From: Child Development, Information, and the Formation of Public Policy - an international perspective, Thomas E. Jordan (Editor) Published by Charles C. Thomas, Springfield, Illinois 62717, USA (1982)

7)

Chapter 3

CHILD DEVELOPMENT STUDIES IN EUROPE

Harvey Goldstein

Introduction. Mednick and Baert (1981), in their recent survey, list nearly 70 child development studies in western Europe, to which can be added several more from my personal knowledge. It would be pointless to attempt to summarise all these studies once more. Moreover, they are not all equally useful or informative and some selection criteria become necessary. The criteria I have adopted are not ideal and they exclude from consideration some studies which are of interest in certain areas. Nevertheless, I have tried to keep in mind the overall theme of this volume in choosing criteria, as well as some more general considerations of methodological soundness.

A useful categorisation of studies is into those which attempt to examine representative samples of well defined populations or subpopulations and those which are 'opportunity' samples, usually to pursue a few topics at some depth. Partly because my own competence and experience largely lies with the former type of study and partly because I believe that policy should tend to be influenced by the former rather than by the latter, that is what I will concentrate on. This is not to say that I consider the latter to be unimportant; far from it. Such studies are probably essential in the development of any field of enquiry, but they are typically exploratory, hypothesis generating and pointers for further, perhaps confirmatory, research. Any conclusions they reach will usually be tentative and therefore not, in general, suitable for policy decisions. Indeed, to base policy decisions on the results of such studies, when they may be highly unrepresentative, can

TABLE 12

REPRESENTATIVE LONGITUDINAL STUDIES

	Study	Population Sampled	Year of Birth	Ages of Measurement	Approx. Sample Size	Types of Measurement
1.	Aberdeen Child Development Study	Aberdeen	1956-1960	6-10, 8-12	14,000	Educational, social, medical, psychological
2.	Aberdeen Delinquency Study	Aberdeen	1950-1954	12-16 17-21 21-25	11,000	Social, criminal, educational
3.	Isle of Wight Study	Isle of Wight (UK)	1953-1955	10,11,15 years	500	Psychological, behavioral, social
4.	London Child Development Study	West Central London	1950	1,3,6,9,12, 18, 24 months and yearly to 18 years	190	Medical, physical, psychological, social
5.	National Survey of Health & Development	Britain	1946	4 - 6 weeks 2,4,6,7,8,9, 10,11,12,15, 16,17,18,19, 20,21,22,23, 24,25,26 years	5,400	Physical, medical, social educational, psychological, environmental
6.	National Child Development Study	Britain	1958	Birth, 7,11,16,19, 23-24 years	17,400	Physical, medical, social educational, psychological, environmental

(Table 12, cont.)

7.	Child Health and Education Study	United Kingdom	1970	Birth, 2 1/2,5, 10 years	17,200	Educational, social medical, psychological, physical, environmental
8.	Census Longitudinal Study	United Kingdom	19th Century onwards	Various, according to significant events	500,000	Medical, demographic, social, environmental
9.	Iceland Child Development Study	Reykjavik	1950-1961	5-15 (various) 25-30 (follow-up)	1,100	Psychological, social, medical, educational
10.	Malmo Study	Malmo (Sweden)	1928	10,14,20, 30-37, 43-45, 46 years	1,500	Psychological, educational, social, criminal, second generation
11.	Project Metropolitan Sweden	Stockholm	1953	Birth, 10,13,15,18, 21 years	15,100	Criminal, psychological, social, educational
12.	Orebro Child Child Development Study	Orebro (Sweden)	1955	10,13,15,16,17 18,19 years	1,000	Educational, psychological physical, social, medical

(Table 12, cont.)

