### DISCUSSION

# Craig T. Ramey, Frank Porter Graham Child Development Center, Univ. of North Carolina

As Cronbach (1957) has pointed out some years ago, research within the domain of the social and behavioral sciences tends to be either correlational in nature or experimental in approach. Long and productive debates have raged concerning which of these approaches generates the more useful kinds of data with which to begin to alleviate major social problems. At the one extreme, complex social systems defy adequate description when elementary, univariate statistics are used as decision-making criteria. Therefore, in many cases, it appears that the best that the biosocial scientist can do, when dealing with large social issues is to find patterns that can be extracted from matrices of intercorrelations of many variables. Yet as every elementary student of statistics knows, the dictum that correlation does not imply causation serves as a continual bane to the full understanding of the directionality of social forces. As a consequence, there are strong arguments for developing an experimental approach to the understanding of human behavior even within the context of complex social systems. Yet as Heisenberg (1962) has pointed out so clearly, the very process of measurement distorts the phenomena to be observed.

A frequent suggestion to those concerned particularly with the understanding of the development of human behavior within complex social systems is to study the development of groups of individuals over time; i.e. to begin a longitudinal investigation of development which will allow one to describe not only the behavior of groups of subjects and the impact that various environmental influences may have on that group's development, but also a longitudinal investigation allows one to examine the course of development of specific individuals within that group. Parenthetically, educators are now strongly suggesting that we spend more time understanding our treatment X subjects interactions than we have historically.

While, in principal such an approach seems to be a reasonable and even highly a desirable method of procedure, the implementation of adequate research designs to understand highly complex social phenomena is extremely difficultat best, and at worst impossible--given our current state of knowledge for decision making rules and techniques for drawing causal inferences. With respect to the desirability of longitudinal research in general, as compared to cross-sectional approaches, it must be pointed out that several large and torturous traps lie in the path of even the most dedicated and competent of researchers. Again, it is elementary to point out that one's findings truly do not generalize beyond the population from which the sample has been selected. Thus a study performed in Milwaukee or in Chapel Hill, or in New York City, or in Berkeley does not of

necessity generate data on human development which is directly extendable to other communities or to other countries; yet, one hopes that certain basic principles of development are so general and so pervasive in nature that one may safely and practically proceed to make such generalizations even though he realizes that in theory he must do so with extreme caution.

A morass of other problems plague the researcher who begins a longitudinal study and many of these have been pointed out in an excellent and thoughtful series of methodoligical papers by Warner Schaie (1965) and his collegues including Paul Baltes and John Nesselrode, among others. However, much as these problems are with us, we are left on the horns of the dilemma. Do we seek not to engage in broad scale research which can in fact generate information relative to the human condition, because we dare not violate serious assumptions? Or do we recognize that in many cases we are violating the assumptions and proceed to make the best analysis that we can of our rather imperfect data?

Rick Heber and Howard Garber and their staff are to be highly commended for having the audacity to design an experimental program which seeks to resolve a large part of the ubiquitous nature-nurture controversy. They are particularly to be applauded for attempting to resolve this problem, not within the hothouse of the sheltered laboratory (which would be hard and meaningful enough if successfully accomplished, even if the generality of the finding were limited) but, rather, they have sought to generate data relevant to the heredity-environment controversy within the most meaningful of all possible realms - the realm of real life, with real people, with very real problems. Further, those problems are embedded within a social context which itself is nearly impossible to comprehend and to describe precisely. Thus, in all sincerity we must praise the magnitude of the effort and, because the problem is so vitally important, we must also bring our best acumen to bear upon the adequacy of their methodology, the precision of their data analyses, and the validity of their conclusions.

#### Methodology

The primary justification for their subject selection rests upon the finding from a cross sectional design study which purportedly shows that the IQ's of children whose mothers score below 80 on an IQ test show a progressive decline in intellectual performance when the children are between 13 and 168 months of age. Warner Schaie (1965) has sufficiently cautioned us against such straightforward extrapolation of longitudinal trends from cross sectional analyses that we must all realize the fragileness of such extrapolation. Although the papers from the Heber-Garber project are not written in sufficient detail concerning the sampling and testing procedures for me to fully evaluate the validity of their conclusion, it seems to be a plausible one. In any event, for the sake of discussion I will tentatively accept their findings as valid.

As a result of this finding the authors selected families for inclusion into their intervention program whose mothers scored less than 75 on an intelligence test. Forty such families were identified and assigned either to an experimental or a control group. The number of families contacted, the refusal rate, the demographic differences between those selected and those who accepted as far as I can determine have not been presented. The method of assignment of these families is equally crucial for the interpretation of their findings. Yet from the papers that I have had access to, there simply is not enough information to evaluate the adequacy of their assignment procedures. Were the families randomly assigned? Were they matched on a priori determined variables and then randomly assigned? Or was some other procedure employed? It is imperative that we know the answers to these questions because almost all subsequent statistical analyses make assumptions in this area. The consequences of violating these assumptions in any particular sample simply cannot be determined.

