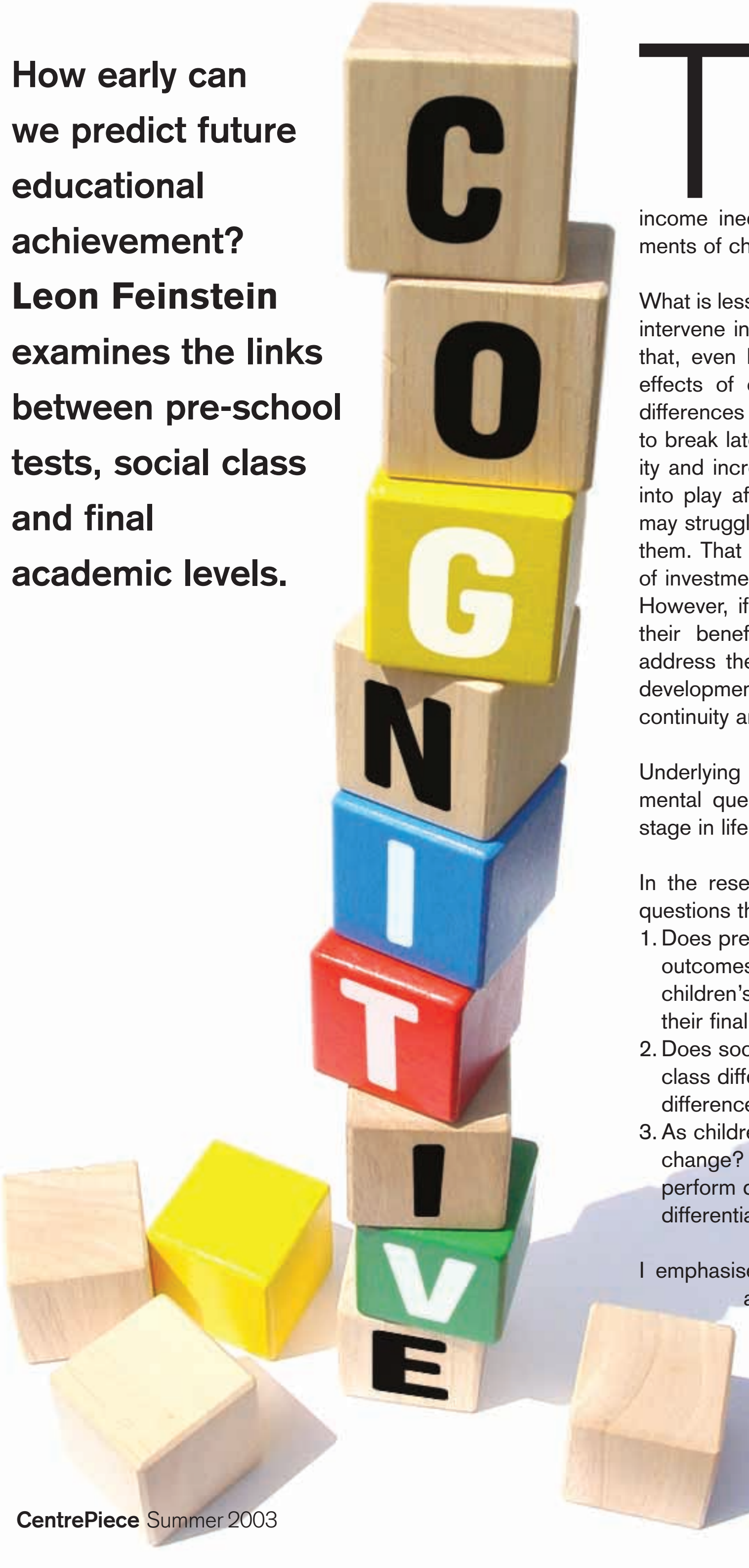
by Leon Feinstein

Very early



How early can we predict future educational achievement? Leon Feinstein examines the links between pre-school tests, social class and final academic levels.

Tony Blair famously made education his government's priority. His motivation seemed to be both economic and moral – the desire to raise the quality of human capital and to reduce inequality. The government also clearly appreciates the inter-generational links between income inequality amongst adults, the educational attainments of children and future income inequality.

