MONOGRAPHS OF THE
SOCIETY FOR RESEARCH IN
CHILD DEVELOPMENT

SERIAL NO. 150, VOL. 38, NO. 3
JUNE 1973

DEVELOPMENTAL CHANGES IN MENTAL PERFORMANCE

ROBERT B. McCALL
FELS RESEARCH INSTITUTE

MARK I. APPELBAUM
UNIVERSITY OF NORTH CAROLINA

PAMELA S. HOGARTY
FELS RESEARCH INSTITUTE

MONOGRAPHS OF THE SOCIETY FOR RESEARCH IN CHILD
DEVELOPMENT, SERIAL NO. 150, VOL. 38, NO. 3
JUNE 1973

CONTENTS

ABSTRACT

McCall, Robert B.; Appelbaum, Mark I.; and Hogarty, Pamela S. Developmental Changes in Mental Performance. *Monographs of the Society for Research in Child Development*, 1973, **38** (3, Serial No. 150).

A distinction was made between the general level of a person's IQ (i.e., the average of his scores across age) and his profile contour (i.e., the developmental pattern of inflections in performance). Most research has focused on the general level of IQ and its correlates; this project systematically investigated patterns of IQ change with data from the Fels Longitudinal Study. Normal middle-class children changed an average of 28.5 IQ points between 2½ and 17 years of age, and one in seven displayed shifts of more than 40 points. While siblings were more similar in their general level of IQ than unrelated children matched or unmatched for social class, the pattern of developmental changes in childhood IQs or infant DQs was not more similar among siblings than among unrelated children. Twins and triplets scored lower in general level and were less variable in IQ over age than singletons, thus potentially qualifying attempts to generalize IQ change results based upon multiple-birth samples to singletons. Parent-child IQ correlations were higher when the parent was assessed as an adult than if both parent and child were tested at the same chronological ages. A sample of 80 subjects who had relatively complete IQ data (maximum of 17 tests) between 2½ and 17 years were clustered into five groups which represented different patterns of IQ change over age. These profiles were relatively simple linear or quadratic trends and not random fluctuations about a constant value. Major inflection points occurred at 6 and 10 years of age. It was not obvious that these patterns were simple products of repeated testing or the changing nature of the IQ test. The children in the five IQ profile clusters had parents who were different in the extent to which they attempted to accelerate their child's development and the severity of punishments they administered. These parental correlates appeared to hold up even when parental education and IQ as well as the general level of the child's IQ were statistically controlled. Given the assumptions prompted by these data, the changing nature of environmental circumstances across the childhood years for a given individual may be as potent in changing IQ as the differences between family environments.

I. THE LONGITUDINAL STUDY OF DEVELOPMENTAL CHANGES IN IQ

Several reviews of the literature on the development of mental performance are available (Bayley 1970; Bloom 1964; Pinneau 1961). However, most of the research surveyed involves correlations between mental test performance at several different ages and correlations between personality and social variables with IQ at one age versus another. Such research strategies are valuable and have yielded important results, but they are "developmental" only in the sense that data are available at several ages on the same subjects and correlational relationships are derived within and across age periods. Because the correlation coefficient is independent of the means of the distributions, it is possible to obtain high correlational stability of IQ across age but marked fluctuations in the group mean. In short, cross-age correlations reveal little about the nature of absolute changes in level of performance, and they provide only partial information about individual differences in developmental pattern. This project focuses on developmental patterns of *change* in the level of mental performance as assessed by traditional intelligence tests.

IQ CHANGE OVER AGE

The Validity of Intelligence Tests

The basic data to be reviewed and analyzed in this *Monograph* consist primarily of IQ scores. These scores have often been interpreted as reflecting a basic, constant, unitary, general mental aptitude—Spearman's *g*. Increasingly, this orientation is being challenged. Many psychologists have pointed out that IQ scores do change, the IQ test reflects only one or two of many possible mental abilities, and the IQ test is not a fair indicator of the potential performance of certain subgroups in the American population (e.g., ghetto blacks).

These criticisms, while still enmeshed in controversy, represent important qualifications on the meaning and interpretation of IQ tests. It is undeniable that IQ scores have been misinterpreted and misused. However, it is probably also the case that the critics who argue that IQ scores have little or no validity and are therefore nearly worthless have also overinterpreted the data they cite for their position.

> Certainly it is true that mental tests have fallen into disrepute in many quarters. But with respect to predicting academic performance, this fate has come about not because the tests have failed to predict, but because they predict too well. It would be difficult to find a better example of a self-fulfilling prophecy than in the use of intelligence tests in the United States. . . . The insidious aspect of the situation is that, once typed by his test performance, it is extremely difficult for a child to overcome the resulting expectations of the educational system (and perhaps his parents) . . . . But however unjust this situation may seem, the association between test behavior and school success cannot be denied, with all it may imply about the place the child will assume in the social order. [Rees & Palmer 1970, p. 2]

Despite the current furor over interpretation and use of the IQ test, the IQ score is still probably the best single predictor of contemporary occupational and scholastic success for middle-class and minority-group children alike (e.g., Kennedy 1969; Kennedy, Van de Riet, & White 1963). Part of its cultural bias is shared with the criteria of scholastic and occupational success.

The present *Monograph* makes no global assumptions concerning the IQ score, except the belief that it reflects an important aspect of mental competence and behavior and that it correlates with and predicts academic and occupational success in contemporary American society. Beyond this, the observations proposed herein are intended to *inquire* about (rather than assume) the stability and proper interpretation of these mental assessments.

Stability and Change in IQ

All major longitudinal studies have reported the correlations between IQ assessed at two or more ages on the same sample of children. These data have been remarkably consistent in showing high correlational stability of IQ from middle childhood. Developmentally, prediction of IQ at 18 years from assessments prior to 3–4 years is not terribly impressive (McCall, Hogarty, & Hurlburt 1972). However, the correlations with 18-year IQ rise quickly between 3 and 6 years, and after age 6 the correlations are .80 and above between IQ in childhood and IQ at age 18 (Bayley 1949). For example, the correlation is approximately .96 between IQ assessed at 15 and 18 years (Bayley 1949).

These data have been interpreted as suggesting that IQ performance is very stable after age 6. Hence, until recently, IQ was considered to be largely unchanging throughout life. However, these correlational data do not merit such an interpretation.

Group changes in IQ.—Despite correlational stability, nearly every study which administered IQ tests to the same subjects at different ages has revealed some change in IQ for the entire group of subjects. As early as 1922, Baldwin and Stecher reported that the mean IQ increased between 5 and 16 years of age for their sample of 143 children. Since that time, the Harvard Growth Study (Dearborn & Rothney 1941), the Berkeley Guidance Study (Honzik, Macfarlane, & Allen 1948), the Fels Study (Sontag, Baker, & Nelson 1958), the University of Colorado Child Research Council Study (Hilden 1949), the Chicago Study (Freeman & Flory 1937), the Brush Foundation Study (Ebert & Simmons 1943), the Berkeley Growth Study (Bayley 1940, 1949), and the London Longitudinal Study (Moore 1967) have all found nontrivial group and/or individual subject changes in IQ. In nearly all of these projects, there was a general rise in IQ over age for the group as a whole, although there was often a small group of subjects who declined in IQ. By contrast, a longitudinal study of black children from predominantly disadvantaged environments has shown a progressive decline in IQ between the ages of 5 and 10 years (Roberts, Crump, Dickerson, & Horton 1965). Similarly, children from rural mountain villages or other isolated or "disadvantaged" environments have also shown (cross-sectional) declines in IQ performance over age (Asher 1935; Sherman & Key 1932; Skeels & Fillmore 1937).

There has been some suggestion that males are more likely than females to show increasing patterns of IQ over the childhood years (Bayley 1968b; Bayley & Oden 1955; Ebert & Simmons 1943; Moore 1967; Sontag et al. 1958), and in the study of poor black children (Roberts et al. 1965) boys were less likely to show the general group decline than were girls.

The IQ also changes in adulthood. Early studies (Jones & Conrad 1933; Miles & Miles 1932; Wechsler 1944) indicated that the highest IQ scores were obtained when individuals were in their early 20s, decreasing thereafter. However, these results are based on cross-sectional data in which the educational opportunities for different cohorts might have favored the younger groups. In contrast, data from longitudinal studies of the same children (Bayley 1955, 1957, 1966; Bayley & Oden 1955; Freeman & Flory 1937; Owens 1953) have indicated that IQ may increase throughout life, and Birren (1968) reports that men who do not suffer from deteriorated health may show increases in IQ quite late in life.

Changes in individual subjects.—Much of the evidence that IQ can change among normal children comes from case studies reported by the longitudinal projects cited above. For example, Sontag et al. (1958) presented individual graphs of 140 children from the Fels study who had been assessed 13 times between 3 and 12 years. The most dramatic increase was from an IQ of 118 at age 3 to 176 at age 11 (these are approximate values based upon averages of three assessments at adjacent ages). Probably the most impressive shift in IQ in the literature is for a boy who went from

78 at age 3 to 151 at 8 years, a shift of 73 IQ points (Moore 1967). This represents a rise from approximately the 8th to the 99.9th percentile.

Both stability and change?—The fact that the same data are used to demonstrate correlational stability as well as group and individual changes in IQ over age may appear anomalous. However, as stated above, the correlation coefficient is independent of the means of the distributions entering into its computation, and thus it is possible to obtain high correlational stability but marked changes in the level of IQ over age.

But, how is it possible to have high correlational stability and also a nontrivial number of individuals who display extreme IQ shifts that do not follow the group trend? This phenomenon may occur if the individual's shift is small in comparison with the total range of scores within each age. In these longitudinal studies the range of IQ scores may be 75–190 IQ points, varying with age. An individual shift of 30 or 40 points spread over 10 years is a meaningful change of approximately 2 standard deviations from the standpoint of the individual but a small inflection relative to the range of scores for the entire group. Thus, it is possible to have high correlational stability and simultaneously find an absolute shift in IQ for the entire group and many large individual changes from age to age.

FACTORS RELATED TO IQ CHANGE

Effects of Repeated Testing

Although the longitudinal method is the only approach that can answer questions about developmental changes in IQ, this procedure has the inherent problem that trends over age may occur as a function of the repeated testing regimen rather than any pervasive change in behavior perpetrated by nontesting factors. That is, subjects may become testwise. This possibility is often invoked to explain longitudinal IQ changes.

Two studies have attempted to enroll new subjects in the middle of their testing program in order to determine if children who have been exposed to repeated testings score higher than do those who are not so experienced. Using the Binet and WISC, Moriarty (1966) introduced new subjects at approximately school age into a testing program begun during infancy and found no difference between the new and the previously tested subjects. Freeman and Flory (1937) used a battery of vocabulary, analogies, completion, and opposite tests (VACO) beginning at 8 years of age and assessed new subjects periodically through the testing program. Again, no differences were observed between experienced and nonexperienced subjects.

If a repeated testing effect were the principle determinant of change in IQ performance, then one would expect to find simple increasing performance with repeated testing or declining increases as the cumulative effects of such experience stabilized. While it is true that many longitudinal studies show progressive increases for their sample as a whole, disadvantaged children tend to decrease (see above). Moreover, examination of the graphs for individual subjects presented by Sontag et al. (1958) indicates a wide variety of IQ developmental profiles. Contrary to these expectations, many children displayed relatively no change during the preschool period with a prominent point of inflection sometime between 5 and 7 years of age. Moreover, some subjects declined before age 6 while others showed large declines during the school years after a preschool period of relative stability. These observations, characteristic of the majority of the sample, correspond more with the developmental dynamics discussed by White (1965) than with predictions based upon repeated testing.

Thus, there is very little evidence in support of a simple repeated testing effect as a major influence in these studies.

General Level and Profile Contour of IQ

It is helpful to distinguish between a subject's general level of IQ and his profile contour. *General level* is the subject's average IQ over all assessments. It is the single value that estimates his approximate elevation on the IQ scale. *Profile contour* or *IQ pattern* refers to the sequence of inflections in the graph of a subject's IQ over age. These two concepts are potentially independent. For example, two subjects could show identically increasing developmental profiles but at different general levels, or two subjects could have opposite profiles at the same general level.

Several investigators have been concerned with the question of whether children who score at a high general level are more likely to show one pattern of IQ change than children who score at another IQ level. The most consistent result is that children with high general levels of IQ tend to change more in IQ over age regardless of the direction of that change than children with somewhat lower-IQ levels (Ebert & Simmons 1943; Hilden 1949; Hirsch 1930; Sontag et al. 1958; Terman & Merrill 1937). There are several possible explanations for this phenomenon. McNemar (1942) and Terman and Merrill (1937) have suggested that higher within-individual variability for high-IQ subjects is inherent in the mental age/chronological age formula for IQ. Moreover, Pinneau (1961) has pointed out that the more advanced items on the test are each worth more months in mental age than items scaled for the younger ages. Therefore, a correct or incorrect response on any single item at the higher levels has a greater impact on the total IQ than a response on the lower-level items. Therefore, guessing and chance responding would have a greater impact on IQ at relatively higher levels than lower ones.

Although high-IQ children display greater variability in performance over age, do they possess different profile contours from children scoring at other levels? Cattell (1931) found that between the ages of 3 months and 6 years low-IQ children tended to decline, average-IQ children showed irregular increases, and high-IQ subjects steadily increased over age. Baldwin and Stecher (1922) observed that the general increase over age occurred equally among low- and high-IQ groups, while Honzik and Macfarlane (1970) reported that gains were greatest for the average-IQ group between 18 and 40 years.

Therefore, although high-IQ children are more apt to show larger fluctuations in IQ over age, it is not clear that their profile contour will be systematically different from the profile contours of children scoring in a lower-IQ range. Because of the possible independence of IQ level and developmental profile, literature on correlates of IQ level and of profile contour will be treated separately below.

Parental Education

The IQ level.—Several investigators have inquired about the age at which the educational level of the parents correlates with the general IQ performance of the child (see McCall et al. 1972). The results from these several studies are quite consistent in suggesting that parental education begins to relate to the child's IQ performance between 1½ and 3½ years for girls but not until 2½ to 6 years for boys (Bayley & Jones 1937; Goodenough 1927; Hindley 1965; Honzik 1963; Kagan & Moss 1959; Moore 1968; Rees & Palmer 1970). Although the precise age at which the correlational relationship is first statistically significant varies among studies (i.e., presumably as a function of the nature of the sample, sample size, specific test instruments, testing regimen, etc.), the sex difference favoring an earlier age for girls is quite consistent.

Some investigators have found higher relationships with mother's education than with father's, especially during the 1½–3-year period (Bayley & Jones 1937; Kagan & Moss 1959). Others do not find this parental sex difference (Goodenough 1927; Honzik 1963; Werner, Honzik, & Smith 1968), although the parental correlations may be higher for boys than girls (Werner et al. 1968).

Interestingly, the relationship between the IQ of adopted children and the education of their biological parents who did not rear those children was strikingly similar to that for natural mothers rearing their own children. Moreover, there was nearly no correlation between the education of the rearing parent and the IQ of their adopted child (Bayley 1970; Honzik 1957; Skodak & Skeels 1949). However, the absolute IQ level of the adopted children as a group was reported to be closer to the general level of their rearing parents than to the level of their biological parents (Bayley 1966; Honzik 1957; Skodak & Skeels 1949). This seems to imply that, while children retained approximately the same ranking within their group as the relative ordering of their biological parents, as a group the children shifted more than 20 points toward the level of their adoptive parents (for a statistical discussion, see McCall [1970c], pp. 205–207).

The IQ profile.—The data on parental education and change in child's IQ is considerably more meager and inconsistent. Ebert and Simmons (1943) found a slight tendency for children who increased in IQ to come from better educated families and children who decreased from relatively less educated ones. Similar results are reported by Härnqvist (1968) for a sample of Swedish boys assessed at 13 and 18 years. Rees and Palmer (1970) noted that father's education related better than mother's to increases in IQ between 6 and 12 years and that this effect was stronger for boys than girls. Finally, Roberts et al. (1965) found some rather complex interactions between socioeconomic class and sex of child with the amount of change in IQ between 5 and 10 years for poor black children.

Thus, while the data are consistent in suggesting a fairly strong correlation (e.g., .50) between parental education and level of child IQ at any single age after 6 years (earlier for girls), the data on the relationship between parental education and developmental changes in IQ is less clear.

Siblings

The IQ profile.—While Rees and Palmer (1970) and Sontag et al. (1958) found little difference in IQ change as a function of the presence and sex of younger siblings, Hindley (1961), in England, observed that boys with younger siblings within 5 years of their age showed more gradual rises in IQ than girls with younger siblings. However, Rees and Palmer (1970) report that, if a boy had an older sibling, the greater the age span between them, the more likely the younger boy was to perform at a higher level and show increases in IQ after socioeconomic factors were controlled. Further, children showing increases in IQ tended to tell TAT themes containing more favorable dispositions toward male figures and unfavorable dispositions toward female figures. Rees and Palmer (1970) suggest that increases in IQ have been associated with masculinity and male roles in our society and that the availability of male models for children might be a mediating circumstance. Such an interpretation integrates their sibling data as well as the tendency for boys to show greater rises in IQ over age than girls.

There is some suggestion that larger families have children with lower levels of IQs (Hirsch 1930), although Reed and Reed (1965) found an IQ difference only when families exceeded five children. The negative correlation of approximately −.30 between family size and the IQs of the children in it led to "Cattell's paradox" in which the level of IQ in the population should, but does not, decline over generations. However, when childless individuals are included in the analysis, the lower-IQ groups produce the fewest and the higher-IQ couples produce the most children (Bajema 1968; Higgins, Reed, & Reed 1962).