13.	Swedish Urban Growth Study	Urban Sweden	1954-55	9-17 years every 6 months, 18 yrs	740	physical, medical, social
14.	Copenhagen Schizophrenia Study	Copenhagen	1950-1960 (approx.)	Birth-25 years (Various)	310	Psychological, medical social
15.	Danish Birth Cohort Study	Copenhagen	1959-1961	Birth, 1 year 11 years	48 (out of 7,400 measured at birth)	Neurological, psychological, behavioural
16.	Project Metropolitan Denmark	N. East Zealand (Boys)	1953	Birth, 12,13 15,22,23 yrs	12,100	Criminal, social, demographic, educational psychological
17.	Danish Twin Register	Denmark (Twins)	1881-1930	Various according to significant events	40,000	Genetic, medical
18.	Danish Adoption Register	Denmark (adoptees)	1910-1945 (approx)	Various according to significant events	5,500	Psychological, social medical
19.	Finnish prospective epidemiological study	Helsinki	1975-1976	Birth, 0-6 months 6-8 months	5,500	Medical, Social Psychological

(Table 12, cont.)

20.	Helsinki Cohort Study	Helsinki	1955	Birth, 14,20 years	6,800	Physical, medical, social psychological
21.	Nymegen Growth Study	Nymegen (Holland)	1961-1967	Every 6 months 3-14 years (overlapping cohorts)	500	Psychological, dental, physical, social, medical
22.	Ghent Developmental Screening study	Ghent (Belgium)	-	3 - 9 months	560	Developmental
23.	Brussels Child Developmental Study	Brussels	1955-1958	1,3,6,9,12,18 24 months and every year to 15 yrs, 18 yrs	270	Social, behavioural, educational, psychological
24.	Swiss Child Development Study	Zurich	1955	1,3,6,9,12,18 24 months and yearly to 18 years	400	Medical, physical, social psychological
25.	Prague Child Development Study	Zizkov (Prague)	1956-1960	1,3,6,12 mths then 1/2 yearly to age 20	290	Psychological, educational, physical, medical, behavioural, social
26.	Romanian Child Growth Study	Romania	1932-1971	0-18 (not longitudinal. Three cross sections 1957, 1964,1971)	approx. 975,000	Physical, Nutritional

be very damaging. In addition, where such studies are longitudinal, the lack of representativeness will often be more severe than for cross-sectional studies.

The distinction I have made is not to be confused with the distinction between large and small studies. While it is probably true that 'representative' studies do tend to be larger, some non-representative ones are also large. For example, the important study of determinants of educational progress by Rutter et al (1979) involved over 2,000 school children, but was confined to twelve schools in one small area of inner London. As such, and despite its relatively large size, it is not representative of any useful population and in my view is rather unsuitable for policy, although of course it raises a number of important research issues which have begun to command the attention of the educational research community.

In addition to requiring representative studies, this review will largely, although not exclusively, be concerned with longitudinal ones. This is not because I believe cross-sectional studies to be unimportant, but because only longitudinal studies can begin to answer the more interesting questions about child development.

Table 12 lists those studies with which I shall deal, along with the population sampled, the year of birth of the sample children, the ages of measurement, sample size and types of measurement. Detailed descriptions of most of these studies are given in Mednick and Baert (1981). Where they are not described there, separate references are given. It is worth remarking that the categorization of studies used by Mednick and Baert is somewhat confusing because some studies, such as

that of Rutter et alii referred to above, are listed as 'representative' which they are not; others, such as the Brussels study which is representative, are listed under 'non-representative' studies.

POPULATIONS REPRESENTED

It will be seen that 18 of the 26 studies in Table 12 are actually representative of only one region, city or area of a city. Furthermore, most of the studies sample only a single birth cohort, so that there is no information on secular trends. Of the remaining 7, four are located in Britain or the United Kingdom, one in Sweden, one in Romania, and the other two are specialized studies of twins and adoptees in Denmark. The Romanian study is only cross-sectional, and the Swedish one is only of Urban children, leaving Britain as the only European country with representative longitudinal studies of the whole population. I shall offer some tentative suggestions as to why this is the case, drawing upon the designs and results from all the studies in Table 12, and attempt to draw some policy conclusions.