What is less clear is the extent to which policy can seriously intervene in these long-term processes. There is evidence that, even before our children reach nursery school, the effects of class difference are already apparent. These differences are not set in stone, far from it, but may be hard to break later . This means that policies to reduce inequality and increase average levels of performance that come into play after children have reached primary school age may struggle to achieve the success that one may wish for them. That is the motivation for the Sure-Start programme of investment in local interventions for pre-school children. However, if the interventions stop at entry to school then their benefits may be equally transitory. Therefore, to address these issues it is necessary to take a long-term developmental perspective with a focus on trajectories of continuity and discontinuity.

Underlying these questions is a more general and fundamental question relating to the timing of policy: At what stage in life are policy resources best applied?

In the research described here I have considered three questions that help in understanding these issues:

1. Does pre-school development matter for long-term outcomes? I assess the extent to which indicators of children's pre-school development are associated with their final educational levels as adults.
2. Does social class matter? I assess the extent of social class difference in these pre-school indicators (a difference sometimes called the "attainment gap").
3. As children get older, how does this stratification change? I assess how high or low initial achievers perform differently as they develop and how this differential development depends on social class.

I emphasise that in this research I make no hypotheses about the underlying causes of these social class effects. Different writers stress different combinations of genetic, economic, psychological or sociological explanations for social class differences but my aim here is descrip-

evidence

tive: to assess the degree of difference in the development of children in broadly defined social class groups. The results are indicative of the policy challenge facing the Government and others concerned with social inequality in education or social exclusion more generally.

The results describe the context within which all interventions in the education system are made but they do not provide answers to other important and perhaps more fundamental questions such as: why do some children do better than others? Do genes matter more or less than environments? Why is there an attainment gap? What particular policies will be most effective?

I have used data from the 1970 Birth Cohort Survey (BCS), a longitudinal dataset of exceptional richness. In particular, it records the results of tests of development given to two sub-samples, when the children were 22 and again when they were 42 months old.

So let us start with some of the evidence (see Table 1). Of particular interest is the data collected when the children were 22 and 42 months old. Due to interest in the effect of foetal malnutrition on brain cell proliferation, a sub-sample of the children was studied at these ages. A 10% random sample of all births was taken, together with those children who were considered to be most at risk from foetal malnutrition. All findings are tested against the control group who had normal levels of risk of foetal malnutrition. The results are found to be representative of this wider population of children.

Although there were over 17,000 children in the full cohort, the research presented here only makes use of information about 2,457 children in the pre-school sub-sample. This includes test scores at four ages (22, 42, 60 and 120 months old) for 1,292 of these children, who form the sample frame used. Analysis suggests that this is a fairly representative sample of the population.

There is one major issue concerning the sample that cannot be overcome. It included only children from two-parent families. This is a potentially serious limit to the representativeness of these results, particularly for those

Table 1. Observations in first four sweeps of BCS

	Total		Test scores	Test scores in common
Birth	17196			
22 month sub-sample	2457		2436	
Random control sub-group		1125		
At-risk sub-group:				
Twins		228		
Post-mature babies		748		
Small-for-dates babies		567		
42 month sub-sample	2315		2297	2045
Random control sub-group		1093		
At-risk sub-group:				
Twins		211		
Post-mature babies		676		
Small-for-dates babies		527		
5 years	13135		11738	1672
10 years	13871		12308	1292
26 years	9003		8395	

The effects of class difference are apparent even before nursery school

concerned with family breakdown. Nonetheless, these data still shed light on questions of the importance and explanation of early ability differences between children of different backgrounds.

At each age, children in the Birth Cohort Survey were assessed by a wide range of tests of intellectual, emotional and personal development. At 22 months the children were asked to complete a range of different tasks: for example, pointing to their eyes to show understanding of language; putting on their shoes, to show personal development; stacking cubes and drawing lines to show locomotor ability. Based on the tests used for screening in child health clinics, these measurements and those made at 42 months were intended to indicate the children's general development. At 42 months, counting and speaking ability were tested, with further copying tests, such as drawing simple geometrical shapes. At age 5, copying was again assessed, together with tests of basic vocabulary.

It used to be thought that cognitive development could not be tested before children were five years old. However, recent work on attention or response to novelty in the first six months of life have shown results here to be correlated with cognitive test scores in later childhood. Nonetheless, intelligence changes qualitatively over early childhood and there is still considerable instability in the recorded scores, because very young children do not stay "on task" for long, because of qualitative changes in the underlying ability being proxied by the different tests, and because all children do not have the same growth rate for cognitive abilities. A child whose IQ score remains the same throughout childhood does not exhibit the same performance at six and 16. Steady gains in ability will be observed, but the relative performance is constant. Conducting analysis on rank position rather than actual scores increases stability in the measurements and that is the procedure followed here.

