

Misunderstanding data: Can researchers simplify longitudinal data for policymakers without it leading to errors?

 blogs.lse.ac.uk/impactofsocialsciences/2015/01/21/misunderstanding-data-feinstein/

1/21/2015

Following the comments made on evidence-based policymaking by the Rt Hon David Willetts MP, [Leon Feinstein](#) provides further background on the longitudinal data discussed and defends the findings against some key misunderstandings of the data. For those trying to enhance the use of evidence, an important question is always how to simplify without introducing error and understating uncertainty.



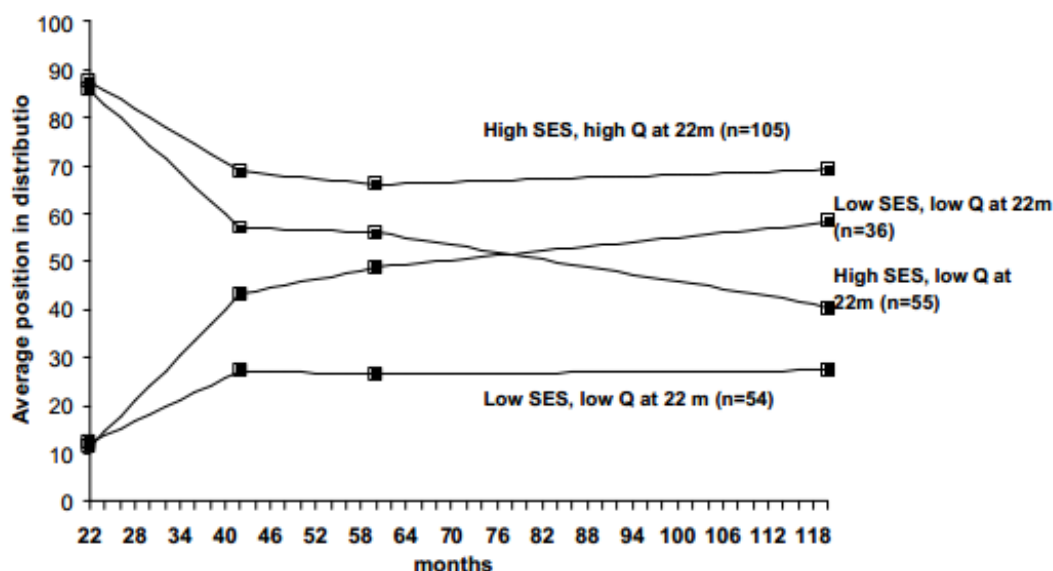
As head of evidence in the Early Intervention Foundation, a What Works centre set up by in 2013 to provide evidence and advice on how to support intergenerational change, I was interested to see remarks by David Willetts in a blog here: “[The messiness inherent to policymaking is a real challenge – can evidence alone outshine tribal instincts?](#)” It became more relevant still when he drew on the example of the public debate concerning Fig 2 in a paper I published in [Economica](#) in 2003, to say that:

Sometimes over-reliance on one specific piece of evidence can leave you vulnerable. I remember being influenced by Leon Feinstein’s very interesting paper for [Economica](#) in 2003 called [Inequality in the Early Cognitive Development of British Children](#). He showed that bright poor kids fell behind rich dim kids by the age of 7. I served on Nick Clegg’s social mobility group and recommended this powerful evidence to him and he too was impressed and cited it. But Leon’s work was challenged by other academics because it was affected by reversion to the mean. The result was that the [Guardian](#) ran a piece that the Coalition’s social mobility strategy was undermined because the research on which it rested had been disproved. That is not, of course, a reason for giving up on evidence-based policy: but it is a reminder of how careful we have to be in using it. [[read full article](#)]

He is kind to say that it is an interesting paper and David Willetts is of course right that over-reliance on one single (he might also have said dated) piece of research is a bad idea. Replication and continual improvement of measurement and evaluation are vital to good quantitative social science. So I have always welcomed the Jerrim and Vignoles correction to the false interpretation of the data that, as he says, “bright poor kids fell behind rich dim kids by the age of 7.” The paper David Willetts cites never says anything about bright or dim kids. Nor there being a specific age of some type of formal crossover. I certainly have never said it. It is a misunderstanding of the data that I was very happy to see Jerrim and Vignoles attempt to correct, but we should be clear that it is a misunderstanding of the data rather than being an error in the data.

For those not familiar, the paper included a graph that shows that children in the 1970 cohort study, who on average scored well on early tests of cognitive development at 22 months in households characterised by low wage and low skill employment, were still on average scoring substantially higher in tests of cognitive skill at 42 months and 5 years than children from households characterised by high wage employment who scored poorly on the tests at 22 months, and were on average overtaken at age 10 in more scholastic tests of school attainment.