Sontag et al. (1958) did not find any relationship between IQ change and family size.

Child Personality Correlates of IQ Change

The IQ level.—Bayley and Schaefer (1964) report the child personality correlates of IQ level at various ages for the Berkeley Growth Study. At the risk of glossing over important details, a major underlying thread in these observations appears to be a dimension of activity and frolicsomeness in infancy and social extraversion and facility in childhood and adolescence. However, there seems to be a developmental reversal in the relationships between these personality dimensions and test performance for boys but not for girls. That is, while social facility and extraversion are related to high test scores for both sexes during childhood and adolescence, male infants who are rated as happy, positively responding, and calm prior to 15 months have high infancy scores but relatively lower IQs by age 4. This shift is not as pronounced for girls. Similar relationships have been observed in the Fels data (McCall 1971; McCall et al. 1972), and analogous findings have been reported by Bell, Weller, and Waldrop (1971) for physiological activity in the neonate and social and mental performance at 2½ and 7 years of age.

While this summary of the Berkeley Growth Study data is not exhaustive of the literature on child personality correlates of IQ (see, e.g., Bayley 1968a, 1968b, 1970; Kagan & Moss 1962), it is representative of the fact that there are personality correlates of IQ level during childhood and that the pattern of relationships may be different as a function of age and sex.

The IQ profile.—Using clinical assessments, McHugh (1943) studied the change in IQ in the first 3 months of exposure to kindergarten. She reports that shy children gained most, perhaps as a function of adapting to the social situation of kindergarten and of the test administration. Moriarty (1966) described her clinical impressions of the personalities of children showing different trends in IQ.

Broader attempts to show personality correlates of IQ change have come from the major longitudinal studies. For example, Sontag et al. (1958) rated prose reports taken from home visits, interviews, and the child's behavior in a nursery school and day camp setting between 3 and 12 years of age. These ratings of a variety of child behaviors and personality dimensions were then examined for differences between children showing marked increases or decreases in IQ between 3 and 12 years of age. For the preschool period, independence and competitiveness characterized the IQ gainers of both sexes. During the elementary school years IQ increasers were described as independent, self-initiating, problem solving, and scholastically competitive and independent. High anxiety for girls and high parental emphasis on achievement for boys were sex specific in their relationship to IQ gain. Preschool ratings of aggressiveness, self-initiative, and competitiveness predicted elementary school gain in IQ.

Honzik and Macfarlane (1970) suggested that changes in IQ after

18 years are also associated with personality traits. People who gained in IQ between 18 and 40 years of age "maintained distance from other people; are likely to turn inward, not outward towards people; and there is some evidence for a lack of nurturance in relation to others." Schoenfeldt (1972) has also found relations between social-personality type and change in performance on the Army Alpha test throughout adulthood.

Despite the differences in procedures and ages between these studies, there appear to be child-personality relationships with IQ level and IQ change; these tend to involve variables that relate to the dimension of extraversion-introversion; and the developmental pattern may be different for boys and girls, at least within infancy and very early childhood.

Parent Behavior and Home Climate

The IQ level.—There have been several attempts to relate assessments of parental behavior and home climate to IQ at different ages (Bayley 1970; Honzik 1967a, 1967b; Kagan & Moss 1962; Moore 1968). For example, Moore (1968) has shown positive relationships between general IQ, vocabulary, comprehension, and reading ability on the one hand with four measures of the home climate on the other: (1) toys, books, and experiences; (2) parental example and encouragement; (3) emotional atmosphere; and (4) child adjustment. These home variables were rated when the child was 2½ years of age and related to IQ at subsequent ages to 8 years. For example, when social class was partialed out, the correlations between toys, books, and experiences with 7–8-year mental test variables were quite high: .70 with 8-year vocabulary for males and 7-year reading for females.

Bayley and Schaefer (1964) found that IQ level for girls was more highly related to the education of their parents; IQ for boys was more consistently related to parental personal-social-emotional characteristics. Bayley (1970) has summarized these results as follows:

> In the first year the boys' scores are seen to be negatively correlated with equalitarian, affectionate maternal behaviors but positively correlated with maternal hostile, rejecting behaviors. For the girls these correlations are reversed. For both sexes controlling, achievement-demanding maternal behaviors are positively correlated with children's scores. As the children grow older, the pattern shifts. At the preschool ages, the boys' correlations become more like those for girls. At school ages, the boys' scores are seen to be clearly related to the early maternal love-hostility dimension, whereas the girls' scores are almost completely independent of these maternal variables. For the boys there is a correlational pattern which may be stated briefly. Early maternal love and acceptance goes with slow development at first but later high achievement in mental abilities, with the reverse pattern for boys with punitive rejecting mothers. For the girls, although early maternal love goes with high scores in infancy, its influence diminishes after three years and then drops out almost completely. [Pp. 1192–1193]

While these are not the only relations nor are they totally supported by other literature (e.g., Crandall & Battle 1970; Kagan & Moss 1962), they do indicate parental behavior does relate to IQ level during childhood.

The IQ profile.—Honzik et al. (1948) present case studies of IQ change in which the inflections in mental performance appear to parallel alterations in the emotional climate at home. Sontag et al. (1958) found no parent variable during preschool to correlate with marked increases or decreases in IQ, but during the school years children characterized by substantial increases in IQ had accelerating parents who used rational and democratic discipline (e.g., justification of policy and readiness of explanation).

Again, there appear to be parental behavior and home climate correlates of IQ performance, and these correlates apparently depend on the age and sex of the child. In addition, sparse data suggest parental correlates of IQ change as well, perhaps focusing on the degree of accelerational attempt and an emphasis on rational and democratic discipline during the childhood years.

SPECIFIC ABILITIES

The literature reviewed to this point has been primarily concerned with changes in IQ and its correlates. However, the IQ scores reflect some types of mental behaviors and not others, and the several aspects of mental performance that are represented in the IQ score might profitably be separated. These concerns are commensurate with the shift away from the conception of "intelligence" as pervasive and unchanging to a dynamic system of many abilities and processes which may be more or less related and more or less constant over age (e.g., see McCall et al. 1972). It is possible for changes and growth profiles to be quite different for one type of mental performance than for another. For example, Cattell (1963) and Horn (1968) have suggested a distinction between crystallized (accumulated and retained knowledge) and fluid intelligence (processes of discrimination and reasoning). They proposed that crystallized intelligence continues to grow throughout life while fluid intelligence may decrease after 18 years, and Horn and Cattell (1966) offer some cross-sectional data in support of this hypothesis. Another example of the importance of specific abilities is the fact that females are somewhat better at verbal skills and boys at spatial-perceptual tasks (Conrad, Jones, & Hsiao 1933; Heilman 1933; Herzberg & Lapkin 1954; Hobson 1947). Perhaps the growth curves for these aspects of mental performance are different for the two sexes.

Is there any evidence that different abilities do show contrasting patterns of growth? Using cross-sectional data, Thurstone (1955) suggested that, while the developmental curves for number ability, memory for paired associates, and verbal comprehension were quite similar to one another, the childhood developmental progressions for perceptual speed, spatial visualization, reasoning, and word fluency were more erratic and dissimilar to one another.

Most of the data on growth patterns for different aspects of mental performance are for adults. Schoenfeldt and Owens (1965) administered the Army Alpha test to the same individuals at 18, 49, and 60 years of age and found that verbal ability increased markedly between 18 and 49 but declined slightly by 60 years. In contrast, numerical performance showed a steady decrease over adulthood and reasoning evidenced no significant change at all.

The subtest scales of the Wechsler test have provided a convenient vehicle for studying growth patterns for these abilities. Honzik and Macfarlane (1970) found that, while there was a gain in general IQ between 18 and 40 years, there was a loss on digit symbol scores. Both men and women gained in performance IQ, but only women increased in verbal ability.

Bayley (1968a) has examined the Wechsler subtest scores at five ages between 16 and 36 years. The subtests were initially subjected to a Tryon cluster analysis and then grouped into dimensions. Dimension A was predominantly verbal and interpreted as crudely representative of Cattell's category of crystallized intelligence (e.g., information, vocabulary, picture completion, picture arrangement, similarities). These scores tended to show monotonic increases during this age period, especially for males. The other subtests (digit symbol, block design, digit span, object assembly, arithmetic, comprehension), which Bayley noted were roughly similar to those classified by Cattell as fluid, showed increases until the early 20s followed by decreases to age 36. Unfortunately, it should be pointed out that the reliability and differential validity of Wechsler's subscales are not firm (Guertin, Rabin, Frank, & Ladd 1962; Littell 1960). Nevertheless, there is a suggestion in these several studies that crystallized intelligence increases throughout middle age whereas fluid performances begin to drop in the mid-20s.

One possible implication of differences in developmental profile for various mental abilities is that children who are competent at manual skills will drop in IQ over age while those with a more verbal orientation will increase, because most IQ tests presumably increasingly emphasize verbal content with age (Anastasi & Foley 1949; Jones 1954; Sontag et al. 1958). While Baldwin (1948) presents some evidence that this may be the case, Baker, Sontag, and Nelson (1955) and Sontag et al. (1958) found that children characterized by marked increases versus decreases in IQ did not differ in their ability to pass one set of items versus another.

It seems that not all abilities show the same growth rate and developmental shifts. On the other hand, changes in general IQ scores over age are probably not artifactual compounds averaged over a variety of different abilities. For the Wechsler tests between 16 and 36 years, there appears to be a fluid versus crystallized distinction among the 11 subtests, but that is all. Thus, longitudinal performance on the Wechsler and Binet IQ tests probably represents one or two developmental dimensions of profile contour over age. However, there are likely to be many other mental abilities, not adequately represented on these instruments, that have quite contrasting developmental profiles.

IMPLICATIONS FOR THIS RESEARCH

This sampling of the literature on the development of mental performance has several implications for the data to be reported in this *Monograph*.

First, most of the work on the "development of mental performance" has involved within- and cross-age correlations between mental test scores. Relatively less research exists on changes in test performance over age and the correlates of those changes. Since there is no necessary relation between cross-age correlations and the nature of developmental profile contours, and since individual subjects can display marked changes even though cross-age correlations are high, the investigation of change in developmental profile of mental performance should be considered a separate and important issue.

Second, most of the studies that have investigated change in mental test performance have looked merely at children who increase or decrease between two assessment ages. However, the use of this strategy ignores possibly important inflections in IQ pattern that occur between the two end points. Moreover, the increase-decrease dichotomy is an a priori classification scheme imposed by the researcher. The predominant IQ profiles determined empirically might not be simple increasing or decreasing functions. Consequently, there is a need to examine patterns of IQ change over age by procedures which consider all of the available assessments and which do not artificially impose a priori classification schemes upon the data.

Third, although profiles for individual subjects have been plotted and many reports of individual cases of change in IQ are available, there is relatively little information on individual differences in the pattern of IQ over age. One might ask: What are some typical developmental profiles that characterize large subsamples of individuals within a research population? Are such typical patterns of IQ change relatively simple, consistent, and meaningful shifts in performance, or are they merely random variability about some constant value? Are siblings and other genetically related children more similar to one another in their pattern of profile contour than are unrelated children? What are the parental behavior correlates of such developmental profiles of IQ change? These issues will be addressed in this *Monograph*.

II. SIMILARITY IN IQ LEVEL AND PROFILE AMONG RELATED AND UNRELATED PAIRS

A recurring issue in psychology concerns the relative heritability of mental ability, especially IQ as measured by standardized tests. The available data suggest that the correlation between the IQs of pairs of individuals increases markedly with the degree of their genetic relationship (Erlenmeyer-Kimling & Jarvik 1963), and the broad-sense heritability for IQ is approximately 0.70–0.80 (Jensen 1969a). On the other hand, the IQs of severely deprived youngsters can be raised by certain programs of stimulation and compensatory education (Deutsch, Katz, & Jensen 1968; Gray & Miller 1967; Heber 1969; Hellmuth 1968; Hunt 1961, 1969).

Despite the attempts of geneticists (including Jensen 1969a) to insist that (*a*) high heritabilities simply reflect the low degree of impact produced by environments currently represented in that particular sample, and (*b*) heritabilities have nothing to say about the potential of environments not represented or the general malleability or mutability of the trait, these concepts have been slow to be internalized by many professionals and lay people alike. Moreover, few people realize how much variability in the phenotype is possible as a function of the environmental circumstances that are sampled even if the heritability of IQ were 0.80. For example, Jensen (1969b) clearly shows that on the basis of a heritability of 0.80 the range of phenotypes for a theoretical group of individuals all having a "genotypic IQ" of 100 would exceed 40 points, or 2½ standard deviations of randomly sampled scores varying as a function of both heredity and environment. Thus, high heritabilities do not obviate change, either by the environment sampled (unless the heritability is perfect) or by environments not sampled, now or in the future (given any heritability).

For example, phenylketonuria is a monogenetic trait whose deleterious effects on the phenotype can be overcome by the institution of a specialized diet at a very early age. Clearly, genotypes do not inevitably produce a given phenotype; rather, a phenotype can be changed given certain environmental conditions at certain points in development. Determining the circumstances and timing required to produce phenotypic changes are empirical issues. The practical problem of changing the phenotype for any given trait may be difficult (e.g., changing the sex of an individual) or easy (e.g., hair color, obesity). Where shall one look for the determinants of IQ change, and how readily and in what ways can individuals shift their relative mental performance?

It is important to study changes in IQ with development, since this information might be quite different than the impression given by heritabilities calculated on the basis of a single assessment at a given age. Moreover, intraindividual shifts in IQ are not directly influenced by the diversity of family environments that exist in the population in the sense that an individual's general family environmental situation remains relatively constant across his childhood. Thus, the environmental factors that govern an individual's developmental change in mental performance could be qualitatively different from the between-family differences that presumably shape the environmental variance of single-age assessments. For example, it could be that under most contemporary family environments one's genotype exerts a relatively powerful control over one's general level of IQ (e.g., what quarter of the distribution one is likely to fall barring extreme circumstances), while the variety of life experiences (e.g., a certain teacher whose stimulation matched a particular interest of the child at a specific age) might contribute a salient determinant of a cluster of life-style attributes which result in a steady 30-point IQ rise over 5 or 10 years with commensurate implications for educational and occupational success.

Thus, it is a primary purpose of this *Monograph* to distinguish between the general level of IQ (child's average score over years) and the developmental profile of his spurts and lags in IQ pattern over age. Most of the evidence on the heritability of IQ has involved correlations between genetically related individuals at a single or at separate ages. There is nearly no evidence available on the possible heritability of the developmental IQ profile. To the authors' knowledge, only two studies have been concerned with the heritability of developmental profile contour as opposed to the general level of IQ. The first (McCall 1970a) examined the IQ profiles of pairs of siblings in the Fels Longitudinal Study between 3 and 12 years of age. For every sibling pair, an unrelated pair of subjects was matched for sex, year of birth, and parental education. As expected on the basis of the usual sibling correlation of .50, siblings were dramatically more similar than matched unrelated children for general level of IQ ($p < .0001$), but when only profile contour was examined there was no difference in the degree of similarity between siblings and matched unrelated children ($t < 1$). Among siblings, the degree of similarity was statistically greater for general level of IQ than profile contour ($p = .03$).

The lack of profile similarity was not because subjects did not change in IQ. On the average, an individual subject shifted approximately 24 IQ points (highest minus lowest score), and the patterns of change were not random fluctuations about a constant value but relatively simple, meaningful progressions. Sibling similarity apparently did not increase when the profiles were compared by equating children for calendar year rather than chronological age nor was it greater when the analysis was confined to the preschool or school years.

A similar analysis was performed for parent-child pairs in which the IQ profile between 3 and 12 years of age for the parent was compared with the profile for the child between those same ages. In contrast to other data in which the parent is assessed during early adulthood, this was a comparison between parent and child when *both were assessed at the same chronological age* (e.g., parent at age 3, child at age 3, etc.). Although one expects a correlation of approximately .50 between the *adult* parent's and the child's IQ (Erlenmeyer-Kimling & Jarvik 1963), the parent-child pairs were no more similar to one another than matched unrelated control pairs for either general level or profile contour. When the simple correlations between parent and child IQ at their respective ages between 3 and 12 years were examined, the median same-age parent-child correlation was $r = .29$ (varying with age), which is somewhat lower than the figure of .56 reported for the same kind of comparison by Burt (1966). Collectively, these data suggest that the possible heritability of developmental profile in IQ is considerably less than for general level of IQ.

Since McCall's report, Wilson (1972a) has addressed the same question using the Bayley infant scores for monozygotic (MZ) and dizygotic (DZ) twins from the Louisville Twin Study. His results show some degree of heritability for both general level and for profile contour assessed at four ages during the first year and separately for three ages in the second year of life. However, there are several differences between the Wilson and McCall studies. For example, Wilson examined Bayley scores between the ages of 3 and 24 months whereas McCall analyzed Stanford-Binet and Wechsler-Bellevue IQs between 3 and 12 years of age. It is possible that performance on infant tests is governed more closely by genetic and maturational factors that decline in influence with development (or are less related to the skills examined on the Binet and Wechsler tests). Alternatively, questions of heritability may be more precisely addressed with MZ and DZ twins than with siblings and unrelated children. This chapter attempts to consider these issues further.