Table 2. Raw correlation of individual test scores with scores at 120 months

		120 months
	Reading	Maths
22 month scores		
Cube stacking	0.20	0.11
Language use	0.22	0.12
Personal development	0.20	0.13
Drawing	0.15	0.14
42 month scores		
Counting	0.29	0.13
Speaking	0.28	0.17
Copying designs I	0.32	0.16
Copying designs II	0.27	0.14
60 month scores		
Copying designs	0.40	0.19
Vocabulary	0.40	0.18
Human Figure Drawing	0.31	0.13

Table 2 shows the raw correlation of each early test score with later test results at age 10. It can be seen that early test scores, particularly those at 42 months, are associated with later ability, but that there is no particular connection between scores in tests of specific abilities at early ages and subsequent performance in more demanding tests of the same abilities. For example, cube stacking and language scores at 22 months are equally strongly associated with reading at age 10. Similarly, age 10 reading ability is as strongly associated with age 5 copying designs scores as it is with the age 5 vocabulary scores. There is, therefore, no evidence of any functional equivalence, or that any single test score by itself is an obvious candidate to proxy development.

Principal components analysis is a statistical technique that enables us to combine scores and to exploit all of the information available in the test scores without discriminating on unsupported a priori grounds between different tests. I created a single index of development at each age in this way, and took each child's position in the distribution at each age as the set of developmental indicators.

Perhaps the clearest picture of the relevance of the early position in the index of development is given by the results in Table 3, which show how the position in the distribution at each age predicts final qualification level at age 26. Starting with the 22 month panel, the first row shows the highest qualifications at age 26 of the children who were in the bottom quartile at 22 months. For example, 15% of the bottom quartile at 22 months obtained no or only "miscellaneous" qualifications. The third row shows that, of those who were in the top quartile at 22 months, only 8% ended up in the lowest qualifications group at age 26. The z-statistic greater than 2 indicates that this difference is statistically significant.

It is striking that, even measured at 22 months, children in the bottom quartile of this development index are significantly less likely to get any qualifications than those in the top quartile. Moreover, at 42 months more than three times as many of those in the top quartile as those in the bottom quartile go on to get A level qualifications or above. Given the young age of the children tested, these are strong findings. They suggest that, even before children have entered school, substantial signals predicting future educational progress are contained in these standard tests of child development.

Next, in order to get a picture of the stability of the distribution of scores as the children develop, let us look at some transition matrices. These highlight the fact that although there are apparent long-term implications of development at 22 months there is also considerable movement so that it would be wrong to conclude from the earlier results that anything has been set in stone.

Table 3. Age 26 educational and vocational qualifications by quartile position in early development scores.

		Age 26 Highest Qualifications			
		None/Misc	Lower/Middle	A level or higher	Total
22 month					
Bottom Quartile	%	15.00	52.80	32.30	100
Top Quartile	%	8.10	48.60	43.30	100
z-stat on difference*		2.6	1.0	2.7	
42 month					
Bottom Quartile	%	25.80	57.30	16.90	100
Top Quartile	%	6.50	41.00	52.50	100
z-stat on difference*		6.4	3.7	8.4	
5 Years					
Bottom Quartile	%	30.20	51.80	18.00	100
Top Quartile	%	5.20	36.40	58.50	100
z-stat on difference*		18.8	8.3	21.8	
10 years					
Bottom Quartile	%	35.00	53.50	11.50	100
Top Quartile	%	2.40	30.70	67.00	100
z-stat on difference*		25.5	12.6	29.8	

The final row for each panel reports a test statistic for the difference between cell proportions. This has a standard normal distribution, under the null. The z-statistic on the difference in proportions in the first column of the first panel is 2.6, i.e. the difference is significant at 1%.

The transition matrices group children by their quartile position at each of two ages, giving a table of sixteen cells. Because of the degree of instability in scores at these ages, movement between adjacent cells may not be solid evidence of genuine mobility. Perhaps more interesting are movements from top to bottom and vice versa. Table 4, therefore, sets out these large movements. The top panel shows movements from the top/bottom quartile positions at 22 months and the lower panel shows movements from the top/bottom positions at 42 months.