I had originally felt that it would be useful to try to make some general statements about knowledge of child development in a comparative analysis. On a more careful look at the various studies, however, it became clear that this was a formidable task which lay beyond my present resources. Indeed, the coordinated International Children's Centre Studies (Falkner, 1980) which were explicitly set up to facilitate comparisons have achieved less than expected in this area despite frequent contacts and an agreed baseline of measurements. The reasons for this, presumably, lie partly in the usual problems of highly stretched

resources, but, I suspect, it is also a product of a current research ethos which places a high value on 'original' data and a 'progression' of research themes.

Why are there so few nationally representative longitudinal studies? In fact, one should probably be more surprised that there are any representative studies at all. To begin with, nationally representative cross-sectional studies are difficult to organise successfully, and the added complexity of a longitudinal study means that a considerable effort has to be maintained over a long period of time. It is of some interest that all three child development cohort studies in Britain have each had one or two dedicated individuals associated with them for nearly all their existence. In addition, the structure of the British welfare state undoubtedly made the study organisation simpler and enables all the studies to be carried out relatively cheaply. As much as anything, perhaps, the existence and success of the very first such study, namely the 1946 cohort, set a pattern which it was subsequently rather easier to follow with other studies. Whatever the reason, it remains the case that these studies are the only ones which enable us to make confident statements about longitudinal changes for a whole nation. While it has long been recognized in the case of cross-sectional studies that inferences for sub-populations in cities or regions have an important but limited usefulness, this is rarely discussed for longitudinal studies. Furthermore, for most of the longitudinal studies in Table 12, the population sampled is that which was resident in the city or area at one time with no follow-up of those who moved away. Since such individuals may possess different developmental patterns, this will

limit the usefulness of many results. A further consideration is that longitudinal studies are usually carried out on 'convenient' and 'privi-ledged' populations - such as those near research institutes or in capi-tal cities. Like much social and epidemiological research, the most deprived areas tend to be omitted because they have few resources and cannot easily support the burden of long term research studies.

An illustration of the hazards of using results from unrepresenta-tive samples can be found in attempts to estimate the correlations between the heights of children and their parents. Specifically, some of the early studies (e.g. Thomson, 1955, Bailey, 1954) reported higher correlations between parents and children of the same sex then between parents and children of opposite sexes. More recent data, however, from the 1946 and 1958 British Birth Cohort Studies (Goldstein, 1971) have shown, at least for mothers' heights, that there is no sex difference in these correlations. It would seem that one important implication for research policy here is the need to consider carefully the establishment of valid total-population inferences from longitudinal research rather than geographically limited ones.

COMPARISONS OVER TIME

One of the justifications for having three British Birth cohorts (1946, 1958, 1970) is the potential they provide for comparisons over time. When one looks for such comparative analyses, however, one finds rather few. As remarked above, this is, presumably, partly to do with resource constraints, but also seems at least in part to be a realisa-tion that making comparisons over time is not as straightforward as might at first sight appear.

One of the two major obstacles to comparisons over time lies in the nature of much of the data collected. When we attempt to measure, say, educational variables we will often need to devise new or special measuring instruments which are valid for the environment or time period when they are used. For example, when measuring reading attainment it would be inappropriate to contrive to use the same test over a long period of time if written language use and school curricula are changing. Also, it may be difficult to justify using the same instrument in different environments, for example countries, at the same time period. Furthermore, as behavioral science develops new and superior instruments may appear, and there may be a strong case for using these rather than the old ones used in previous studies. With medical measurements also, definitions can change as can diagnostic practice both across time and nations. Basic anthropometric measurements on the other hand tend to remain stable, and it is no accident that for these we have the most extensive and reliable data across time and populations. In some ways these are more reliable even than mortality statistics, especially when determining 'causes of death' which will often reflect definitional differences between countries. The other obstacle is the general diversity of procedures for collecting, recording and processing information which can easily frustrate any attempts at comparisons.