DEVELOPMENTAL PROFILES OF IQ AMONG SIBLINGS

The McCall (1970a) analyses described above have some limitations (Cliff, personal communication, 1970; Wilson, personal communication, 1972). First, in longitudinal studies, subjects often miss assessments for a variety of reasons, most of which are unrelated to the behavior in question.

Nevertheless, it is possible that subjects who have sufficient data for analysis are somehow different from those who missed too many testings to be included in the analysis. Fortunately, in this case, the mean IQs of subjects included in the analysis are comparable to the means for the entire Fels population.

Second, the data base includes 14 consecutive IQ determinations throughout childhood; one cannot realistically demand complete data from every subject, which raises the statistical issue of handling missing observations. McCall (1970a) used a moving mean technique in which a subject received a score equivalent to the mean of the available tests for each set of three adjacent assessments between 3 and 12 years of age. It has been argued that such a "smoothing" technique would obscure minor inflections in the developmental profile. Although it is difficult to see how this would bias the sibling versus the control pairs, it would indeed produce somewhat smoother and simpler IQ profiles than might actually be the case. Better procedures for handling missing data are now known to the author.

A third criticism focused on the control pairs which were matched for sex, year of birth, and parental education. From a genetic standpoint, the comparison between siblings and unrelated pairs matched for parental education biases the result against obtaining greater similarity for related individuals because of assortive mating. That is, individuals tend to marry people with similar IQs and therefore presumably self-select a certain communality of genes related to mental performance. Hence, a control group matched for parent education (thus imposing artificial assortive mating) would mask the genetic effect to a certain degree. It would be best to add another control group composed of unrelated children matched for sex and year of birth but not for matched midparent education.

Finally, there has been some disagreement over the proper method of analysis of profile similarity. A variety of techniques has been used, and some psychometricians have pointed out deficiencies in nearly all of them (e.g., Cronbach 1958). Thus, a discussion of such techniques would be helpful.

Wilson (1968, 1972a) has proposed an analysis of variance procedure to evaluate similarity within pairs. In the case of MZ and DZ twins who have several mental assessments over an age span, Wilson employed a mixed model two-factor univariate analysis of variance with repeated measures performed separately on the MZ and DZ samples. Twin Pairs was the random factor with $N = 2$ per level, and Tests was the repeated fixed factor. Since the Pairs main effect was the ratio of between-pair relative to within-pair variability, the intraclass correlation based upon this effect was interpreted as a measure of the extent of within-pair similarity in the general level of performance. Analogously, the Pairs × Tests interaction and its accompanying intraclass correlation were used to evaluate within-pair similarity in IQ profile over age.

This application of the analysis of variance raises at least two questions with regard to its appropriateness and interpretation (McCall 1972a, 1972b). The first concern is whether one should compare similarity within a pair to between-subject variability before making a MZ-DZ comparison, and the second issue involves violation of the assumptions of homogeneity of covariance.

In Wilson's approach, the MZ-DZ comparison is made by testing the intraclass correlations for MZ versus DZ twin pairs. These correlations are expressions of within-pair similarity relative to between-subject variability, and they were calculated separately for MZs and DZs. While this is an acceptable technique, it is usually used only when between-subject variability is different for the two groups, which was not the case in this situation. A more direct approach would be to take the ratio of the mean squares within pairs for DZs and MZs as an expression of the relative similarity within pairs for the two groups. When this is done for Wilson's (1972a) data, there is greater MZ similarity for general level in the first and second years ($p < .0001$, $p = .03$) and for profile contour during the first year ($p < .0001$) but not during the second year ($p = .25$ approximately). Such a result is conceptually reasonable in that one might expect maturational circumstances to operate in the first year and progressively become mixed with nongenetic factors as the organism grows more complex.

The second concern with Wilson's (1968, 1972a) approach is the fact that the result from a univariate repeated measures analysis of variance is not accurate to the extent that heterogeneity of covariance (or correlation) exists between the several tests assessed on the same subjects (Box 1954). Such heterogeneity almost always characterizes serial measurements of this sort since adjacent-age scores usually correlate more highly than developmentally distant ones. Unfortunately, departures from homogeneity of covariance can rather markedly inflate F values and intraclass correlations, and such violations cannot be treated as lightly as researchers do when faced with heterogeneity of within-cell variance (Box 1954; Davidson 1972; McCall & Appelbaum 1973). Moreover, with $N = 2$, the corrections that exist for this situation (Box 1954; Geisser & Greenhouse 1958) are of dubious value. The best alternative in such cases might be to employ a multivariate analysis of variance (e.g., Bock 1963; Bock & Vandenberg 1968; McCall & Appelbaum 1973). Wilson (1972b) has responded to these and other criticisms (Wachs 1972), and the interested reader is encouraged to evaluate these contrasting positions.

McCall's (1970a) analytic strategy employed the square root of the sum of squared distances between corresponding points of the profiles of pair members as a measure of their similarity (actually dissimilarity). If the raw test scores are used, this measure reflects differences in general level as well as profile contour. If each subject's scores are taken as a deviation about his own mean, the index reflects differences in profile contour apart from general level. This approach was offered by Cronbach and Gleser (1953) as a general model for studying profile similarity after they evaluated its advantages and disadvantages relative to other techniques (e.g., factor models, correlational approaches, Mahalanobis distance function, etc.).

However, this procedure has its liabilities. It is most suitable for the case in which the variances of the several measures are equal and those measures are uncorrelated (Overall 1964). In McCall's (1970a) study the variances were homogeneous, but the measures were differentially correlated. For longitudinal data, the cross-age correlation matrix approaches a simplex, and thus the first and last tests in the developmental battery likely would be weighted more heavily in the dissimilarity measure for profile.

Although Cronbach (1958) was generally dissatisfied with all analytic procedures available at the time, more recently Bock and Vandenberg (1968) have offered another alternative using multivariate analysis of variance in a study of similarities in twin profiles on the Differential Aptitude Tests. This procedure analyzes the difference score between twin members for each available variable with a multivariate analysis of variance which compares MZ and DZ twins on these several difference scores in a multivariate framework. The analysis assumes multivariate normality and weights the several test scores differentially to produce a maximum multivariate effect. The test statistic is then evaluated against a sampling distribution which takes into account the number of twin pairs, the number of measurements, and the fact that the discriminant function has been maximized. It has the advantage of providing an exact probability statement of MZ-DZ differences and makes no assumptions about the homogeneity of covariance.

In view of the disagreements regarding method of analysis, both the squared deviations technique and the multivariate tests are reported in this paper. The authors feel the assumptions of the univariate repeated measures analysis of variance technique are untenable, and hence it was not used.

The meaning of the term profile as defined by these two techniques should be clarified. Both statistical approaches are sensitive to phase differences in essentially the same geometric trend. For example, suppose one child scored 100, 105, 110, 115 over a period in which his sibling scored 115, 110, 105, 100 at the same ages. Although both showed a "linear trend," the two analyses would indicate a sizable *difference* in the linear trend of their profiles. Similarly, for two children who possessed the following scores: 100, 105, 110, 105, 100, 105, 110 versus 120, 115, 110, 115, 120, 115, 100, respectively. These two would be regarded as having a difference in profile with respect to their cubic trends (the mean difference would be ignored). Finally, suppose one child has scores of 100, 105, 110, 115, 120, etc., while his sibling also shows an upward pattern but one which is not monotonic: 100, 110, 105, 120, 115, etc. From the standpoint of the analyses, these two subjects would have different trends despite their "similar" increasing tendency. However, the latter subject, in addition to possessing a large contrast for the nth polynomial component, would also have a strong linear trend. Thus, although the subjects do have different trends and this would be reflected in the data, neither analysis would regard them as being as dissimilar in total "profile" as in the first example. Thus, both pattern and phase differences are considered "dissimilarity in profile contour" by both analyses—two children are similar in profile only when they have the same relative trend at the same ages.

Sample

The sample included those subjects enrolled in the Fels Longitudinal Study (for a more complete description of the Fels population, see Kagan & Moss 1962; and Sontag et al. 1958) who had missed no more than two consecutive tests of the 14 assessments made between 3 and 13 years of age and who had at least one sibling who met the same data requirements. Forty-eight males and 66 females qualified for analysis and produced 18 male-male pairs (13 families), 28 female-female pairs (19 families), and 54 male-female pairs (30 families). When more than three siblings from one family were involved in any single sex grouping, one pairing was randomly dropped from the analysis in order to minimize this lack of independence. These children were born between 1930 and 1957, they have above average intelligence (approximate Binet mean of 117, varying somewhat with age) but average variability at any given age (average standard deviation, s = 16.9), and the average parental education on Hollingshead's seven-point (7 is high) educational scale was 4.57 (s = 1.28).

Three groups of subject pairs were developed. The sibling pairs and the control group matched for sex, year of birth, and parental education were identical to those used in the McCall (1970a) paper described above. The matching was done in the following manner. If A and B constituted a sibling pair, then the corresponding unrelated control pair, A-C, was determined such that A was arbitrarily selected from the two choices of A and B, and C was unrelated to either A or B but matched with B with respect to sex, year of birth, and average parental education level on the basis of the educational scale of the Hollingshead two-factor index of social position. All subjects in the control groups came from among those subjects in the sibling group. A third control group was formed in the same way, but the C member was matched for birth year and sex but not parental education. The differences in birth year and parental education between pair members for these three groups are presented at the top of table 1.

The IQ assessments were made every 6 months from 3 to 6 years of age and then every year through age 13. Unfortunately, the particular test used at any given age differed over the years of the Fels project. Thus, after Pinneau (1961) corrections were performed on the Binet test results, scores were standardized (M = 0, s = 1) for each test at each age on

TABLE 1

MEAN BIRTH YEAR AND PARENTAL EDUCATION DIFFERENCES BETWEEN MEMBERS OF PAIRS

	BIRTH YEAR		PARENTAL EDUCATION	
	M	*s*	*M*	*s*
Children ($N = 100$):				
Siblings	4.48	3.68	0	0
Matched	4.15	3.63	0.59	0.52
Unmatched	4.41	3.69	1.85	0.82
Infants ($N = 142$):				
Siblings	3.85	2.62	0	0
Matched	3.87	2.85	0.38	0.45
Unmatched	3.77	2.62	1.80	0.63

the basis of the mean and standard deviation for the entire Fels population (average N per age = 151). Table 2 presents the frequency for each sex of each test at each age. The number of missing data points is indicated in the last line for each sex. In the entire data base, 11% of the assessments were missing, and 85% of the available test scores were from the Binet.

Missing data points were estimated by a LaGrangian linear interpolation-extrapolation procedure which leaves existing data points intact and provides missing points in accordance with the existing data trend for the individual subject. The general rule was that interpolation is performed when possible, but if necessary extrapolation is executed until interpolation is possible. When one data point is missing between two extant points, the missing point is filled with the number halfway between the two existing points. If two consecutive points are missing between existing data, the first filled point represents an extrapolation of the preceding two data points $[X'_3 = X_2 + (X_2 - X_1)]$, and the remaining missing point is then halfway between that filled point and the next real data point. If a missing data point is the first point in the sequence, it is filled with the first existing data point. This procedure makes a reasonable compromise between not disturbing existing data and filling in data points in a manner that preserves the existing preceding trend within a subject.

Results

Similarity (actually dissimilarity) in general level was examined by analyzing the absolute value of the difference between the mean IQs over age of pair members. Profile similarity was assessed first by calculating the square root of the sum of squared deviations between corresponding points of the profile of the pair members after each subject's mean had been subtracted from each of his scores. These similarity scores for the three groups

TABLE 2

DISTRIBUTION OF SPECIFIC TESTS FOR SUBJECTS IN CHILDHOOD SIBLING COMPARISONS

	AGE													
TEST	3	3½	4	4½	5	5½	6	7	8	9	10	11	12	13
Males ($N = 48$):														
Merrill-Palmer	1	10	...	6	...	2	...	...	...	...	...	...	...	...
Binet, SR	10	...	10	...	4	...	2	1	...	...	...	...	...	...
Binet, L, M, LM	35	38	35	37	39	37	45	35	38	34	47	31	28	...
WISC	...	...	...	...	4	...	...	11	...	5	...	14	...	...
W-B	...	...	...	...	...	...	...	...	...	...	...	...	...	38
Missing	2	...	3	5	1	9	1	1	10	9	1	3	20	10
Females ($N = 66$):														
Merrill-Palmer	1	15	...	12	...	3	...	...	...	...	...	...	...	...
Binet, SR	14	...	15	...	12	...	8	3	...	...	...	...	...	1
Binet, L, M, LM	46	50	45	46	49	50	58	47	54	48	62	43	44	...
WISC	...	...	...	...	5	...	...	14	...	5	...	19	...	2
W-B	...	...	...	...	...	...	...	...	...	...	...	...	...	49
Missing	5	1	6	8	...	13	...	2	12	13	4	4	22	14
Total missing of 114	7	1	9	13	1	22	1	3	22	22	5	7	42	24

(siblings, matched, and unmatched unrelated controls) were subjected to an analysis of variance, and the sibling-unmatched contrast was tested a priori. The mean dissimilarity scores for these groups are presented in table 3.

TABLE 3

MEAN DISSIMILARITY INDEX FOR CHILDHOOD IQ FOR SIBLING, MATCHED, AND UNMATCHED PAIRS

Group ($N = 100$)	General Level M	General Level s	Profile Contour M	Profile Contour s
Siblings	0.51	0.44	2.32	0.62
Matched	0.75	0.51	2.53	0.91
Unmatched	0.75	0.56	2.44	0.61

The general analysis of variance on the dissimilarity scores for general level was significant, $F(2,297) = 7.48$, $p < .001$, implying that siblings were more similar in general level than unrelated children (table 3). Moreover, sibling pairs were more similar than matched, $F(1,297) = 10.97$, $p < .005$, or unmatched, $F(1,297) = 11.47$, $p < .001$, groups.

With respect to profile contour (table 3), the overall analysis of variance was nonsignificant, $F(2,297) = 2.09$, as was the a priori contrast between siblings and unmatched unrelated controls, $F(1,297) = 1.36$. Therefore, there were no differences between groups in within-pair similarity for profile contour as assessed by this method.

Profile similarity was also analyzed with a multivariate approach adapted from Bock and Vandenberg (1968). This involved calculating the orthogonal polynomial contrasts for each subject (McCall & Appelbaum 1973; Winer 1971, pp. 170–185), and then taking the algebraic difference between corresponding contrasts for subjects within a pair. The sampling distribution used to evaluate the probability of the observed multivariate F takes into account the number of measures in the analysis, and the critical value rises with the number of dependent variables. Thus, since the predominant trends in IQ change over age are relatively simple (see McCall 1970a; and Chap. IV), and since differences in higher-order polynomial contrasts may be reflections of measurement error and not contribute to group discrimination, the multivariate tests were made on the linear, quadratic, cubic, and quartic orthogonal polynomial contrast differences only.

Thus, the unit of analysis was a set of four-pair differences for subjects in the sibling, matched, and unmatched unrelated control groups. The analysis was a one-factor multivariate analysis of variance with the four orthogonal polynomial contrast differences for each pair of subjects as the dependent variables (see Bock 1963; Bock & Vandenberg 1968; McCall & Appelbaum 1973). The multivariate test of the Groups factor constitutes a test of whether the relative within-pair similarity in IQ trend over age was different for siblings, matched, unmatched relatedness groups. One multivariate analysis was performed on all three groups and another just on the sibling-unmatched comparison.

The multivariate F's for these tests were both less than 1, thus failing to indicate differences in profile similarity within pairs for the three relatedness groups or for the sibling-unmatched contrast.

It was also possible that siblings might be more similar than unrelated children in the variability of IQ performance over age regardless of its pattern or direction. That is, there might be heritability to "plasticity" in mental performance or to susceptibility to changing environmental circumstances regardless of the direction, form, or chronological timing of their influence on IQ. This hypothesis was tested by computing the standard deviation of the 14 IQ scores for each subject as a measure of IQ variability and correlating those standard deviations between sibling pair members. The resultant correlation was .08, indicating that siblings were not similar in the variability of IQ performance over age.

The data indicate that, while siblings are more similar than matched or unmatched unrelated controls in terms of general IQ level during childhood, they were not more similar in developmental profile contour, a result concordant with the previous report (McCall 1970a). There was no greater similarity among siblings for variability in IQ regardless of the nature of the change.

PROFILES OF INFANT DEVELOPMENTAL QUOTIENTS (DQ)

The above analyses indicate that the difference between the McCall (1970a) and Wilson (1972a) results are not purely a function of the particular method of handling missing data or assortive mating. However, these two studies differ in the ages of the subjects investigated, and there can be no pretense that IQ in an infant is qualitatively similar to what is reflected on a childhood IQ test. Consequently, an effort was made to examine sibling pairs, matched unrelated infants, and unmatched unrelated infants for relative similarities in general level and developmental profile for Gesell scores obtained at 6, 12, 18, and 24 months.

Subjects

All subjects in the Fels Longitudinal Study who had at least three of the Gesell assessments made at 6, 12, 18, and 24 months of age and who had at least one sibling who fulfilled these same data requirements were selected for analysis. This sample included 72 males and 74 females born between 1940 and 1959 with a mean parental education level of 5.26

(s = 0.94) on the Hollingshead scale. This sample produced 37 male-male (23 families), 36 female-female (24 families), and 69 cross-sexed (32 families) sibling pairs.

The sibling, matched unrelated, and unmatched unrelated groups were formed in a manner comparable to that described above for the childhood data. The birth year and parental education differences between pair members for the three groups are presented in the bottom half of table 1.