It might be argued that the degree of movement observed would be affected by the over-sampling of children at risk from foetal under-nourishment. If such children were hindered in their early years but subsequently caught up, mobility would be overstated in this sub-sample relative to that in the population. On the other hand, if such children were persistently affected, mobility might be understated. Chi-squared tests have been applied and the results are shown in the Table. These suggest that there is no significant difference between the transition matrices for the full sub-sample and those for the control group.

Standard Errors are in brackets. Only extreme quartile cells are reported, i.e. top and bottom quartiles. The reported chi-squared test is a test of the difference between transition matrices of the control group and full sub-sample. The critical level at 5% with 15 degrees of freedom is 25.0

The first row in Table 4 shows that of the 25% children scoring lowest at 22 months, 39.9% were still in the lowest quartile at 42 months. On the other hand, 13.7% had entered the top quartile. By 10 years, from these sample data, a child in the bottom quartile at 22 months would have a probability of 0.42 of still being in the bottom quartile at 10 years and only a probability of 0.15 of reaching the top quartile.

There is even clearer persistence of scores between 42 months and ten years, particularly in terms of the proportion of large movements. However 10% of the bottom group at 42 months had reached the top quartile by age 10. This emphasises the need to interpret the development indices as signals of development and not as stronger classifying mechanisms. Children can still catch up and overtake those who have out-performed them early on. Nonetheless, the 22 month and 42 month scores provide a powerful guide to subsequent performance. This is particularly the case for the 42 month index and other experiments have shown that the index is slightly more predictive for girls than for boys.

Our second basic question was whether the results of these pre-school indicators are stratified by social class. Figure 1 maps the average position of children from different social backgrounds in the distribution of test rankings at the four survey ages.

Social class classifications are made here on the basis of both parents' occupational classification (Socio-Economic Status, SES) at the child's birth. No allowance is made for

Social class stratification becomes more extreme by ten

Table 4. Selected cells from quartile transition matrices

		Quartile at 42 months		Quartile at 10 years		
		Bottom	Top	Bottom	Top	Obs
Quartile at 22 months	Bottom	39.9	13.7	41.5	15.4	304
		(1.4)	(1.0)	(1.4)	(1.0)	
	Top	10.8	43.5	13.4	34.6	306
		(0.9)	(1.4)	(1.0)	(1.4)	
Chi-squared (dof=15)			2.5		8.7	

		Quartile at 10 years		
		Bottom	Top	Obs
Quartile at 42 months	Bottom	44.8	10.1	306
		(1.4)	(0.9)	
	Top	6.2	43.8	306
		(0.7)	(1.4)	
Chi-squared (dof=15)			10.8	

Notes: Standard Errors are in brackets. Only extreme quartile cells are reported, i.e. top and bottom quartiles. The reported chi-squared test is a test of the difference between transition matrices of the control group and full sub-sample. The critical level at 5% with 15 degrees of freedom is 25.0

their changing occupational classification over time because it is not possible to differentiate between genuine changes and the results of miscoding. In any case, social class at birth provides a fairly good indicator of the material, genetic and educational inputs that the children can be expected to receive through childhood.

Dotted lines represent intervals of two standard errors. The definition of categories with sample observations is as follows: High SES – father in professional/managerial occupation and mother similar or registered housewife (307 obs.) Low SES – father in semi-skilled or unskilled manual occupation and mother similar or housewife (171 obs.) Medium SES – those omitted from the high and low SES categories (814 obs.) Thus, children whose mothers were housewives were categorised by the SES of fathers.

At 22 months, the difference between the average rank of children in the top and bottom social class groups is 13 percentiles. At 10 years, the difference between top and bottom groups is 28 percentiles. The average rank of the low SES group falls over time, but this does not mean that actual development has been retarded. The interpretation of this change as one of increasing polarisation must be tempered by the fact that the ranks are positions within the distribution of different tests at the two ages. The findings may be explained either by declining relative performance or by the fact that the later tests are more effective at discriminating between children. It may also be that the later tests assess tasks at which children in the low SES group are less able.