### Data Analyses

Although one wants to ask many detailed questions about specific data analyses, there appears to be one pervasive issue which leaves one unsettled. The essential nature of this project is to take many measures on the same individuals over time in what is essentially a repeated measures design. Although it is not spelled out in the written version of Dr. Garber's paper, I presume that Winer (1971) type analyses of variance with repeated measures formed the bulk of their analyses. The criticism which I am about to make must be understood in context. It has been seven years since this project was begun and many statistical innovations and refinements, thanks to many of you in this audience, have been made in the meantime. Nevertheless, if such analyses were performed, they must be reevaluated in light of recent criticisms by McCall and Applebaum (1973) among others. For example, unless one can assume homogeniety of the covariance matrices in a repeated measures design, then it is likely that one inflates his alpha level beyond what he has actually chosen. Various methods of protecting one from resulting type I errors exist to minimize such risks. For example, one can use the Box (1954) correction for the degrees of freedom in univariate analyses or one can use the MANOVA program with the Wilks Lamda Criterion which does not require an assumption of homogeneity of covariance. Several other analysis questions could be offered but I shall limit myself to these two because of time constraints.

#### The Intervention Program Itself

Seven years ago the prevailing professional opinions on the expectations of the possible outcomes of intervention programs were divided into basically two different camps. On the one hand there were those who felt that specific deficits in language and cognitive development could be remediated through some relatively short term compensatory experiences which would significantly enough alter the course of intellectual development, such that children who were already developmentally retarded would be remediated and innoculated against failure in future academic settings.

On the other hand there were those who felt that only by a drastic alteration of the base and deplorable social forces which press unbearably upon the oppressed and the poor, could the bright futures, which socially conscious Americans advocate as the birthright of all citizens, be achieved.

The disappointing results from the short term remedial intervention attempts lay waste to the ideal of a simple solution to a complex social problem. Even before the Headstart results were tabulated, the Milwaukee group was quietly moving toward the enormous and consuming task of beginning at the beginning. An in medias res solution was given way to a reexamination of genesis. With little in the way of articulated theory of early and global human development and with even less help in the form of curricular products from educators, the bold conception of starting essentially at birth was born. As one who is now currently trying to evolve a theoretically generated curriculum for ensuring optimal early development, I can more than sympathize with Heber's often quoted and much criticized statement that the Milwaukee curriculum consisted of every thing but the kitchen sink.

Because the curriculum and the other components of the project are essentially the independent variables in this and other intervention programs, we, as scientists, are left with feelings of dissatisfaction about the project and its replicability. Harsh and severe criticism can and will be levied against the imprecision of the specification of the treatment variables. It is, and will continue to be, nearly impossible to determine which variance components have been accounted for by the improved nutrition which the experimental group undoubtedly received; by the potential Hawthorne effects associated with being in the experimental group and by the myriad other variables which might have influenced the differential performance of the two groups.

Such criticisms can and should be made. They should be made because only through such a process of critical examination will the truly important manipulable variables be isolated and examined. Yet one should temper one's criticisms with the realization that science proceeds from the molar to the molecular levels of explanation. Newton's work had to preceed Einstein's and Einstein's had to preceed Heisenberg's. So too we in the bio-social sciences can and must proceed.

We now know empirically through projects such as this one that the course of human behavior can now be profoundly altered; that developmental retardation is not the birthright of even our most oppressed and disadvantaged children. What remains to be done now is to move from the level of description of the consequences of social and individual change to an understanding of the processes and mechanisms whereby such change is theoretically understandable and exportable for the elimination of preventable retardation.

## References

Box, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems, II: Effects of inequality of variance and of correlation between errors in the two-way classification. <u>Annals of</u> <u>Mathematical Statistics</u>, 1954, <u>25</u>, 484-498.

- Cronbach, L. J. The two disciplines of scientific psychology. <u>American Psychologist</u>, 1957, <u>12</u>, 671-684.
- Heisenberg, W. <u>Physics</u> and <u>philosophy</u>: <u>the</u> <u>revolution</u> in <u>modern</u> <u>science</u>. New York, 1962.
- McCall, R. B. and Appelbaum, M. Bias in the analysis of repeated measures designs: some alternative approaches. <u>Child Development</u>, 1973, <u>44</u>, 401-415.
- Schaie, K. W. A general model for the study of developmental problems. <u>Psychological</u> <u>Bulletin</u>, 1965, <u>64</u>, 92-107.
- Winer, B. J. <u>Statistical principles in experi-</u> <u>mental design</u>. New York: McGraw Hill, 1971.