The data (Gesell DQs) were not standardized, and missing data were filled in by the linear interpolation-extrapolation procedure described above. Only 5% of the assessments were missing.

Results

The analyses were identical to those described above for the childhood data. The mean dissimilarity scores are presented in table 4.

TABLE 4

MEAN DISSIMILARITY INDEX FOR GESELL SCORES FOR SIBLING, MATCHED, AND UNMATCHED PAIRS

Group (N = 142)	General Level M	General Level s	Profile Contour M	Profile Contour s
Siblings	3.85	3.22	9.54	6.15
Matched	5.45	4.22	9.29	3.85
Unmatched	4.77	3.88	10.34	5.21

For general level, the overall analysis of variance indicated that the relatedness groups differed in the degree of within-pair similarity, $F(2,423) = 6.33$, $p = .002$, and that siblings were more similar than matched, $F(1,423) = 12.64$, $p < .001$, and unmatched, $F(1,423) = 4.16$, $p = .04$, unrelated control pairs.

When profile contour was considered, the overall test was not significant, $F(2,423) = 1.60$, nor was the a priori contrast between siblings and unmatched unrelated pairs, $F(1,423) = 1.72$. The multivariate analyses comparing the three groups and comparing siblings with unmatched controls yielded multivariate F's < 1. These same results were also found when the analyses were separately performed on the first-year (6- and 12-month Gesells) and second-year data (12-, 18-, and 24-month Gesells).

The correlation between the standard deviations for individual subjects within pairs was .00, indicating that siblings were not similar in the degree of variability in DQ over age regardless of the form or timing of performance shifts.

These results are identical to those found for the childhood data. There was greater similarity for siblings than matched or unmatched controls for general level but not for profile contour, and siblings were not more similar in their variability of IQ over age regardless of direction.

SIMILARITY IN IQ PATTERN OVER AGE FOR TWINS AND TRIPLETS

The data reported above suggest that, while siblings are more similar in general level of performance on standardized mental tests during infancy and childhood, they are not more similar than matched or unmatched unrelated controls with respect to profile contour of performance over age. These findings are consistent over different methods of handling missing data, checks on assortive mating, and two statistical techniques, yet they are at variance with Wilson's (1972a) report on MZ and DZ infant twins. While the preceding discussion of techniques for assessing profile similarity suggests that there may be statistical qualifications and biases that would explain these contradictory results, another factor may be that twins and singleton births represent sufficiently different populations that generalization between them is restricted in some ways.

We now turn to some very limited data from the Fels study on a few sets of multiple births. Unfortunately, the Fels files do not contain enough measures during infancy on these twins to be comparable to Wilson's analyses, and the Louisville twins are not yet old enough to provide childhood data on a sufficiently large scale. Thus, despite the limitations of the data, serial IQ tests on twins throughout the childhood period are rare, and these data can alert us to hypotheses and methodological issues that will be valuable in considering the issue of IQ change now and in the future.

Subjects and Zygosity

All twins and triplets enrolled in the Fels study were investigated. Unfortunately, (*a*) there were only eight sets of twins and four sets of triplets with any IQ information at all, (*b*) the timing of IQ assessments during childhood was not the same for all pairs, (*c*) different IQ tests were used at different ages for different pairs, and (*d*) the zygosity of some sets was open to question. Nevertheless, the data, such as they are, are unusual and capable of provoking hypotheses about IQ change and the twin situation. For example, there seemed to be differences between multiple and singleton births which are important and could be interpreted without being certain about zygosity within the multiple-birth sample.

The 28 subjects were born between 1928 and 1964 with a median birth year of 1939. Their parents had an average of 3.94 on Hollingshead's educational index ($s = 1.14$).

The patterns of IQ change for multiple births in the Fels sample are presented in figures 1–3. The axes for each set are identical across families so that the graphs can be compared directly. However, the specific IQ tests administered at different ages varied from set to set. A key is presented on

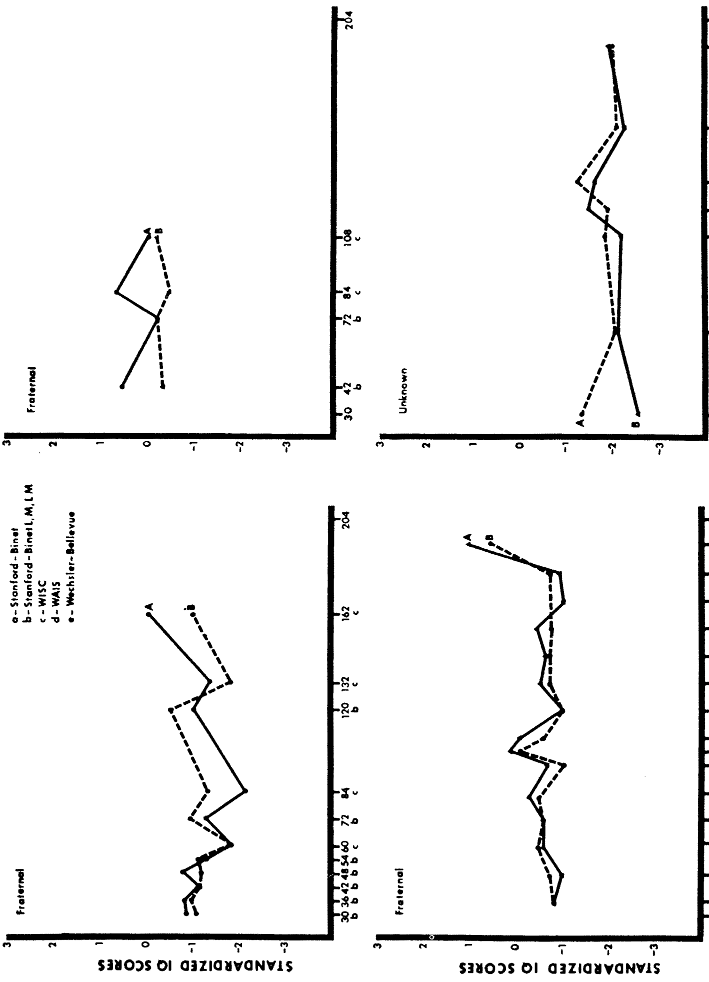


FIGURE 1.—Standardized IQ scores over age for three sets of fraternal twins and one set of unknown zygosity.

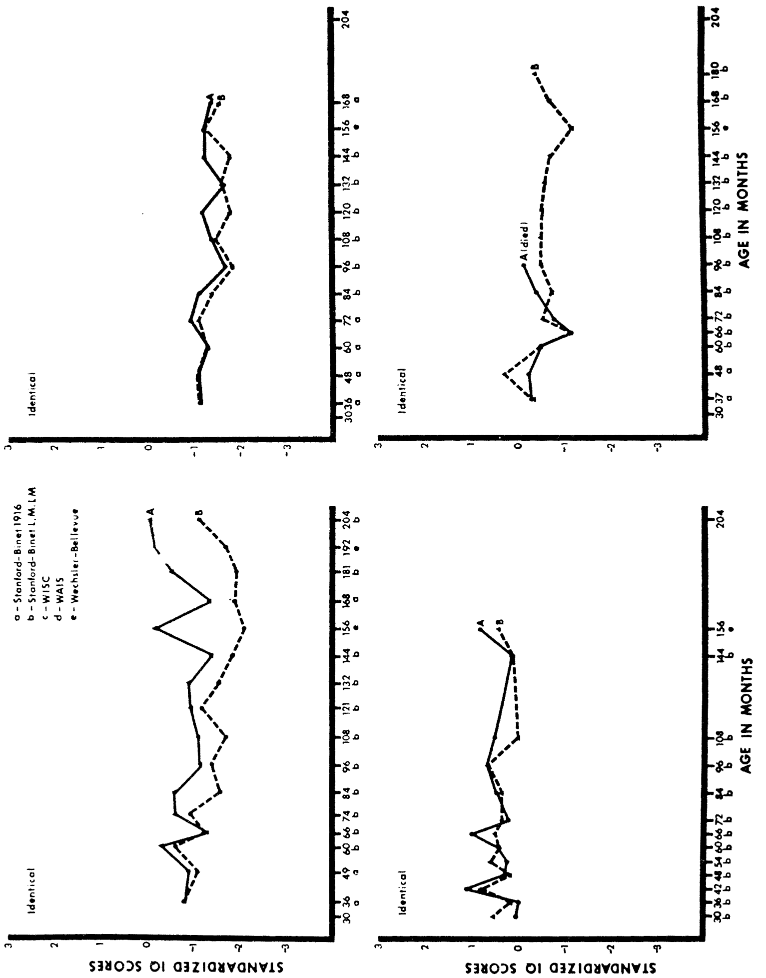


FIGURE 2.—Standardized IQ scores over age for four sets of identical twins

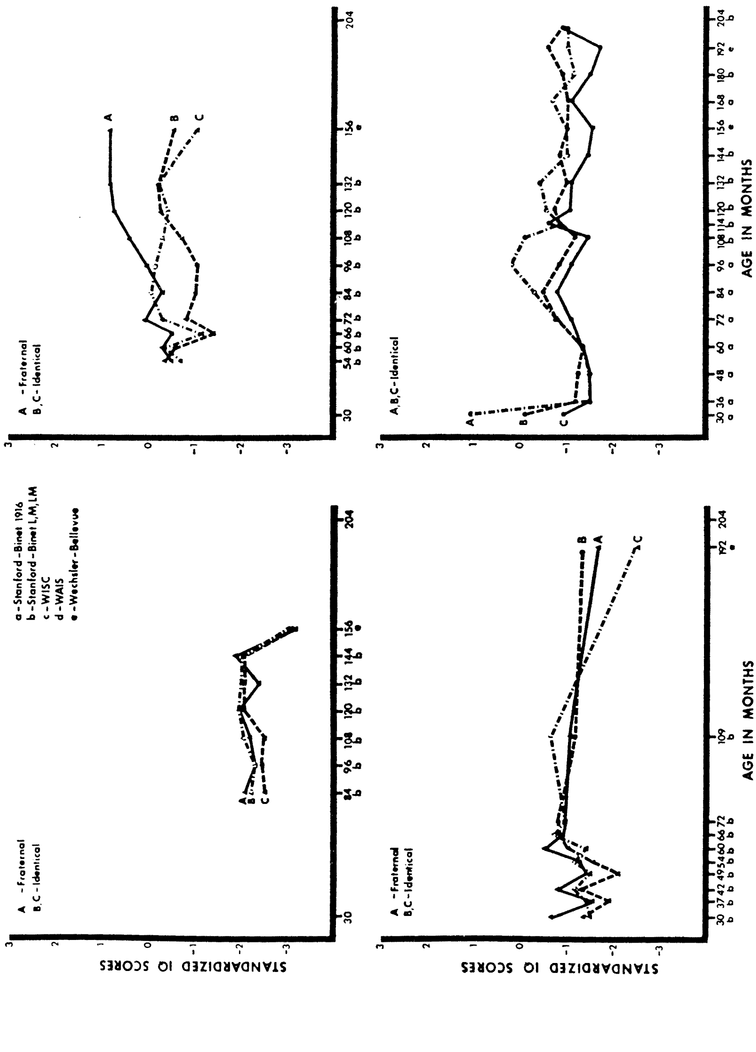


FIGURE 3.—Standardized IQ scores over age for four sets of triplets

each figure and the letters below the ages indicate which IQ test was used for that determination. The ordinate represents standardized IQ ($M = 0$, $s = 1$), where the standardization was based upon the mean and standard deviation for all subjects in the Fels population given that test at that age (average N per age = 151). Thus, zero on the ordinate reflects the Fels mean for each test at each age. Consequently, these data are roughly comparable to those presented above for childhood siblings in that any general trend in the Fels sample has been extracted in both cases and graphs represent relative departures from the performance of the Fels population as a whole.

Unfortunately, when these subjects were enrolled in the Fels study the determination of zygosity was not as precise as it is today. Moreover, some of these subjects can no longer be reached since they have graduated from the Fels program. Consequently, the determination of zygosity leaves something to be desired in several cases. In figure 1, the three sets of fraternal twins are composed of a male and a female, and the unknown set is a pair of males (not used in zygosity comparisons below). All but one set of twins in figure 2 were judged to be identical. Considered from left to right, top to bottom, the zygosity judgment for the first set was based totally on appearance. The second set was described at 11 years by a pediatrician as "undoubtedly monozygotic" and having "practically identical growth trends, anthropometric measurements, ossification, dentition, hair color, eye color, and two identical hereditary defects of red/green color blindness and missing permanent bicuspid teeth." The members of the third set were identical on 22 blood group factors (one ambiguous) and the fourth set on appearance and similarity for the blood group antigens A, B, M, and N. The first set of the four like-sexed sets of triplets in figure 3 were girls, one of whom was larger than the other two, and inspection of X-rays of bone joints at 7 years of age suggested two MZs and a fraternal twin. The zygosity of the second, third, and fourth sets of triplets is more certain, being based on an assessment of at least 11 blood group antigens. In short, of the 19 usable *pairs* of multiple births, the zygosity of 14 was reasonably certain (i.e., made on the basis of sex or 4–22 blood group factors), four were based on some, but inadequate, evidence (dentition, color blindness, skeletal similarity), and one was judged on appearance only.

Results

The square root of the mean squared difference between corresponding deviation IQ points was calculated for each pair in this sample of multiple births as an index of the degree of dissimilarity in profile contour. Because of the limited number of subjects and the diverse distribution of tests, the multivariate procedures could not be used. All pairs of subjects were analyzed, and no data points were filled.

When profile contour was examined and zygosity accepted as described above, the 10 identical pairs were not more similar in their profile contour over age than the nine fraternal pairs. In fact, the direction of the difference in mean dissimilarity score was opposite to the genetic prediction, $t(16) = -0.74$.

It was interesting to observe that the mean dissimilarity index for profile contour for both MZ (0.42) and DZ (0.37) pairs was lower than the average of 0.62 for singleton siblings. The DZ-sibling difference, $t(107) = 1.76$, $p < .05$ one-tailed (see McCall 1972a), is particularly instructive since the degree of genetic overlap among pair members is the same for these two groups. This difference presumably reflects the fact that environmental events have their effect at the same age for multiple births but at different ages for singleton siblings.

Unfortunately, the Fels childhood twin and triplet data must be viewed with considerable caution since the zygosity determinations were not as accurate as desirable in some cases. However, other data point to the same conclusion. If one takes Wilson's (1972a) DZ infant twins and compares the intraclass correlations for general level (.75, .79 for the first and second year, respectively) and profile contour (.52, .50) with those obtained by the same analysis on the Fels infant sibling data for general level (.24, .44) and profile contour (.09, .14), the siblings are not as similar as Wilson's twin pairs assessed by this method (see discussion of method above). Moreover, since DZ twins and siblings both share half their genes, the theoretical genetic average correlation is .50 (Jensen 1969a), and a figure higher than this must reflect nongenetic (presumably environmental) factors operating in concert for pair members. Thus, similarity in profile contour may be greater for twins than siblings (see below) because of greater nongenetic similarities within pairs.

Firmer conclusions can be drawn from the Fels data by comparing singletons with multiple births, ignoring zygosity. For example, the general level of IQ was lower for multiple births than for singletons in the Fels sample. When the mean standardized IQ for each subject was computed, multiple births were more than 1 standard deviation below singletons in IQ (0.14 vs. $-$ 0.91), $t(141) = 6.60$, $p < .0001$. Moreover, only six of the 30 multiple births had an average IQ above the Fels mean (approximately 117 on the Binet at age 10), and 14 had averages lower than 1 standard deviation below the Fels mean. Note that in terms of unstandardized IQ values, the multiple births averaged approximately 102, although the difference between singleton and multiple birth means is somewhat greater than usually found (Record, McKeown, & Edwards 1970).

Moreover, twins and triplets did not change very much in IQ from one age to the next. If a standard deviation is calculated for the scores possessed by each individual as a measure of that subject's variability in IQ over age, the average standard deviation was 0.41 for multiple births and 0.47 for singletons when no filled data points were included, $t(142) = 2.27$, $p < .03$. Regardless of zygosity, multiple births varied less in IQ performance from age to age than did singletons.

There may be several reasons why multiple births evidence less change in IQ over age than singletons, and one possibility concerns the nature of IQ tests. The Stanford-Binet, for example, assigns 1 month of mental age per item at the easy and middle difficulty level, but items regarded as difficult frequently have more than 1 month of mental age associated with them. Consequently, as indicated in the literature review, there is greater variability over age for subjects with a high level of IQ performance than for those with a lower-average score. Therefore, it is possible that multiple births have less variability than siblings merely because they score lower on the test. To assess this, the variability of the multiple births was compared with only those singletons who fell within the range of individual average standardized IQs for multiple births (-2.38 to $+0.46$). The multiple births again showed somewhat less variability than the singletons (0.41 vs. 0.44), but the difference was not at all significant ($p = .42$). Thus, it seems likely that multiple births show less variability in IQ over age than singletons, partially as a function of scoring lower on the tests.

DISCUSSION

Summary

The data presented previously (McCall 1970a) and in this chapter indicate that normal middle- and upper-middle-class children change an average of 24 points on standardized IQ tests between 3 and 12 years. These shifts were not random fluctuations about a developmentally constant value but relatively simple and consistent progressions over age (McCall 1970a; see Chap. IV). Moreover, while siblings possessed some resemblance in the general level of IQ performance relative to unrelated children matched or unmatched for parental education, there was no evidence that siblings exhibited greater similarity with respect to the pattern of IQ over age. These results held for IQs assessed between 3 and 13 years as well as for Gesell scores obtained during the first 2 years of life.