However, in a seminal 1983 paper, R.S. Wilson looked at the IQ scores of twins between the ages of 3 months and 15 years. For identical twins, the correlation between scores at 24 months was 0.81. For non-identical twins the correlation was 0.73. By age 15, the correlation coefficients were 0.88 and 0.54, respectively. Wilson argued that, as children mature, the genetic component of performance in the tests becomes more dominant and that, as with the development of height, the action of the genes in this respect is not completed until adulthood.

The pattern of polarisation shown in Figure 1 is, therefore, not surprising whether one tends towards a genetic or environmental explanation. Crucially, the graph clearly shows that, although children are already stratified by social class in standard tests of intellectual and personal development at 22 months, this stratification has become more extreme by age 10 as assessed by the standard tests for academic development appropriate at that age. There is certainly no evidence here from the late 1970s that entry into schooling in any way overcame the polarisation of children's educational achievement linked to the deepening effects of parental background.

Figure 1, however, only shows the average rank positions of three fixed groups of children as they matured. There is considerable variation within the groups. This is brought out in Figure 2, which groups children not just by their family background but also by their 22 month quartile position. The advantage over Figure 1 is that this shows something of the distribution within the SES groups and so helps us see how social class impacts on children's developmental trajectories.

The definition of SES categories is as for Figure 1, with medium SES children omitted. Children in the 2nd and 3rd

Turnover of staff must be kept to a minimum

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

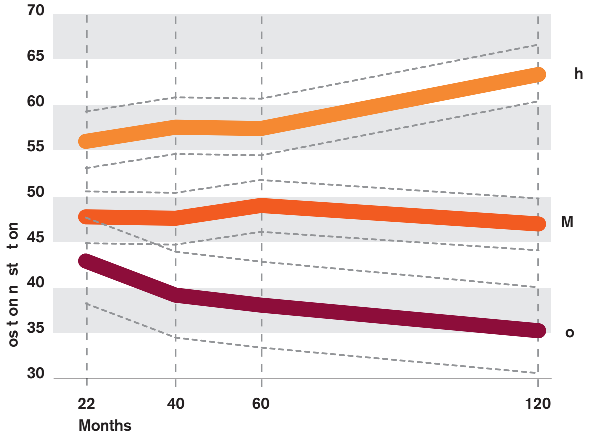


Figure 2. Average rank of test scores at 22, 42, 60 & 120 months by SES of parents and early rank position

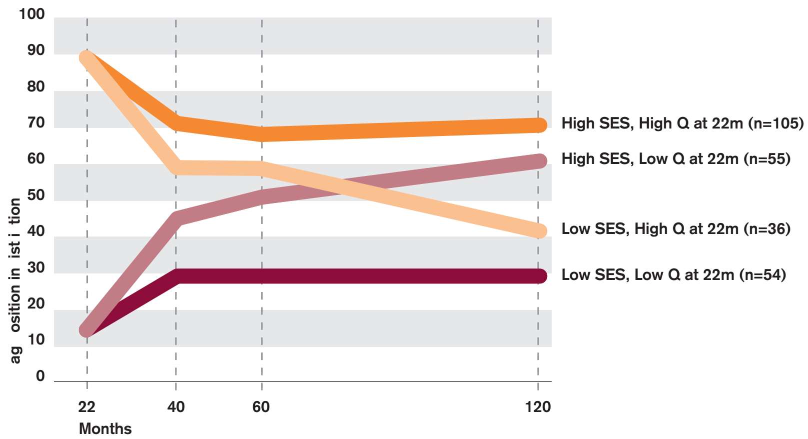
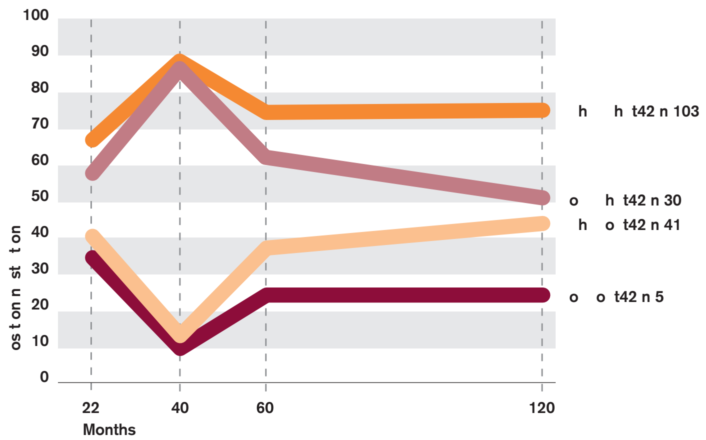


Figure 3. Average rank of test scores at 22, 42, 60 & 120 months by SES of parents and 42 month rank



quartile at 22 months are also omitted. Standard error intervals are not shown increase clarity.

