond this to the assertion that social class itself is genetic is a further false extension. Such a strong hypothesis would surely need substantial evidence including detailed information on how the heritability of diverse characteristics such as intelligence, motivation, character, physical health and beauty interact in practice with actual contexts to generate social outcomes that are correlated across generations. Until we have a clear specification of this social scientific question the heritability estimate is a number in search of a theory as far as its application to the average continuity of social position across generations is concerned.

This extension is social scientific rather than biological because of the lack of common heredity; no claim is made that poor children have a common

and distinct gene pool. The claim that the transmission of SES across generations is explained to any degree by genes is a statement about how children with different genetic inheritances come to achieve common outcomes that are socially structured in times and places by social processes interacting with biological heredity. Yet there is no theory in structural genetics about social process, about how capabilities interact with resources and contexts at multiple levels to influence outcomes, nor about how this changes over time. Nor are there very much data on these things. It is odd that a supposedly biological underpinning of social outcomes should be put forward without reference to evolution as though genes are fixed and capabilities carry value at all times and places in unchanging ways. The nature of the relationship between the social class of one generation and that of the next depends heavily on the nature of the society in which the two generations are studied, so there are no universal truths in the few studies conducted so far – features of society that include banking systems, laws and schools, amongst much else not reducible to genes. Nor is there a theory of how evolution relates to history, of how diverse capabilities play different roles in changing social structures that interact with genes in the generation of biological and social change. The biological sciences of genetics and epigenetics are fascinating and important. Used carefully they yield important clues for social policy and social science but we must look to social and economic research to understand how societies operate in the sharing of wealth and opportunity, the justice or efficiency of our current allocation and what scope there is for change.

Acknowledgements

I would particularly like to acknowledge comments from Stephen Peck on an earlier draft and the ever-wonderful Judith and Max for patience over too many weekends.

References

- Atkinson, A.B. (2015). *Inequality; what can be done*. Harvard.
<http://dx.doi.org/10.1111/j.2050-5876.2015.00834.x>
- Bowles, S., Gintis, H., & Osborne Groves, M., (2008) (Eds.) *Unequal Chances: Family Background and Economic Success*. Princeton
- Carey, N. (2012). *The Epigenetics Revolution*. Iconbooks.co.uk