A small number of multiple-birth pairs also failed to display a difference in the similarity of IQ profile contour during childhood as a function of zygosity. The degree of profile similarity within pairs was substantial for multiple births during infancy and childhood. The greater profile similarity for DZ sets than for singleton siblings despite their presumed equivalence in genetic overlap argues for the operation of nongenetic factors in determining profile contour. Unfortunately, the zygosity information was less precise than desirable, and thus these results must be viewed only as provocative of hypotheses to be investigated further.

Relatively more confidence can be placed in comparisons between single and multiple births regardless of zygosity. Multiple births had lower IQs and less variability over age than singletons, the latter result possibly deriving from the lower scores characteristic of multiple births.

In contrast to these data, Wilson (1972a) has reported similar analyses on Bayley scores during the first year and separately during the second year of life, comparing MZ and DZ twins for similarity in general level and profile contour. The similarity for MZ twins was greater than for DZ twins for general level and for profile contour. The discrepancy in results between the twins and the present data may reside in the difference in infant tests, methods of analysis, and in the difference between twins and singleton children.

Tests and Analyses

While there is considerable item overlap in the Gesell and Bayley scales, a review of the literature indicates that the means and correlations of the two tests as well as their efficiency in predicting later IQ are far from identical (McCall et al. 1972). Moreover, while the items on the Gesell are elaborately described, the scoring is not as precise and mother reports are sometimes included. In contrast, the Bayley administrative and scoring instructions are more definitive and less subject to "clinical judgments." Since the Gesell protocols from Fels were obtained principally by one examiner, they are further open to criticisms of examiner bias, although it is difficult to predict whether this would result in more or less similarity in MZ versus DZ profiles.

The qualifications on the analysis of variance model used by Wilson (1972a) have been described above and Wilson's (1972b) rebuttal noted. Given these data one might argue that, while MZs are more similar than DZ twins in general level for the first and second year, they may be more similar for profile contour only during the first year. Moreover, if "broad-sense heritabilities" are calculated for Wilson's data (McCall 1972a), the estimates for general level in the first and second year are 0.30 and 0.20, respectively, and 0.50 and 0.30 for profile contour. Given these heritability figures, a good deal of variability within infant twin pairs is apparently determined by nongenetic circumstances. Moreover, these results are subject to the limitations imposed by violations of the assumption of homogeneity of covariance required by the repeated measures model. Data presented here for siblings and unrelated infants and children do not provide any evidence for greater profile similarity among siblings than unrelated children.

Implications of the Twin Situation

It may be that for psychological or psychometric reasons, generalizations cannot easily be made between singleton and twin populations. For example, evidence has been presented which suggests that not only do twins score lower in childhood than singletons but they may also be less variable in IQ over age. A similar situation may exist in infancy. An examination of Wilson's (1972a) infancy data indicates that, in comparison with Bayley's (1965, 1969) norms, the mean score for twins was lower at every age and significantly lower at three of the six ages during the first 2 years. Moreover, while the standard deviation for the 24-month assessment for twins was comparable to that found for singletons (Bayley 1969), the within-age variability among singletons was from 16% to 40% greater than among twins at the other ages. Since the mean scores for twins, though lower, followed the same pattern over age as singletons, it is possible that intraindividual variability was also less for twins than singletons.

The fact that twins apparently vary less in performance across age may suggest that twins do not show the range of developmental patterns that singletons do and that their profiles may be more restricted in the magnitude of change in relative performance over age. Thus, the two samples might not be comparable with respect to the extent and variety of IQ profile, and generalization across populations might be thereby limited.

Growing out of this discussion is the point that all the analyses presented here and by Wilson (1972a) regarding similarity in profile contour have assumed that similarity within pairs for general level and for profile contour of IQ over age are independent. In view of the past literature, sparse as it is (see Chap. I), this was a reasonable assumption. However, suppose, for example, children who scored between 120 and 150 tended to display one kind of pattern of IQ over age and those scoring between 90 and 120 were characterized by a different IQ pattern. Then, general level and profile would be related. Moreover, in such a case similarity in profile contour would be obtained to the extent that the general levels of the pair members were similar. Since it is well-known that MZ twins are more similar in general level than DZ twins, one might also expect them to be more similar in profile contour under these circumstances, not as a function of a genetic factor governing IQ change but because pattern over age might be related to general level. The "independent" evaluation of these two effects by the analysis of variance procedures would not eliminate such a "natural" relation between general level and profile type.

Is it the case that general level and IQ profile are related in the manner suggested above? Although the past literature reviewed in Chapter I is ambiguous, data presented in Chapter IV of this *Monograph* strongly suggest for the present sample that children who display different patterns of IQ change over age also differ in general level of IQ ($p < .001$). Thus, pairs of subjects may have profile contour similarity in crude proportion to the degree of their similarity in general level.

One reason similarity in general level might not be independent from similarity in profile contour concerns the relation between mean and variance on the Binet. Consider that the less variability in IQ over age within a subject, the more that subject's profile becomes a straight line. Note that individual variability becomes less on the Binet as general level decreases, and thus two low-scoring pair members are more likely to share a common pattern (i.e., a straight line) than higher-scoring pairs, and this would probably be more characteristic of multiple births who score lower, have less variability over age, and possess greater similarity in general level than singletons. Thus, one would expect a correlation between within-pair similarity in general level and within-pair similarity in profile contour, a fact that would question whether any similarity observed in IQ profile was not merely an artifact of similarity in general level. In the Fels data, if the absolute value of the difference between pair-member means over age (a measure of general level dissimilarity) is correlated with the square root of the mean sum of squared deviations (index for profile contour), the correlation is .56 ($df = 17$, $p < .02$) for the multiple births regardless of zygosity but only .14 ($df = 98$, $p = .16$) for siblings for the childhood years. Since MZ twins are more similar to one another in general level than DZ twins, within-pair concordance in profile for MZ versus DZ twins might be at least a partial function of similarity in general level.

Since psychometric factors associated with the Binet scoring system may operate in this relationship, similarity in general level may not be related to similarity in profile contour for infant test scores. Indeed, the corresponding correlation for the 142 infant-sibling pairs was .00. However, since low within-subject variability across age may also be involved, this possible relationship should be routinely investigated in infant twin samples as well.

The lower variability within individual twins could also have implications for heritability estimates based on a single assessment, depending upon the mechanism operating. Loehlin (personal communication, 1972) has viewed the following possibilities:

> If the twin situation acts as a buffer to environmental inputs in general, so that twins' phenotypes reflect their genotypes more directly than do non-twins', this could lead to an overestimate of heritability from twins; if the twin situation also acts as a counterforce against genetic variability (for example, by tending to hold members of a fraternal pair to an internally-developed pair norm), it could produce the opposite effect, or no net effect at all, depending upon the relative strengths of the forces involved.

However, one would need to postulate that whatever is operating in the twin situation makes a different impingement on MZ than DZ twins, otherwise most conventional correlational methods of assessing heritability at a single age would not be influenced very much by low intraindividual variability. Moreover, estimates of IQ heritability based upon nontwin data agree reasonably well with those derived from twin protocols, and thus estimates of heritability for general level of IQ rooted in single-assessment data from multiple births are probably not influenced by factors uniquely associated with the twin situation.

However, the situation remains ambiguous when determining the heritability of developmental profile apart from general level from twin samples. These results suggest that heredity-environment investigations of IQ profile contour over age must demonstrate that the similarity in general level and similarity in profile contour are either uncorrelated or equally correlated within each relatedness group. This is not guaranteed by the statistical techniques usually used to analyze such data. If the relatedness groups show differential similarity among pairs for general level, the variability in individual IQ over age is not too great, and/or there are differing degrees of association between the similarity in general level and in profile for relatedness groups, one may obtain spurious results. These potentially biasing circumstances are more likely to be present when studying multiple births and when current tests of childhood IQ are employed.

Conclusion

The analysis of profile similarity in mental test performance for related groups of paired individuals is a highly technical enterprise. A number of statistical techniques exist which have different features, qualifications, and limitations. The researcher must be careful to alert his reader to these circumstances and possible biases. The multivariate approach may be the best alternative to more traditional techniques. Moreover, similarity in developmental profile may not be independent of similarity in general level, especially for childhood IQ data and for twin samples. A check on this assumption is necessary to insure that profile differences between relatedness groups are not psychometric or statistical artifacts of similarity in general level.

The data on similarity in profile contour for mental test scores suggests that there may be some degree of profile similarity among related individuals during the first year of life, especially for twins. The data are ambiguous at best and probably unsupportive of similarity in the second year and not at all positive with respect to the childhood period (though only siblings have been adequately examined at the older ages). Thus, the data suggest the tentative hypothesis that similarity in developmental profile among related individuals, especially twins, is associated with genetic differences in birth condition but that, as the determinants of test behavior become more complex and more "mental" in nature, the degree of association with such genetic factors becomes blurred by nongenetic correlates.

III. CORRELATIONS BETWEEN PARENT AND CHILD IQ ASSESSED AT THE SAME AGE

Almost all correlations in the literature between the IQs of parents and their children are for the case in which the parent is assessed as an adult and the child is assessed at some age in childhood. Sometime after the child reaches 2½–4 years of age, these adult-parent and child IQ correlations are approximately .50 (Erlenmeyer-Kimling & Jarvik 1963).

Very little data exist that describe the degree of relationship between the IQ of a child and the IQ of his parent assessed at the same chronological age. For example, what is the parent-child correlation between IQ when both parent and child are assessed at age 6 years? Obviously, only a few longitudinal studies are able to furnish such information. Eichorn (1969) has reported such data on infants from the Berkeley Growth Study, and table 5 presents her same-age and adult-parent and child correlations for the Bayley infant tests administered at various points during the first 3 years of life. The Stanford-Binet was given to the parents at age 17. The two sets of correlations are quite similar during the infancy period, and neither the same-age or adult-parent assessments are related to the child's score before his second birthday. When the WAIS, assessed at parental age 36, was used as the adult-parent measure, the adult-parent and child correlations were somewhat higher at the younger ages.

Relatively less data are available for same-age parent-child IQ correlations for the childhood years. Burt (1966) and Reed and Reed (1965) suggested a correlation of approximately .56, but the age of the subjects at testing was the "early school years." McCall (1970a) reported same-age parent-child correlations for a sample of 35 pairs at 11 ages between 3 and 12 years. The median correlation was .29, but it was from .43 to .50 between 5½ and 7 years of age. However, these correlations were between "IQ scores" that were averages of raw scores at three adjacent ages. Moreover, missing data were filled by this moving mean technique, and several different IQ tests were used at different ages. The purpose of this section is

TABLE 5

PARENT-CHILD CORRELATIONS FOR MENTAL PERFORMANCE FROM THE BERKELEY GROWTH STUDY

Age (Months)	Same Age	Parent Age 17
6	—.21	—.05
12	—.06	.03
18	.04	—.05
24	.28	.13
36	.29	.35

SOURCE.—Eichorn 1969.

to eliminate these qualifications and provide a description of the degree of relationship between parent and child IQ when assessed at the same age in development and compare these figures with the correlation between child and adult parent (age 17 years).

Subjects and Assessments

All parent-child pairs in the Fels sample who had at least one Stanford-Binet assessment at comparable ages in development were selected for analysis. Only Binet test scores and no filled values were used. Following correction of the scores by Pinneau's (1961) method, all scores were standardized within each age on the basis of the entire Fels sample (average N per age = 151).

The method of subject selection meant that a particular parent-child pair might provide data at more than one age but not necessarily at all ages. Moreover, there was not total overlap in the samples used for same-age parent and the adult-parent correlations. Unfortunately, if complete data had been demanded there would not have been sufficient numbers of cases to perform a meaningful analysis.

Results

Table 6 presents the results for each parent-child sex group separately at several childhood ages. At the top are the same-age parent-child correlations at three ages during development. Thus, the correlation of .71 between mothers and their sons represents the correlation between Pinneau corrected, standardized Stanford-Binet IQ scores for mothers and sons obtained when both were between 40 and 45 months of age. At the bottom half of table 6 are the correlations between parents' 17-year Stanford-Binet and their child's Binet IQ assessed at four different ages during childhood. The median same-age r was .29, while the median adult-parent r was .53 for correlations assessed on the child between 40 and 126 months.

It should be noted that many of the correlations in table 6 are based upon statistically miniscule samples. They are included in this form because

TABLE 6

PARENT-CHILD STANFORD-BINET IQ CORRELATIONS BY SEX

	MOTHER				FATHER			
	Daughter		Son		Daughter		Son	
	r	*N*	*r*	*N*	*r*	*N*	*r*	*N*
	Same-Age Parent and Child							
Parent and child age (months):								
40–45	.37	16	.71***	13	—.73*	7	—.57	8
70–78	.55**	18	—.38	16	—.19	9	.49	12
115–126	.54**	18	.16	14	.21	14	.46*	14
	Adult Parent (Age 17) and Child							
Child age (months):								
28–33	.50**	17	.18	8	—.39	7	.31	8
40–45	.54**	21	.38	14	.42*	19	.70***	17
70–78	.59***	22	.13	12	.44*	15	.54*	13
115–126	.51**	18	.58*	11	.43	9	.65*	8

* $p < .10$ (two-tailed).
** $p < .05$ (two-tailed).
*** $p < .01$ (two-tailed).

of the rarity of the data and because some readers are concerned with possible dependencies between sex of parent, sex of child, and age, which are not discussed below. Sample sizes can be increased by collapsing over sex or age. This was statistically possible since the scores were standardized, but it would obscure any qualifications that sex and/or age might impose. Table 7 presents the data for child and same-age parent and child and adult parent for the case in which the sex of parent was ignored and then when the sex of child was ignored. Age has been ignored in table 8 which presents the parent-child correlations for the four sex groups when the parent is assessed at the same age as the child or as an adult. For this table, if a parent-child pair had data at more than one age, a random number table was used to select the one age used for that subject pair.

Discussion

Several consistencies in these results can be discerned. First, correlations between child and adult parent were generally higher and more consistent than child and same-age parent. For example, if the adult-parent data for the 28–33-month assessment are not considered (there were no same-age parent-child correlations available for this age period), then nine of the 12 adult-parent versus five of the 12 same-age parent-child correlations were beyond the $p < .10$ level when the individual sex groups are considered (table 6). Moreover, 10 of the 12 comparable *r*'s are higher for adult-parent than same-age parent-child correlations, although statistical

TABLE 7

PARENT-CHILD IQ CORRELATIONS IGNORING CERTAIN SEX CLASSIFICATIONS

	PARENTS				CHILDREN			
	Daughters		Sons		Mothers		Fathers	
	r	N	r	N	r	N	r	N
	Same-Age Parent and Child							
Parent-child age (months):								
40–45	.21	23	.21	21	.50***	29	—.59**	15
70–78	.33*	27	—.04	28	—.01	34	.32	21
115–126	.33*	32	.31	28	.30*	32	.29	28
	Adult Parent (Age 17) and Child							
Child age (months):								
28–33	.22	24	.20	16	.42**	25	—.05	15
40–45	.46***	40	.62****	31	.48***	35	.60****	36
70–78	.47***	37	.42**	25	.42**	34	.49***	28
115–126	.43**	27	.61**	19	.51***	29	.43*	17

* $p < .10$ (two-tailed).
** $p < .05$ (two-tailed).
*** $p < .01$ (two-tailed).
**** $p < .001$ (two-tailed).

tests cannot be performed because there is overlap between the subjects in these distributions. When the sex classification for either parents or children is ignored, this trend is even more emphatic. For example, when parental sex group is ignored, all six correlations are higher for adult parent than same-age parent (table 7), and when sex of child is ignored five of six *r*'s were higher for adult-parent relationships. Moreover, the correlations are more consistent over age and sex groups for adult parent than for same-age parents. For example, 10 of the 12 adult-parent *r*'s (table 7) but only three of the 12 same-age *r*'s are within ±15 points of their respective medians.

From an environmental standpoint it is not surprising that the 17-year

TABLE 8

PARENT-CHILD IQ CORRELATIONS DURING CHILDHOOD BY SEX GROUP

	MOTHER				FATHER			
	Daughter		Son		Daughter		Son	
	r	N	r	N	r	N	r	N
Same age	.38*	35	.04	27	.14	21	.22	22
Adult parent	.53**	25	.31	16	.45*	19	.57*	17

* $p < .05$ (two-tailed).
** $p < .01$ (two-tailed).

IQ of a parent should be more highly correlated with his child's performance than that same parent's performance as a child on the identical test. Adult-parent correlations are a function of data gathered at more nearly the same period in the family history. If the parents' behavior has any influence on the child's performance, one should expect a higher degree of relationship for the adult-parent correlations. This holds true despite the fact that the adult-parent responds to different test items than the child. These data suggest that parent-child correlations are apparently higher when the assessments are made at nearly the same point in calendar time than if they are widely separated in time but both individuals respond to the same item type.

Another interpretation involves the fact that the reliability of the test changes somewhat with age. Terman and Merrill (1937) report that Binet reliabilities for assessments at ages younger than 6 years are slightly lower (.88) than for older ages (.93). Thus, same-age correlations, especially at the younger ages, might be lower than adult-parent *r*'s because of the relatively lower reliability of the parent's score. However, since the adult-parent relationships are higher than the same-age parent-child correlations even at the older ages, it is unlikely that differential reliabilities would account totally for this trend.