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



The definition of SES categories is as for Figure 1 with medium SES children omitted. Children in the 2nd and 3rd quartile at 22 months are also omitted. Standard error intervals are not shown to ease clarity of exposition. Important details of significant differences are given in the text.

Source: Feinstein (2000). Inequality in the early cognitive development of British children in the 1970 Cohort.

In this context reversion to the mean concerns the fact that there is a lot of error in the early test scores, so classifying children into stable “groups” is prone to error. Some of those that score well might do so because they had a good day, but the next day they might not score as well. There is a lot of reversion to the mean in the period between 22 and 42 months. This was always recognised by those who used the graph in policy circles, at least when I was privy to the conversation.

It is not necessary to believe that the groupings are stable, innate or fixed to find the data interesting. The graph shows what happens to the average test scores of different clusters of children in an interaction between an indicator of family of origin and average measures of cognitive development, starting very early in childhood. Low Social Economic Status (SES) children in the UK tended (and still tend) on average to fall back relative to middle class children, whatever the early levels of measured ability.

One thing interesting for policy about the graph was how much change there was between ages 5 and 10 years in the average scores of the different groups. It is important to remember that the tests change through childhood because cognitive development is not a simple, linear, uni-dimensional process. There are no repeat measures of cognitive capability through early childhood. How cognitive capability is manifested changes through development. This may always be true, but is particularly evident in the period before children start school. Motor skills in the data are better predictors of maths and numeracy than more narrowly cognitive and social skills such as use of language. Only at age 10 did the children in the 1970 Cohort Study sit narrow scholastic tests of literacy and numeracy. What the data show is that working class children in the 1970 cohort didn’t translate their early signals of cognitive ability into scholastic achievement at anything close to the same rate as the middle class children.

Now, some might argue from reversion to the mean that this is because the early (even age 5) signals of ability were

simply error, noise. I think it is mainly because the cultures, economy and society into which they were born didn't so effectively lend themselves to school achievement. The graph doesn't prove one way or another, though subsequent work ([Crawford et al. SMCPC](#)) suggests that mean reversion doesn't explain away this remarkable level of impact of social structure on children's development and life chances in this country. It is still a barrier to social mobility, equality and economic growth.

It is clearly true the graph was not a good estimate of the extent to which those with persistent high early scores from low SES backgrounds in middle childhood performed relative to those with persistent low scores from high SES backgrounds. The intended correction is useful for those interested in that question. However, a new element introduced in the Guardian article and related publications was that working class children who score well in the early years are much more likely to be measured in error than middle class children who score well. It is important to say this is entirely unproven as no-one has "true" scores but, to be clear for politicians and others concerned with policy rather than an over-hyped graph, the policy issues were not disputed in the Guardian article. In their corrections for regression to the mean Crawford et al. are very careful to label this "high early performance," to distinguish it from anything that might be thought innate.

There is I think now fairly general agreement that intelligence and school achievement have sufficient fluidity and malleability that only in rare cases is school achievement fixed. [Heckman \(2007\)](#) puts it very clearly, based on his model of the production of capability:

The nature versus nurture distinction, although traditional, is obsolete. Abilities are produced and gene expression is governed by environmental conditions. Behaviours and abilities have both a genetic and an acquired character. Measured abilities are the outcome of environmental influences, including in utero experiences, and also have genetic components.[\[full article\]](#)

It is wrong to interpret this type of longitudinal interaction between early scores and late scores (even if corrected for early reversion to the mean) as the later outcomes of dim or bright children. Shorthand is necessary to engage with busy people. For those trying to enhance the use of evidence an important question is always how to simplify without introducing error and understating uncertainty.

Note: This article gives the views of the author, and not the position of the Impact of Social Science blog, nor of the London School of Economics. Please review our [Comments Policy](#) if you have any concerns on posting a comment below.

About the Author

Leon Feinstein is Director of Evidence at the Early Intervention Foundation (EIF) and Visiting Professor at the LSE's Centre for the Analysis of Social Exclusion. Before joining the EIF Leon was an academic and civil servant. Between 2008 and 2013 Leon worked in the Treasury and the Cabinet Office on policy implementation and performance policy. Until 2008 Leon was Professor of Education and Social Policy at the Institute of Education and Director of the Centre for Research on the Wider Benefits of Learning, undertaking inter-disciplinary quantitative and qualitative research on education and social policy.

- Copyright © The Author (or The Authors) - Unless otherwise stated, this work is licensed under a Creative Commons Attribution Unported 3.0 License.