- Diez-Roux, A.V. (1998). Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *American Journal of Public Health*, 88,(2), 216-222. <http://dx.doi.org/10.2105/AJPH.88.2.216>
- Feinstein, L., (2004). Mobility in Pupils' Cognitive Attainment During School Life. *Oxford Review of Economic Policy*, 20(2) Education. <http://dx.doi.org/10.1093/oxrep/grh012>
- Feinstein, L. & Peck, S. (2008). Unexpected Pathways Through Education: Why Do Some Students Not Succeed in School and What Helps Others Beat the Odds? *Journal of Social Issues* 64,(1), 1-20. <http://dx.doi.org/10.1111/j.1540-4560.2008.00545.x>
- Feinstein, L., Jerrim, J. & Vignoles, A., Goldstein, H. & French, R., Washbrook, E. & Lee, R. & Lupton, R. (2015). Social class differences in early cognitive development debate. *Longitudinal and Life Course Studies*, 6, 331-376. <http://dx.doi.org/10.14301/llcs.v6i3.361>
- Goldstein, H. & French, R. Differential educational progress and measurement error (2015). In Feinstein, L., Jerrim, J. & Vignoles, A., Goldstein, H. & French, R., Washbrook, E. & Lee, R. & Lupton, R. (2015). Social class differences in early cognitive development debate. *Longitudinal and Life Course Studies*, 6, 331-376. <http://dx.doi.org/10.14301/llcs.v6i3.361>
- Jerrim, J., & 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 (Statistics in Society)*, 176(4), 887–906. <http://dx.doi.org/10.1111/j.1467-985X.2012.01072.x>
- Jerrim, J. & Vignoles, A. (2015). Socioeconomic differences in children's test scores: what we do know, what we don't know and what we need to know. In Feinstein, L., Jerrim, J. & Vignoles, A., Goldstein, H. & French, R., Washbrook, E. & Lee, R. & Lupton, R. (2015). Social class differences in early cognitive development debate. *Longitudinal and Life Course Studies*, 6, 331-376. <http://dx.doi.org/10.1111/j.1467-985X.2012.01072.x>
- Lupton, R. (2015). The practice of policy-related research. In Feinstein, L., Jerrim, J. & Vignoles, A., Goldstein, H. & French, R., Washbrook, E. & Lee, R. & Lupton, R. (2015). Social class differences in early cognitive development debate. *Longitudinal and Life Course Studies*, 6, 331-376. <http://dx.doi.org/10.1111/j.1467-985X.2012.01072.x>
- Neale, B. (2011). Qualitative Longitudinal Research. In S. Becker, A. Bryman and S. Ferguson (Eds) *Understanding Research for Social Policy and Social Work*. Bristol: Policy Press.
- Neisser, U., Boodoo, G., Bouchard, T.J. Jr., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D, Loehlin, J.C., Perloff, R., Sternberg, R. & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51(2), 77–101. <http://dx.doi.org/10.1037/0003-066X.51.2.77>
- Nisbett, R.E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D., & Turkheimer, E. (2012) Intelligence: New Findings and Theoretical Developments. *American Psychologist* 67(2), 130 –159. <http://dx.doi.org/10.1037/a0026699>
- Saunders, P., (2011). *Social mobility delusions*. Civitas: London.
- Tu, Y.K., & Law, J. (2010). Re-examining the associations between family backgrounds and children's cognitive developments in early ages. *Early Child Development and Care* 180,1243-12. <http://dx.doi.org/10.1080/03004430902981363>
- Washbrook, E. & Lee, R. (2015). Beyond the Feinstein chart: Investigating differential achievement trajectories in a US cohort. In Feinstein, L., Jerrim, J. & Vignoles, A., Goldstein, H. & French, R., Washbrook, E. & Lee, R. & Lupton, R. (2015). Social class differences in early cognitive development debate. *Longitudinal and Life Course Studies*, 6, 331-376. <http://dx.doi.org/10.1111/j.1467-985X.2012.01072.x>
- Weaver, I.C.G, Cervoni, N., Champagne, F.A., D'Alessio, A.C., Sharma, S., Seckl, J.R., Dymov, S., Szyf, M., & Meaney, M. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience* 7 847-854. <http://dx.doi.org/10.1038/nn1276>
- Willetts, D., (2011). *The Pinch: How the Baby Boomers Took Their Children's Future - And Why They Should Give it Back*. Atlantic Books.

Annex

Figure 1: Average rank of test scores at 22, 42, 60 & 120 months, by SES of parents

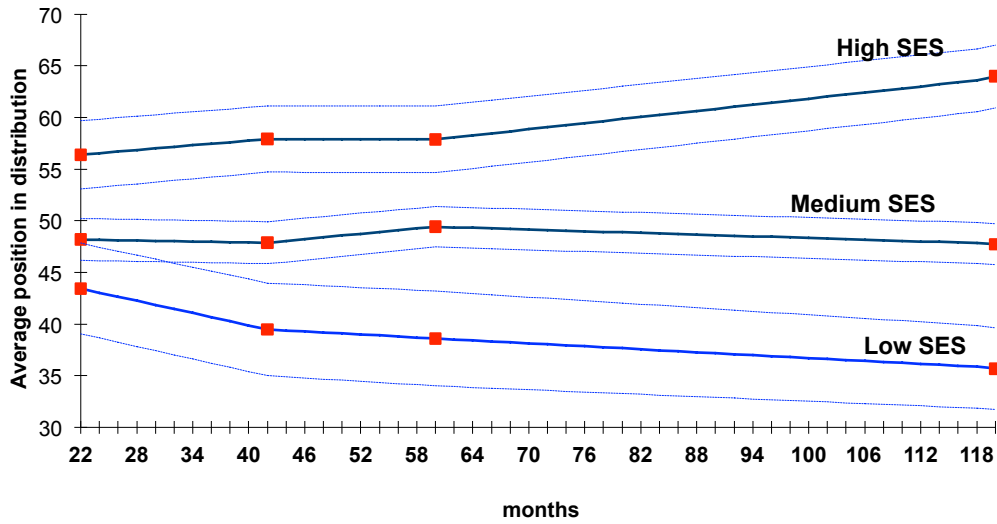


Figure 2: Average rank of test scores at 22, 42, 60 & 120 months, by SES of parents and early rank position