After the child reaches 3 years of age there is very little change in the adult-parent correlations, though there appears to be an increasing degree of relationship for cross-sexed parent-child pairs when parent and child were assessed at the same age. Finally, the correlations are higher and more consistent in either case (but especially for the same-age *r*'s) for like-sexed parent-child pairs than for cross-sexed. At the very least, these sex differences highlight the need to analyze separately for different sex pairings.

IV. PATTERNS OF IQ CHANGE OVER AGE

As indicated in Chapter I, most longitudinal studies have reported that subjects changed in IQ throughout childhood; however, the description of these changes has been somewhat limited. Many studies simply report that the average IQ for the entire sample shifted over age (usually upward for middle-class samples) and that a few individual subjects showed dramatic changes in performance. The most comprehensive display of individual patterns of IQ change was presented by Sontag et al. (1958) in which smoothed developmental profiles for IQ were graphed for 140 Fels subjects between 3 and 12 years of age. In searching for correlates of change in IQ, most investigators have a priori divided subjects into "increasers" or "decreasers" in IQ, and frequently this dichotomy of subjects has been based on only the two end scores during childhood, ignoring the pattern of performance in between. Consequently, except for Sontag et al., there has been no comprehensive attempt to describe individual differences in the nature of IQ change throughout the entire childhood period on more than a case-study basis and using more than two assessments at different points in development.

The analyses presented in this chapter attempt to describe predominant subgroups of subjects in the Fels sample having different trends of IQ change between the ages of 2½ and 17 years. In a subsequent chapter, various correlates of these patterns will be reported.

METHOD

Subjects

Subjects were drawn from the Fels Longitudinal Study (see Kagan & Moss 1962; and Sontag et al. 1958). Children in this study were administered 17 IQ tests during childhood. Subjects were selected for analysis if they possessed Stanford-Binet assessments between the ages of 2½ and 17 years such that no more than three consecutive assessments were missing. A subject with complete data would have a Binet IQ score at 2½, 3, 3½, 4, 4½, 5, 5½, 6, 7, 8, 9, 10, 11, 12, 14, 15, and 17 years of age. Testing was customarily done within a week of the child's birthday, and more than 90% of the assessments were made by the same examiner (Virginia Nelson).

A total of 80 subjects, 38 males and 42 females, qualified for analysis on the basis of this criterion. These children were born between April 1930 and July 1938, and the group included 14 sibling dyads, three sibling triads, two sets of twins, one set of triplets, and four adopted (as infants) children. Two subjects were Negro, and all others were white. The subjects were from small to medium-sized mid-Ohio towns, 37% were firstborn, and they ultimately acquired 14.46 years of education ($s = 2.56$; 12.0 represents high school graduation). The parents of the subjects had a mean midparent IQ (Otis test) of 103.28 ($s = 15.20$), and 13.07 ($s = 3.06$) years of education. These sample characteristics were comparable for male and female subjects.

At age 10 the average IQ for the subjects in the sample was 118.76 ($s = 19.30$). A total of 160 subjects in the entire Fels population were administered a Binet IQ at age 10, and their average was 118.29 ($s = 19.64$). While the total group includes the sample to be studied in this report, the figures are so close that it is reasonable to assume that children who had relatively complete data in the Fels sample were not unique in the level and variability of IQ from children who missed too many assessments to be included.

The major longitudinal studies are often criticized for having samples biased toward the upper-middle class. While it is true that the average IQ of the children studied is substantially above 100 (e.g., Fels = 117, Berkeley Growth Study = 127, depending upon the age of the assessment), these samples are nevertheless representative of a large segment of the American population. First, the Fels sample has normal within-age variability in IQ despite its relatively high mean. Second, the parents of the children in this sample did not score disproportionately high on the Otis test. Third, the variability in education and socioeconomic circumstances of the parents is considerable. Consequently, while no claim is made that the Fels sample is representative of the entire American population, it is representative of a substantial segment of it, and this group is considerably less biased than the usual psychological research sample composed of college sophomores or albino rats.

IQ Assessments and Missing Data

Subjects were administered the Stanford-Binet at each of the 17 ages listed above. Initially in the Fels study the 1916 version of the Binet was used. The 1937 revision was employed as soon as it was available, and the scores on Forms L and M were adjusted by Pinneau's (1961) correction. Furthermore, all children having had at least one 1916 assessment were given the 1916 test again at age 14. Consequently, 46 subjects (22 males, 24 females) born prior to November 10, 1934 were given at least one 1916 Binet assessment while the remaining subjects were administered only the 1937 Binet.

The fact that two different test forms (the 1916 vs. the 1937 Binet revisions) were used meant that some adjustment to the raw data had to be made. Consequently, after Pinneau corrections, all 1916 assessments within each age were converted to distributions having a mean and standard deviation equivalent to the 1937 revision for the Fels subjects at that age. The statistical correction was made by using means and variances for the entire Fels population rather than only those subjects included in this subsample.

The criteria for including a subject in this sample permitted some amount of missing data. The 80 subjects included in the sample possessed an average of 14 of the 17 possible IQ assessments. The protocols of 17 subjects contained one instance in which two consecutive tests were missing, and one subject had one instance in which three consecutive tests were missing. No subject had more than one instance of a run of either two or three consecutive tests missing. Missing data were handled by the linear interpolation-extrapolation procedure described previously on page 20.

Statistical Analyses

General purpose.—The purpose of the analyses was to cluster subjects together into groups such that the patterns of IQ change over age within a group were relatively homogeneous while the general IQ pattern for one group was different from that for another group. Hopefully, the result of such analyses would be a small number of groups which would represent the predominant patterns of IQ change over age in the sample. The task was akin to the issue once raised about "the learning curve" (e.g., Tucker 1960). Just because a group of subjects shows a smooth ogival increase in average performance on a learning task does not mean that all of the subjects in that group possess an individual pattern similar to the average group trend. Rather, the general pattern may be composed of several "component trends" which characterize subgroups of subjects within the sample and which produce the general trend when averaged. What is the nature of such component trends for IQ profile?

Statistical details.—Several different statistical procedures could have been applied directly to the subject's 17 raw IQ scores. However, the nature of such techniques is that one is likely to get many small groups of subjects which differ in their pattern of IQ change by small inflections at two or three ages. Rather, it seemed desirable to concentrate on the more general characteristics of shifts in IQ, ignoring small isolated deflections which may be attributed more to measurement error than real shifts in mental performance.

A procedure that would represent an intermediate step in reducing the raw IQ scores into a few major developmental trends has been described by Tucker (1960, 1966). This procedure decomposes the Subjects × IQ Tests matrix by the Eckart-Young procedure. The result is a set of components which constitute generalized patterns of IQ change over age plus a matrix of component scores for each subject which represent the direction and extent to which that subject's actual pattern of IQ change is approximated by the generalized components. Since there were 17 data points for each subject, 17 components would completely determine the system such that for any single subject the sum of the product of his component scores times the IQ value of that generalized component at a specific age would equal that subject's actual IQ at that age. This process could be repeated for each age, and thus a subject's entire pattern could be reproduced. However, the purpose of the analysis was to reduce the system from 17 components to the fewest number which account for the largest segment of the total variability.

By examining relative root size, the first four components were selected for further analysis. The first component in such an analysis is often related to the grand mean of the entire sample, and a subject's component score reflects not only the extent to which his developmental IQ profile follows the group trend but also how much above or below in general level that subject is from the typical value. Component scores for the subsequent, mutually orthogonal components reflect the extent to which that subject's actual IQ pattern deviates from the main trend in the manner described by each component.

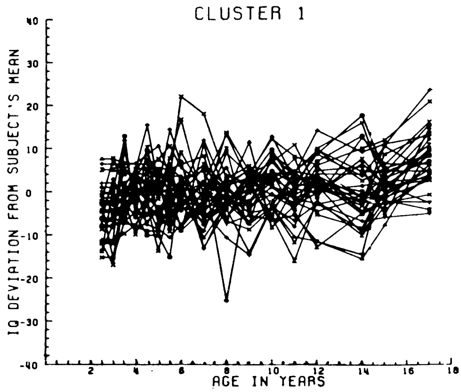
The psychological meaning of the components themselves is controversial. Presumably, like factors from a factor analysis, these components represent prototypical patterns of IQ change over age and would themselves constitute the desired result. However, certain major limitations on such an interpretation have been pointed out by Cronbach (1967). As is true with any factor analysis, the composition of the factors is a function of the variable set submitted for analysis (in this case, the ages at which subjects were tested). Moreover, factor-analyzing data that approximate a simplex structure (i.e., correlations between adjacent assessments are high and progressively decrease with increases in test intervals) is known to produce certain inevitable consequences. The present data were "simplex-like" but did not fulfill all of the requirements and expectancies of a simplex structure. Although there is controversy surrounding the interpretation of the components themselves, the component scores for each subject represent a convenient summarization in a few measures of a subject's general pattern with relatively less imposition of a priori structure than many other techniques (polynomial curve fitting, orthogonal polynomial components, etc.). Moreover, the components could have been rotated to one of several criteria. Since the purpose of using this intermediate procedure was data reduction and not the interpretation of the components per se, rotations were not used.

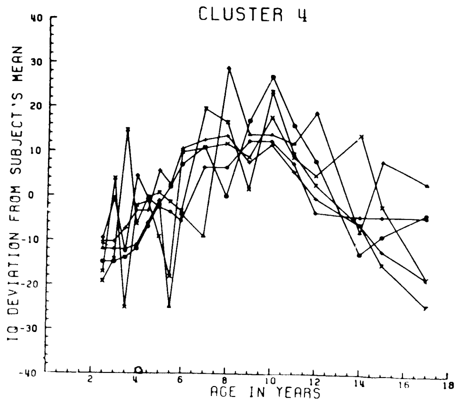
The scores for each subject for the first four components were then subjected to a direct cluster rotation described by Overall and Klett (1972, p. 207) which groups subjects into clusters on the basis of their locus in the four-dimensional space defined by the four Tucker component scores. The technique involves "progressive transformation of a matrix of Q-type vector products until the individuals are tightly grouped into clusters. . . . The rank of the original matrix is reduced through the successive transformations since, in the final matrix, several rows associated with individuals who belong to a single cluster will have identical 0–1 elements. The rank of the cluster-transformed Q-type matrix determines the number of clusters" (Overall, personal communication, 1972). Geometrically, an individual is represented by a vector emanating from the origin to his data point in four-space. "The individual axes are progressively rotated towards other axes with which they have large cosines and away from axes with which they have smaller cosines. This process tends to make the larger cosines larger and the smaller ones smaller until all of the cosines approach either zero or one" (Overall, personal communication, 1972). In the present case, the result should be groups of subjects having similar patterns of IQ change over age within groups but different trends between groups.

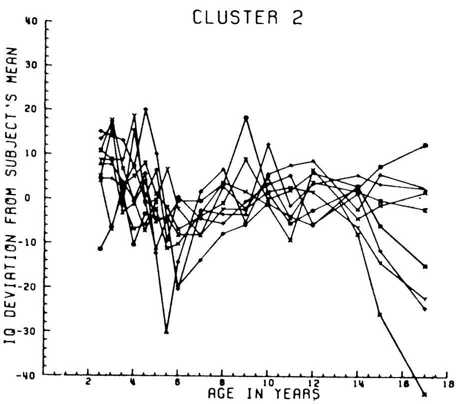
RESULTS

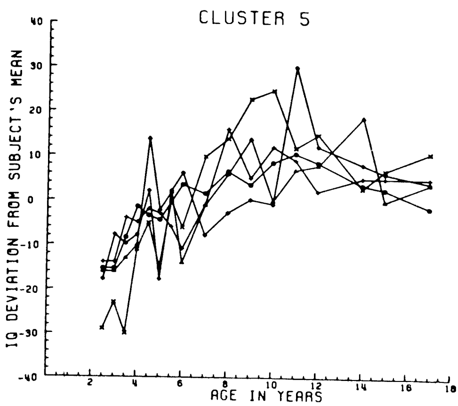
Description of IQ profiles.—The cluster analysis on the four component scores yielded five groups plus 13 subjects who were considered isolates (i.e., were assigned to groups of three or fewer subjects). For the purpose of evaluating the distinctiveness of these clusters and the homogeneity of IQ pattern within clusters, the average of each subject's 17 scores was subtracted from each of his scores producing deviation IQs about that individual subject's mean. Then, the subjects within each cluster were plotted so one could visually assess the homogeneity of subject patterns within a group as well as the differences in IQ pattern between groups.

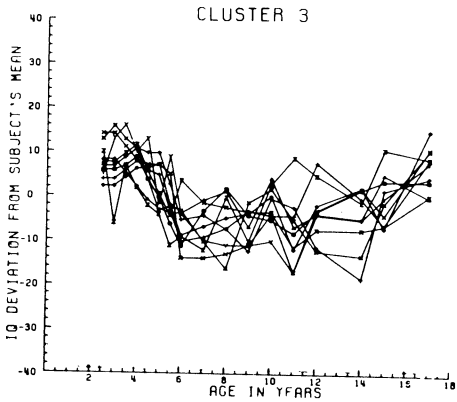
These plots are presented in figure 4. Cluster 1 (cluster numbers are arbitrary) was the largest group with a total of 36 of the 80 subjects (17 males, 19 females). The general trend was one of minimal systematic deviation from a slightly rising pattern throughout childhood. Interestingly, each member of the two sets of twins, two of the three members of the set of triplets (the odd member was an isolate), and all four adopted children in the sample belonged to this group. Cluster 2 contained nine subjects (five males, four females) and was characterized by a sharp decline in IQ performance between the ages of 4 and 6, followed by a slight recovery, and considerable heterogeneity of pattern after age 14 with the principal trend downward. Cluster 3 also showed a preschool decline, but then relatively level performance between 6 and 14 years followed by an upswing. Ten

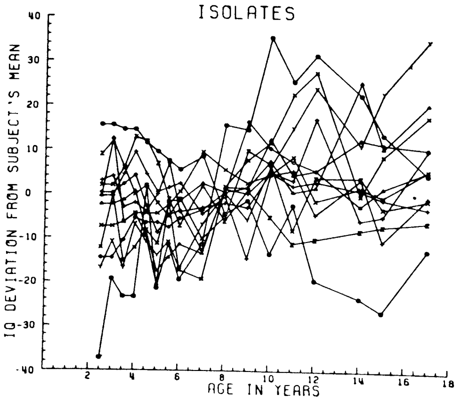

CLUSTER 1
IQ DEVIATION FROM SUBJECT'S MEAN
AGE IN YEARS


CLUSTER 4
IQ DEVIATION FROM SUBJECT'S MEAN
AGE IN YEARS


CLUSTER 2
IQ DEVIATION FROM SUBJECT'S MEAN
AGE IN YEARS


CLUSTER 5
IQ DEVIATION FROM SUBJECT'S MEAN
AGE IN YEARS


CLUSTER 3
IQ DEVIATION FROM SUBJECT'S MEAN
AGE IN YEARS


ISOLATES
IQ DEVIATION FROM SUBJECT'S MEAN
AGE IN YEARS

subjects composed this group with a predominance of females (two males, eight females). Cluster 4 displayed an inverted-*U* trend, peaking between 8 and 10 years of age. The seven subjects in this group (three males, four females) showed the greatest magnitudes of shift in IQ. Cluster 5 (two males, three females) was characterized by a strong and steady rise in IQ until ages 8–10, similar to that observed in Cluster 4 but with less of a subsequent decline. The isolates showed considerable variability at the young and old ages, and the individual trends are not nearly as homogeneous as in the other clusters.

One of the purposes of this analysis was to derive clusters that were homogeneous for pattern of IQ within a cluster but heterogeneous between clusters. Figure 4 displays the degree of homogeneity that existed within each cluster for relative changes in IQ over age. Since subjects' scores were taken as deviations about their own means, these graphs display only homogeneity for *relative* pattern and conceal any differences in general level that may exist between subjects and clusters. Thus, figure 5 displays the average IQ (*not* deviation IQ) for each cluster at each age. The ordinate is labeled "Adjusted Binet IQ" to reflect the adjustment of 1916 scores to the 1937 revision scale of measurement.

Analysis of the difference between clusters.—To determine whether the clusters derived by the above analyses were significantly different in general level of IQ and/or pattern over age, one would customarily perform a Clusters × IQ Assessments repeated measures analysis of variance. However, such analyses are dependent upon the assumption of equality of covariance between all pairs of repeated assessments, and violations of these requirements inflate the size of F (Box 1954; Davidson 1972; McCall & Appelbaum 1973). Since these assumptions usually cannot be met by developmental data, one alternative approach is to determine the orthogonal polynomial contrasts for the linear, quadratic, etc., components for each subject, just as one would do if one were performing a trend analysis on these data. These several variables embody the information in the degrees of freedom for the repeatedly measured variable and can be submitted as dependent measures to a multivariate analysis of variance with Clusters as the independent variable. The multivariate test of the grand mean provides a test of the IQ Assessment main effect (i.e., general IQ trend over clusters), and the multivariate test of Clusters can be interpreted as an evaluation of the Clusters × IQ Assessment interaction (i.e., clusters differences in IQ profile contour). A univariate test between Clusters using the mean of each subject's scores as the dependent variable yields a test of the main effect for Clusters. This multivariate approach does not assume homogeneity of covariance. The rationale and details of such analyses have been presented by Bock (1963),

FIGURE 4.—Plots of each subject's deviation (from his mean) Binet IQ over age for the five IQ clusters and the isolates.