The Figure shows that having a low test ranking at 22 months does not matter decisively for a child's future position in the distribution unless the child has low SES parents as well, in which case the position is unlikely to improve greatly. Furthermore, a low SES child with a top quartile score at 22 months is predicted to fall behind high SES peers who had low quartile scores at 22 months. Thus, early scores matter and low SES children are less likely to have high early scores but even if they do they are very likely to lose this early advantage.

This conclusion is supported by the separate transition matrices of low and high SES children. Of low SES children who were in the bottom quartile at 22 months, 60% were still there at age 10. On the other hand, high SES children, who happened to be in the bottom quartile at 22 months, were more likely to be in the top quartile at 10 years than still to be in the bottom quartile.

Does this suggest that early scores do not matter? The answer is no for two reasons. First, as Figure 2 showed, it is still the case that children within each of the SES groups who are in the top quartile at 22 months score better at age 10 than children in the same SES group who were in the bottom quartile at 22 months. At age 10, the difference is still 13 points in the distribution for the low SES group and 11 for the high SES group.

Second, as Table 4 showed, for the middle SES group (i.e. the majority of children), those in the bottom quartile at 22 months are significantly more likely to get no qualifications than children in the top quartile and significantly less likely to get A levels or higher qualifications. For the top and bottom SES groups, differences at 42 months strongly predict final educational qualifications. So, in combination with the SES factor, the pre-school scores still matter. As well as influencing early performance, family background clearly plays an important role in determining the continued development of children's ability.

So early scores do matter but so does social class after early childhood. The lesson for policy makers is clear. There is mobility (as one would expect) after 22 or 42 months, but upward mobility is mainly for high or medium SES children. Low SES children do not, on average, overcome the hurdle of lower initial attainment, combined with continued low input. Furthermore, social inequalities appear to dominate the apparent early positive signs of academic ability for most of those low SES children who do well early on.

In answer to the questions posed at the beginning of this article, I find that pre-school, academic development does matter in the sense that it predicts final educational success. However, I also find that large social effects continue to impact on children's development after they have entered school. The benefits of good early development persist but can be substantially eroded by social class effects.

This is a measure of the challenge facing the government in addressing inequality through the education system. It must intervene early to impact on these important features of early development but even having done so it must continue to intervene.

The Sure-Start programme is, therefore, to be welcomed but redistributive educational policy cannot stop at entry to school. Research on the effects of pre-school initiatives has shown that significant success can be had but not on the cheap. Staff must be well paid, highly motivated and able. The turnover of staff must be kept to a minimum. Most importantly, the effects of programmes that start pre-school but do not continue for at least the first two years of primary school are likely to be transitory. It is to be hoped that Sure-Start can reproduce the positive features of successful pre-school programmes (on-going evaluation will report on this) but also that it will establish a platform for the continuation of policy into primary school. To address the social class attainment gap, programmes must be developed that confront the apparent disadvantages of children from low SES families even after they have entered school.

Leon Feinstein is Research Director at the Centre for Research on the Wider Benefits of Learning at the Institute of Education and a research associate in the CEP labour markets programme.

This article is based on his Discussion Paper No 404, available from the Centre for Economic Performance.

References & further reading

M.H. Bornstein and M.D. Sigman, "*Continuity in mental development from infancy*", Child Development, 57, 251-74.

S.L. Ramey and C.T. Ramey, "*Early childhood experiences and developmental competence*", in J. Waldfogel and S. Danziger (eds), Securing the Future: Investing in Children from Birth to College, Russell Sage Foundation, 2000.

J. Waldfogel, "*Early childhood interventions and outcomes*", Paper 21, Centre for the Analysis of Social Exclusion, LSE, 1999.

R.S. Wilson, "*The Louisville Twin Study: developmental synchronies in behaviour*", Child Development, 54, 298-316, 1983.

C.H. Zeanah, N.W. Boris and J.A. Larrieu, "*Infant development and developmental risk: a review of the past ten years*", Journal of the American Academy of Child and Adolescent Psychiatry, 36(2), 165-78, 1997.