The above comments are not intended as an argument against attempts to make comparisons between studies, rather as a note of caution about what can be achieved. It is particularly pertinent to ask whether a sharing of experience might not result in better designs which can allow valid comparisons. In order to see what might be achieved it is worth

looking at perhaps the most concerted attempt to provide international comparisons in Europe, namely the International Children's Centre coordinated growth and development studies (Falkner, 1980).

In the early 1950's, a series of longitudinal studies (five in Europe) were begun, using a commonly agreed baseline of social, medical, psychological and physical measurements, with common ages of measurement. It was planned that the study teams would meet regularly to pool their experience. These first objectives seem to have been fulfilled, and a measure of the success of the whole venture is the continued existence of the studies into the late 1970's. The regular meetings in particular led to a considerable and useful exchange of ideas resulting in a great deal of learning from each other. In fact, my own experience from attending six of these meetings is that the most useful lessons learned from comparisons of populations are to do with approaches and methodology rather than with comparisons of numerical estimates. Of course, there are exceptions to this and some of the comparisons of the physical and developmental measurements are interesting (Falkner et al, 1980; Hindley, 1980). Unfortunately, however, even these are of limited utility without adequate studies which probe reasons for the differences found. Such studies would not only need to collect relatively large amounts of more detailed data, they would also need to be representative of their respective populations, as I have argued. An interesting series of studies which moves some way in this direction is the 1973 WHO Perinatal Collaborative Project (WHO 1978) which used vital registration data from eight countries (four of them in Europe) for a whole year to compare mortality rates, birthweights etc., and to see how far any

differences could be explained by socio-demographic factors. While it is still too soon to judge the success of this project it would seem to offer a methodology for a series of representative longitudinal studies, possibly along the lines of the U.K. Census longitudinal study, which is following a 1% random sample of individuals from each national decennial census.

FUTURE DIRECTIONS

A glance at Table 12 and the knowledge that most of the studies listed there remain uncoordinated, lack regular contacts between researchers, collect data in different ways and often are unaware of each other's existence, should make us pause for thought. This is especially so when we recall that these studies have been selected because they are those which were actually intended by their founders to be representative of well-defined populations. The International Children's Centre studies succeeded in achieving a constructive interchange between studies without severely limiting the ambitions and development of each. To do this again on a larger scale with representative samples would not be impossible, although expensive. Nevertheless, it may be no more expensive than the sum total of the separate studies now in existence.

One possible prototype, the U.K. longitudinal study, has already been mentioned. This has certain difficulties in that its use of census data requires the close cooperation of government departments, but its design seems worth exploring further. Another model is the large scale cohort study along the lines of the 1946, 1958 and 1970 British cohorts. Because the size of such a collaborative undertaking and its expense are somewhat daunting, the coordination of such a project would probably

require the backing of international organizations such as WHO, UNICEF and UNESCO, but the potential payoffs could be very large indeed. Of course, the precise design of the studies probably should not copy that of the British cohorts, and in particular should involve more sophisticated population and time sampling methods. There is also no reason why such studies should be confined to Europe; for example Cuba now has a cohort study modelled on the British cohorts. At the very least, like the International Children's Center studies, the process of running such studies and meeting to discuss them ought to provide a more efficient use of research resources than is presently the case. A reasonable expectation is that a series of such studies would, like their British prototypes, enormously increase our overall knowledge of the development of child populations. While they would not entirely replace more limited specialized studies, they could act as reference points from which such studies could develop.

With modern computing techniques there should be little difficulty in rapid interchange and comparisons of data. Because the studies are long term they will need to have secure local institutional support, and University departments would be obvious locations, although well established independent research organizations might equally be suitable. A secure long-term source of funding at least for core staff seems essential, and is something which only the 1946 British Cohort seems to have achieved successfully.

There is, I believe, sufficient expertise available to initiate and sustain a programme of internationally coordinated representative child development studies and, at the very least, a small initiative to explore its feasibility seems well worth funding.

ACKNOWLEDGEMENT

I am most grateful to Neville Butler for his helpful comments on an early draft of this paper.