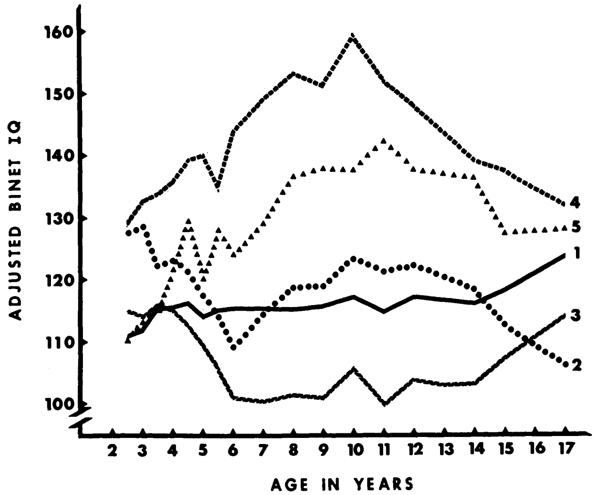


FIGURE 5.—Mean IQ (adjusted for differences between Binet revision) over age for the five IQ clusters.

McCall (1970b), McCall and Appelbaum (1973), and Morrison (1967).

When the linear, quadratic, cubic, and quartic contrasts were entered into the multivariate analysis (the higher order trends were not included since there was little interest in differences between clusters for these polynomial degrees), the clusters were found to differ in their pattern of IQ change over age, $F_{mult}(16, 181) = 16.41$, $p < .001$. Moreover, the univariate linear, quadratic, cubic, and quartic contrasts were all significant at $p < .01$. Consequently, the clustering did produce groups having significantly different patterns of IQ change over age.

In addition, there was a main effect for IQ assessment, $F_{mult}(4, 59) = 2.85$, $p = .03$, which was predominantly linear in form, $p = .006$. This result implies that the sample as a whole (excluding the isolates) had a linear increasing pattern of IQ change. Moreover, the mean IQ for the five patterns also differed, $F(4, 62) = 9.05$, $p < .001$. Although one must be careful in interpreting main effects when a higher-order interaction is significant, it is clear from figure 5 that the groups did differ not only in pattern of IQ but also in general level. This result implies that pattern of IQ change is apparently not totally independent from general level: children having one pattern of IQ change tend to have higher general levels of IQ than those having other developmental profiles. This result has implications for studies of heritability of IQ pattern over age (see Chap. II), since one would expect greater similarity between related individuals for pattern of IQ change merely because of similarity in general level and the association between general level and specific patterns.

Comparisons between specific IQ clusters.—The nature of the five IQ groups discerned above suggested more specific comparisons between one or several of the five. For example, Cluster 1 represents relatively little systematic change in IQ while Clusters 2 through 5 display considerably greater year-to-year fluctuations. Thus, when correlates of patterns of IQ change over age are considered (e.g., parental behavior), one might wish to compare Cluster 1 with the mean of the other groups to determine if that variable is different for children showing a minimum amount of IQ change over age as opposed to those showing more marked fluctuations.

Another contrast of interest was the observation that Clusters 2 and 3 showed essentially declining preschool patterns while Clusters 4 and 5 showed markedly increasing preschool patterns. To determine whether or not these contrasting trends were indeed significantly different from one another, Clusters 2 and 3 were combined and compared with the combination of Clusters 4 and 5 using a special contrast option in the multivariate analysis on all five groups reported above. The analysis indicated that the pattern of IQ change was different for these two subsets, $F_{\text{mult}}(4, 59) = 45.01$, $p < .001$, as well as their average IQ, $F(1, 62) = 21.53$, $p < .001$. These results indicated that it was reasonable to compare preschool IQ increasers and decreasers.

One could also ask whether Clusters 2 and 3 differed from one another. While they were both declining patterns, their trends diverged somewhat in adolescence (see fig. 5). The multivariate analyses again indicated a difference in IQ pattern, $F_{\text{mult}}(4, 59) = 13.78$, $p < .001$, and in mean IQ, $F(1, 62) = 4.29$, $p = .04$.

Finally, Clusters 4 and 5 also showed similar patterns until approximately age 8, after which Cluster 4 lost its relative advantage while Cluster 5 did not. The profile contours, $F_{\text{mult}}(4, 59) = 6.28$, $p < .001$, as well as the general level, $F(1, 62) = 6.57$, $p = .01$, of these two clusters were significantly different.

DISCUSSION

The procedures described above distinguished five groups of children representing different patterns of IQ change over age. These clusters were relatively homogeneous for IQ pattern for subjects within a cluster, and the five clusters were significantly different from one another in their IQ pattern over age. Thus, significantly different patterns of IQ change over age can be found in such a sample of middle-class children.

The predominant pattern involving 45% of the subjects was one of relative stability in IQ over age with a slight increase in performance with development. The remaining patterns were more variable in IQ over age; two were predominantly decreasing and two largely increasing during the preschool period. The decreasing patterns were significantly different from the increasing, and the two patterns within each of these sets were different from each other.

It should be pointed out that several rather arbitrary decisions were made in performing these analyses, and the results are dependent upon those choices. The use of (1) Tucker's procedure as an intermediate step, (2) 17 IQ assessments, (3) unrotated components, (4) four components subjected to the clustering routine, (5) Overall and Klett's clustering technique, etc., all represent decisions which affected the outcome. Consequently, it is not accurate to say there are five types of IQ change over age or that the profiles observed are *the* patterns of IQ change. Rather, these clusters represent one of several schemes for classifying groups of individuals, and the test of their value rests in whether other variables differentiate between these clusters and supply interpretations for these IQ patterns.

V. CORRELATES OF PATTERNS OF IQ CHANGE

It was of interest to determine whether the five IQ pattern clusters developed in the previous chapter could be distinguished on the basis of the behavior of the parents of the subjects. Although a considerable body of literature exists on parental correlates of IQ (e.g., Freeberg & Payne 1967), relatively little information is available on the parental correlates of pattern of IQ change (e.g., Sontag et al. 1958).

PARENTAL BEHAVIOR CORRELATES OF IQ CHANGE

Procedure and Variables

As indicated above, a trained observer was sent to the home every 6 months during the first 6 years and every year thereafter to make global assessments of parental (usually maternal) behavior. The details of this procedure and these variables are described by Baldwin, Kalhorn, and Breese (1949) and Champney (1941). The original Fels Parent Behavior Rating Scales contained 30 dimensions, but a series of factor analyses (Baldwin et al. 1949; Crandall & Preston 1955; Lorr & Jenkins 1953; Roff 1949) reduced the original collection to 10 which will be considered here. These variables were rated independent of any knowledge of the child's IQ scores. A brief description of these dimensions follows.

The *adjustment of the home* was a rating of the general internal adjustment of the family as a whole in its day-by-day relationships. A well-adjusted home (high rating) was characterized by satisfaction, stability, achievement, and happiness while a maladjusted home (low rating) was thwarting, unpleasant, repressive, and insecure.

A second variable was the *restrictiveness of the regulations* set up or implied by the parents as standards to which the child was expected to conform. A home was judged to be restrictive (high rating) if the requirements on the child were numerous and severe and if the child would be highly circumscribed in his behavior by these standards. Low ratings were assigned if there were few and mild requirements and the child had a great measure of freedom.

The *severity of actual penalties* reflected the severity of penalties imposed when the parent took an official note of misconduct. High ratings were given when the penalties were severe and occasionally produced fear or personal resentment in the child, while low ratings were assigned if the penalties were light and inconsequential and a great deal of misbehavior provoked only weak verbal remonstrances.

The *clarity of the policy of regulations and enforcement* was an assessment of the clarity with which the parents' standards of child conduct were communicated to the child. High ratings were given if the regulations and the requirements were clearly formulated and consistently executed so that the child was able to know what was expected of him and what would happen if he failed to conform. Low ratings were assigned if the parents' standards and policies were so vague or fluctuating that the child had little chance of adjusting.

The *coerciveness of suggestion* was a rating of the parents' dictatorial quality in dealing with the child's immediate behavior. High ratings were given if the parent attempted to control a situation by issuing orders or commands which he expected to be obeyed. Lower ratings were given if the parent made mere optional suggestions and avoided coercing the child.

The parents' *accelerational attempt* reflected the degree to which the parents strove to increase the rate at which the child's behavior matured. High ratings were assigned if the parent deliberately trained the child in various mental and motor skills which were not yet essential, whereas low ratings were given if the child was left to "grow naturally" or was even shielded from acceleration influences.

General babying consisted of the parents' tendency to help the child through the ordinary difficulties of everyday life. High values were assigned if the parents insisted on helping in situations where the child was quite capable, while relatively low values were given if the parent withheld aid even in major difficulties.

The *general protectiveness* of the parents characterized the parental response to threats and hazards to the child's well-being. High ratings indicated that the parent tended to keep the child unnecessarily sheltered and prevented the child from encountering difficult circumstances, while low ratings were given if the parent allowed the child to be exposed to dangers, perplexities, and difficulties.

The *direction of criticism* reflected the relative emphasis on reward versus punishment. High ratings were given if the parent favored approval, praise, and acclaim, whereas low ratings were given if parental responses typically expressed disapproval or blame.

Affectionateness was the extent of the parents' expression of affection for the child. High values were assigned for parents who manifested a warm, personal affection toward the child, while low ratings were given for parents who were matter-of-fact, unemotional, or outright hostile toward their children.

Unfortunately, the parental ratings were not begun until 1938, and consequently subjects varied in the number of such assessments (some had to be eliminated for insufficient data) and in the age period of greatest concentration of such assessments. Thus, the average of all available ratings between 3 and 13 years was computed for each child for each of these 10 variables. Therefore, over the entire group these parental assessments are somewhat more typical of parental behavior during the older childhood years.

The intra- and interrater reliability of these 10 variables has been reported in detail by Baldwin et al. (1949). The median interrater reliability for a single assessment of these variables was .76. The reliability of the average over childhood which was used in the following analyses might be somewhat higher.

Parental Behavior Differences for the Five IQ Clusters

Cluster differences.—The number of subjects in the five clusters having parent behavior ratings were 34, 9, 10, 6, and 5, respectively. The 10 parent behavior ratings were submitted to a Cluster × Sex multivariate analysis of variance, an extension of univariate analysis of variance to handle more than one dependent variable at a time. A nontechnical description of multivariate analysis of variance has been given by McCall (1970b).

No analysis reported in this chapter revealed a sex-of-child difference or an interaction between sex and IQ cluster for these 10 variables. However, a significant main effect for IQ cluster was found, $F_{\mathrm{mult}}(40, 172) = 2.05$, $p < .001$, indicating that the five clusters could be differentiated on the basis of the parent behavior variables. Significant univariate effects were observed for severity of penalties ($p = .007$), clarity of policy ($p = .002$), coerciveness ($p = .018$), accelerational attempt ($p < .001$), and direction of criticism ($p = .032$). An examination of the weightings assigned to each variable in the analysis, the within-cells correlations (correlations between the parent behavior ratings within each IQ cluster pooled over clusters), and the univariate effects suggested that the severity of penalties and accelerational attempt were the two principal contributors to this multivariate effect. The severity of penalties variable was positively correlated (within cells) with restrictiveness ($r = .59$) and coerciveness (.69) but negatively related to babying (−.60), protectiveness (−.41), direction of criticism (−.59, implying more negative criticism), and affectionateness (−.56). Accelerational attempt was positively related to clarity of policy (.54) and direction of criticism (.54, implying more positive than negative criticism) but negatively related to coerciveness (−.57).

One strategy in attempting to substantiate the selection of particular variables as major contributors to a multivariate result is to demonstrate that (*a*) the selected subset of variables will produce a significant differentiation between groups when considered separately from the remaining variables, and (*b*) the remaining variables do not differentiate between groups if the subset is covaried. If, in addition, the subset differentiates the groups after the remaining variables have been covaried, the remaining variables apparently contribute little in addition to the special set.

The severity of penalties and accelerational attempt met these conditions. When the cluster differentiation was attempted on the basis of the severity of penalties and accelerational variables alone, the multivariate test was still significant, $F_{\text{mult}}(1, 06) = 5.07$, $p < .001$. If the remaining eight variables were assessed covarying severity of penalties and accelerational attempt, the multivariate test was not significant, $F_{\text{mult}}(32, 168) = 1.32$, $p = .13$, and severity of penalties and accelerational attempt still differentiated the IQ clusters even after the other eight parent variables were covaried, $F_{\text{mult}}(8, 90) = 4.13$, $p < .001$. Thus, severity of penalties and accelerational attempt were necessary and sufficient parental variables for differentiating the five IQ pattern clusters.

Table 9 presents the means for the 10 parent behavior ratings for the five IQ clusters. Cluster 2 was the least severely penalized and had parents who provided minimal acceleration. Cluster 3 was the most severely penalized and was offered the lowest level of accelerational stimulation. Thus, the two generally declining patterns were both characterized by very little accelerational attempt on the part of the parent but either the severest or mildest penalties for the sample. In contrast, the two clusters demonstrating increasing preschool IQ were characterized as having parents who were very accelerating, substantially rewarding, clear in their policies, but medium to fairly severe in their penalties. They seemed to present an encouraging and rewarding atmosphere but one with some structure and enforcement of policies. Cluster 1, which showed very little change in IQ, did not have extreme parents on any of the dimensions measured.

Since Cluster 1 represented relatively little change in IQ over age while Clusters 2–5 evidenced considerably more developmental variability, one might inquire whether the parents of children in Cluster 1 differed from those in the other groups. When this contrast (Cluster 1 vs. the mean of Clusters 2–5 for the 10 parent variables) was performed, there was no multivariate significance for these 10 parent behavior ratings, $F_{\text{mult}}(10, 45) = 1.23$, $p = .30$. Moreover, there were no significant univariate tests for this comparison.

In summary, the five clusters representing different patterns of IQ change over age could be discriminated from one another on the basis of the 10 parent behavior ratings. This was not simply a distinction between the behavior of parents having children who changed a great deal in IQ versus those who did not but rather appeared to be related to the nature

TABLE 9

MEANS FOR PARENT VARIABLES FOR THE FIVE IQ PATTERN GROUPS

Cluster	Variable										Mid-parent Education	Mid-parent Otis IQ
	Adjust.	Restrict.	Sev. Pen.	Clarity	Coerciv.	Acceler.	Babying	Protect.	Dir. Crit.	Affection		
1 ($N = 34$)	66.9	57.5	55.4	66.8	58.9	56.8	53.4	58.7	52.2	62.7	12.3	101.1
2 ($N = 9$)	63.5	52.1	47.7	60.3	51.6	49.6	57.0	57.5	56.4	67.0	13.9	109.4
3 ($N = 10$)	60.8	61.9	58.2	57.9	64.9	49.2	54.1	55.6	47.5	61.0	11.4	92.2
4 ($N = 6$)	67.0	54.8	53.8	73.5	51.7	72.2	54.5	60.0	56.5	64.3	15.9	123.0
5 ($N = 5$)	67.0	53.0	51.8	69.4	52.4	64.2	52.2	54.2	58.4	64.2	14.6	106.0

of the change in IQ over age. Specifically, children who evidenced predominantly declining IQ patterns in the preschool years tended to come from homes which did not make an attempt to accelerate or stimulate them and which had either very severe or very mild punishment regimes. In contrast, children showing increases in IQ until approximately age 8 came from homes which emphasized the acceleration of their child in an encouraging and clear manner but with structure and enforcement.

Secular trends.—Since sampling in the Fels study has occurred continuously since 1929, the patterns of IQ change described above might be associated with different secular periods. Thus, the year of birth for the children in each of the IQ patterns was subjected to an analysis of variance and a significant result obtained, $F(4, 62) = 4.13$, $p = .005$. Despite the fact that a highly significant difference was observed, the average year of birth differed by less than 3.5 years for the five IQ clusters. The average year of birth for the five groups was 1932.6, 1936.0, 1932.7, 1934.7, and 1934.6, respectively. Relatively little consistency appears in these differences, and the authors do not have a firm interpretation.

Speculatively, however, the end of World War II (i.e., 1945) seems to constitute an inflection point in each of the patterns (see fig. 5). On the average, 1945 occurred at age 13½, 11, 13½, 10½, and 10½ for the five IQ clusters, respectively. The end of the war marked the start of the largest sustained rise in IQ performance for Clusters 1 and 3, and the longest sustained decline in IQ for Cluster 4. The year 1945 also points to a major inflection point in the IQ patterns of Clusters 2 and 5. Unfortunately, because these secular differences between groups were not very great, it is difficult to tell whether inflections in IQ patterns are really associated with the termination of World War II or whether such inflections commonly occur between 10½ and 13 years of age. Nevertheless, the proposition that major social-political events and their consequences (i.e., the return of fathers to families, additional siblings, etc.) have some influence on mental performance represents an interesting hypothesis.

Parent education and IQ.—The five IQ groups were compared for midparent (average of both parents) years of education and parental IQ based upon the Otis test. Parental education was defined as the average number of years of education for the two parents where 12 years represented high school graduation. Parental IQ was determined by the first Otis test administered to the parents by Fels (all before 45 years of age), and the average for the two parents was taken unless only one parent was tested in which case that score was used. The means are presented in table 9.

There were significant differences between IQ clusters for both parental education, $F(4, 62) = 3.61$, $p = .01$, and parental IQ, $F(4, 57) = 4.98$, $p = .002$. The results for the two measures were similar but not totally consistent. Generally, the parents of subjects in the two declining patterns

(Clusters 2 and 3) had somewhat lower parental education and IQ than for the two increasing patterns (Clusters 4 and 5), but this was not uniformly true for both variables. Children in Cluster 3 had parents who had the least education and the lowest IQ, while the parents of children in Cluster 4 had parents with the most education and substantially higher IQs than the other groups.

Given that the IQ groups differed with respect to the educational and IQ levels of the parents, one might wonder whether the parent behavior ratings were direct reflections of the mental characteristics of the parents, or whether the parental variables displayed relationships with IQ pattern that surpassed those that might be expected on the basis of differences in parental education and IQ. Consequently, the two parental variables which were found to be the most salient discriminators between IQ groups, accelerational attempt and severity of penalties, were analyzed for cluster differences covarying parental education and parental IQ. The analyses indicated that the groups could still be differentiated on the basis of parental acceleration, $F(4, 47) = 3.42$, $p = .016$, but the test for severity of penalties only approached significance, $F(4, 47) = 2.07$, $p = .10$. Thus, even when cluster differences in parental education and IQ were covaried, accelerational attempt differentiated the five IQ clusters.

General level versus IQ pattern.—As indicated above, the five IQ groups differed from one another both in terms of general level of performance as well as the pattern of IQ over age. Therefore, although there are parental behavioral correlates of these IQ groups, it is not clear whether those concomitants are associated with the pattern of IQ change over age or with the fact that the clusters differed in the general level of IQ. Ideally, one would like to find children in each of the five groups who had similar general levels but showed different patterns of IQ change over age. These groups could then be compared on the parental variables and it could be decided straightforwardly if there were parental correlates of IQ pattern per se. Unfortunately, the size of the groups and the limited availability of the parental behavior ratings did not permit such an analysis.

There is no completely satisfactory method under these circumstances to determine if parental variables relate to IQ profile per se. One could attempt to adjust the parent behavior ratings for differences in the general level of IQ for subjects in the five groups and then ask whether significant differences between profile groups could still be observed. Thus, a multivariate analysis of covariance was performed using accelerational attempt and severity of penalties as the criteria and the child's mean IQ over the 17 assessments as a covariate. The multivariate test of group differences was still significant, $F_{\text{mult}}(8, 2) = 2.89$, $p = .006$, as was the univariate effect for severity of penalties, $F(4, 53) = 4.30$, $p = .004$. The accelerational attempt variable approached univariate significance, $F(4, 53) = 2.36$, $p = .065$. Thus, the parental variables of severity of penalties and accelerational attempt still discriminated between the five IQ groups even after these variables were adjusted for the fact that the groups differed in the general level of IQ performance, although the logic of such an analysis may be somewhat tenuous.

Special Group Comparisons

Increasing versus decreasing IQ trends.—Of the five IQ patterns developed in Chapter IV, two were predominantly increasing in the preschool years and two were decreasing. Therefore, in addition to asking whether the complete set of five clusters differed from one another on various parental measures, it was of some additional interest to ask whether or not the increasing differed from the decreasing patterns. It has already been demonstrated above that the IQ patterns themselves are different for these specific comparisons, and attention now turns to whether these particular sets of clusters differed on parent behavior ratings.

The 10 parent behavior ratings were employed in a multiple discriminant analysis which attempted to differentiate the two increasing IQ clusters (Clusters 4 and 5) from the two decreasing groups (Clusters 2 and 3). This analysis indicated a significant discrimination could be made, $F_{\text{mult}}(10, 45) = 3.61$, $p < .001$, with significant univariate effects for clarity of policy, $F(1, 54) = 14.75$, $p < .001$, and accelerational attempt, $F(1, 54) = 22.20$, $p < .001$. Subsequent analyses determined that the acceleration variable was necessary and sufficient to make the discrimination between increasers and decreasers, since when accelerational attempt was covaried from the other nine variables no significant discrimination between the IQ groups was found ($p = .22$), but accelerational attempt alone was a significant discriminator between these IQ patterns. Moreover, when the nine other parent behaviors were covaried, accelerational attempt was still different for the IQ groups, $F(1, 45) = 10.81$, $p = .002$.

It was then desirable to investigate whether accelerational attempt would discriminate between increasing and decreasing IQ patterns when the education and IQ of the parents were covaried. Thus, an analysis of covariance on the accelerational attempt variable with parental education and IQ covaried was performed. The results, $F(1, 52) = 15.06$, $p < .001$, indicated that children demonstrating an increasing IQ trend through the preschool years as opposed to those displaying a decrease had parents who provided more encouragement and accelerational attempts in amounts in excess of what one would predict on the basis of the parent's education and IQ.

The increasers and decreasers differed in general level of IQ as well as in their developmental pattern. When parental accelerational attempt was adjusted for child mean IQ differences, differences in accelerational attempt still existed between children showing an increasing versus a decreasing IQ pattern, $F(1, 53) = 6.44$, $p = .014$.

These data suggest that the salient parental difference between children who displayed markedly increasing patterns of IQ during the preschool and early school years as opposed to those who showed decreases rests in the accelerational attempts made by the parents. The fact that accelerating parents have children who show increases in IQ performance appears to go beyond the fact that those parents were also more highly educated and had higher IQs. Moreover, these differences in accelerational attempt are greater than might be expected on the basis of cluster differences in general level of child IQ. It is interesting to observe that the children demonstrating increases in IQ do so only until age 10, after which they level off or lose some of their relative advantage. Presumably, performance during the early years is under more direct control of the parent and that influence continues into the first several years of schooling in which the emphasis is on basic skills in reading and mathematics.

Increasers: Clusters 4 versus 5.—The groups showing an increase in IQ pattern differed in the extent of the decline in IQ following age 10. It was of interest to ask whether these two increasing groups could be discriminated on the basis of the parent behaviors. The multiple discriminant analysis did not reveal a significant discrimination ($F < 1$), and there were no significant univariate effects for any of the 10 parent variables. Thus, the parents' behaviors could not differentiate between the two increasing IQ patterns observed in the sample.

Decreasers: Clusters 2 versus 3.—A similar question was posed with respect to the two decreasing patterns of IQ. Both groups showed a marked decline in IQ until age 6, but then Cluster 2 showed an increase followed by a decrease in performance and Cluster 3 remained relatively low until age 15 when their performance increased. Would the parent variables discriminate between these two patterns?

The multiple discriminant analysis revealed a significant group effect, $F_{\text{mult}}(10, 45) = 3.04$, $p = .005$, and there were univariate effects for severity of penalties, $F(1, 54) = 14.12$, $p < .001$, restrictiveness, $F(1, 54) = 6.11$, $p = .016$, coerciveness, $F(1, 54) = 8.87$, $p = .004$, and direction of criticism, $F(1, 54) = 6.94$, $p = .01$. Subsequent analyses determined that the severity of penalties was the major variable. When severity of penalties was covaried from the other nine variables they failed to show a significant discrimination ($p = .13$, no significant univariate effects), and when the other nine variables were covaried severity of penalties still showed a significant discrimination between these two IQ groups, $F(1, 45) = 8.24$, $p = .006$.

The severity of penalties discriminated between the groups even after parental education and IQ were covaried, $F(1, 52) = 11.43$, $p < .001$, and after child mean IQ was covaried, $F(1, 53) = 8.75$, $p = .005$.

These data imply that for children showing marked declines during the preschool period, subjects whose parents were most severely penalizing tended to remain low in IQ throughout the elementary and middle school years but showed some increase in performance during high school. In contrast, decliners during the preschool period who experienced the least severe penalties regained some of their early relative losses in IQ during the early school years but declined to their lowest performance by the end of high school. The differences in the handling of penalties between these groups apparently supercede any differences in parental IQ and education as well as child mean IQ. Thus, extremes in parental behavior with respect to the severity of penalties has a common association with IQ during the first 6 years of life in which a steady declining pattern is observed. However, differences in the severity of penalties are related to differing patterns of IQ between 6 and 17 years of age.

IQ PATTERN AND SPECIFIC TEST COMPONENTS

Composition of the Binet over Age

The most obvious interpretation of the IQ patterns presented in this *Monograph* is that children shift in their mental performance as reflected on these general tests in different ways and at different ages. However, one of the hazards of longitudinal testing is that the assessment instrument used at one age is different, regardless of the constancy of its name and intent, than the assessment instrument used at another age. Obviously, the items on the Binet at age 2½ are considerably different in their character than those at age 17. Moreover, it is possible that the items on the Binet reflect different aspects of mental performance, and these aspects are not necessarily highly correlated. Consequently, a child could maintain the same relative position throughout childhood on each aspect of mental performance that is reflected on the Binet (for example) and yet show dramatic shifts in performance if the Binet changed its composition with respect to these aspects of mental performance from one age to the next. For example, suppose one child were outstanding in vocabulary and verbal skills and relatively poor in spatial-perceptual behavior. If the Binet places an increasing relative emphasis on vocabulary and verbal skills with increasing age levels, then this child should show an increase in IQ performance even though his relative abilities were constant throughout childhood. Similarly, a child not gifted in verbal behavior but skilled in spatial-perceptual tasks would show a declining pattern over age merely because the test progressively de-emphasizes spatial-perceptual performance.

Such a possibility would represent a succinct explanation of the results reported above. For example, children showing increases in IQ had parents who were more highly educated, had higher IQs, and provided a great deal of accelerational encouragement. One would expect these parents to promote verbal skills and behavior, and thus if the test showed an increasing emphasis on verbal performance one would expect these children to rise in their relative standing for such psychometric reasons. Conversely, children having parents of relatively less education and lower IQ might show the opposite trend.

It is extremely difficult to examine these interpretations since no other mental test data assessed independently of the IQ scores presented above were available. However, partial information on the plausibility of these hypotheses can be obtained by examining the item content of the Binet at different age levels. Meeker (1969) has scored the items on the Binet at each age level according to Guilford's (1967) structure of the intellect model. Thus, any item might contribute to particular Content, Operation, or Product in Guilford's terms. Unfortunately, according to Meeker's system a single item may be scored in several cells of the Guilford model, and subscores for each of these aspects of the intellect will be correlated to some degree among themselves and with the total IQ score. Although far from perfect, it would be possible to examine whether the test changes its relative emphasis over age and whether the pattern of IQ change for any one of the five IQ clusters is produced by some Guilford item sets more than by others.

To examine whether the test changes dramatically in its content emphasis according to the Guilford model, the accumulated totals of the Guilford-Meeker expected scores at each age were plotted as a percentage of the total expected score for each of the three facets of the Guilford cube. This was done separately for Forms L, M, and LM. For example, the accumulated maximum expected score for the "figural," "symbolic," and "semantic" aspects of the Contents facet of the structure of the intellect model were plotted over age as a percentage of the total Contents contribution to the IQ score. The same was also done for the several aspects of the Operations and Products facets. These graphs would indicate whether one aspect of Guilford's model changed its relative contribution to the total IQ score over the entire age range studied. The plots for Forms L and M were nearly identical, and the curves for Form LM were a good "average" of the two earlier forms; thus the LM curves will be discussed.

These plots of the relative contributions of various aspects of IQ according to Guilford and applied to the Binet by Meeker show considerable inflection and shifts in relative position through the fourth-year level. However, thereafter, the relative contribution of the various Guilford aspects is very constant over age. Specifically, with respect to Contents, the test is from 50% to 60% semantic, 26% to 40% figural, and 0% to 15% symbolic. These relative contributions are fairly steady throughout the entire age range and never change or overlap in their relative position. The six products are more evenly distributed in their relative contribution to the Binet score and show some shifting in relative position through age 4. The test is predominantly influenced by units, systems, and relations, and least representative of classes and transformations. Implication shows a sharp rise in influence from 0% to 15% between the ages of 3 and 5. After that, the relative contributions of the six Products are quite stable. The five Operations are relatively constant in their contribution after age 4 when cognition contributes 35% to 40%, covergent production 25% to 32%, memory and evaluation 12% to 20%, and divergent production 0% to 6%. Major shifts in relative emphasis before age 4 exist for convergent production which declines dramatically and evaluation which increases during these early age levels.

Consequently, after age 4 the Binet appears to be quite stable in the relative contribution of aspects of the three facets of Guilford's structure of the intellect. Moreover, since the Fels children scored above average, shifts in IQ deriving from changes in content emphasis on the test should be confined to the ages prior to approximately 3½. However, the major points of inflection in the observed IQ patterns were at ages 6 and 10, developmental levels that are well within the age range of stability of relative item contribution. Thus, given this analysis, it would appear that the inflections in IQ pattern described previously are not obviously simple products of changes in Binet item content and emphasis.

Homogeneity of Guilford Facets within IQ Patterns

A second analysis was performed to determine whether each of the 14 Guilford subscores plus the vocabulary score were homogeneous within each cluster in following the general IQ pattern for that group. Consequently, these 15 scores were standardized over the 80 subjects within each age, interpolated-extrapolated within each subject over age to fill missing points, and then the mean over subjects within a cluster plotted as a function of age separately for each cluster. With only one exception the developmental pattern for each of the 15 scores per IQ group were remarkably homogeneous and reflected the general IQ pattern characteristic of that group. The plots for each group were even more homogeneous with respect to pattern than the graphs depicted in figure 4. These data do not contradict the suggestion that the IQ plots in figure 5 are characteristic of a variety of mental functions tested by the Binet and in the Guilford-Meeker subscales. However, it must be acknowledged that the subscales are not mutually independent nor independent of the total IQ score. Nevertheless, the IQ patterns presented in figures 4 and 5 are not obviously functions of the changing nature of the test, or more on one facet of mental performance than another as reflected on the Binet at some or all ages.

There was one exception to the above conclusion, and this was for the vocabulary score. For Cluster 1, the vocabulary score was nearly as constant and as integrated with the other trends as any single Guilford-Meeker subscale pattern. However, for the four groups showing marked changes in IQ the vocabulary score ran somewhat counter to the IQ trend after age 6 and particularly between ages 7 and 11. Thus, children in Clusters 2 and 3 displayed relatively higher vocabulary scores than expected on the basis of their IQ trends and Clusters 4 and 5 had relatively lower vocabulary scores for their IQs. This was especially marked for Cluster 4 which showed the highest IQs at age 10 but whose vocabulary scores were below the sample's mean at ages 8, 9, and 10. Although their vocabulary scores were lower than their relative performance on the Guilford sums during the preschool period, they were rather high after age 11.

The fact that the trend for vocabulary ran counter to the general IQ trend is opposite to what one might expect given the parental information on IQ pattern. That is, one might anticipate a relatively high performance for Clusters 4 and 5 on vocabulary, both because of their high IQ scores in middle childhood and also because of an expected emphasis on verbal skills by their better-educated and higher-IQ parents. Just the reverse might have been expected for Clusters 2 and 3. In point of fact, the vocabulary score was more homogeneous from group to group and ran opposite to these speculations.

These data are certainly not conclusive with respect to the contribution of individual mental skills to these patterns of IQ change, since the Guilford-Meeker components are not mutually independent. Nevertheless they fail to provide support for arguments suggesting that the IQ patterns observed above are simple epiphenomena of the changing nature of the IQ test and have very little to do with possible changes in the subject's mental skills and performances.

DISCUSSION

The data presented in this and the previous chapter indicate that normal middle-class children display different patterns of IQ change over age and that there are global parental behavioral differences associated with these various developmental profiles.

Amount of change.—These data and previous research indicate that normal children do change in mental performance. For example, if filled data points are disregarded and a simple range of observed IQ scores is taken, the average child in this sample shifted 28.5 IQ points between 2½ and 17 years, which is 1.78 standard deviation units. While the individual ranges of 21% of the sample were less than 20 points and an additional 43% shifted from 21 to 30 points, more than one of every three children (36%) displayed performance jumps of more than 30 points and 14% showed IQ trends covering more than 40 points. One child increased 74 points, or 4.63 standard deviation units. This case plus Moore's (1967) report of a child who moved 73 IQ points represent the largest individual performance shifts among essentially normal children known to the authors.

The patterns of IQ change developed in Chapter IV indicate that, while 45% of the sample evidenced a relatively constant pattern over age, the remaining subjects displayed marked changes in IQ, and these shifts were not simply random fluctuations about a constant value. Rather, the IQ trends over age could be clustered into a few groups that represented relatively simple developmental trends—linear, quadratic, and cubic functions—although the statistical analyses emphasized gross trends, and larger samples might disclose subgroups having more complex profiles. Moreover, the overlapping nature of the Binet will insure that shifts in performance are relatively gradual; in fact, sudden marked fluctuations might be viewed as reflecting motivational and situational factors rather than valid changes in mental ability.

It should be pointed out that the cross-age correlation matrix for these data (e.g., see Sontag et al. 1958) was very comparable to that reported by others (e.g., Pinneau 1961). For example, the median correlation was .90 for adjacent-age assessments (6-month to 2-year interval between testings), .81 for intervals of 2–6 years, and .73 for 3½–9-year spans. Correlations from the youngest to the oldest ages were consistently in the .40s and .50s, and beginning at 4½–6 years the correlations were above .70 with IQ at the oldest ages. Viewed by themselves, these data apparently argue for high stability and constancy of IQ throughout the childhood years, especially after age 6.

However, a major methodological implication of these data is that relatively high correlational stability over age does not imply that such a trait is developmentally constant or that substantial numbers of subjects do not display sizable and meaningful shifts in such behavior (unless the correlations are 1.00). This situation can exist because a correlation is independent of the means of the distributions involved and it expresses the similarity of an individual's position in the group relative to the variability across individuals within those distributions. Thus, even when cross-age correlations are extremely high (and the correlational stability of IQ is among the highest of behaviors in the developmental literature), the investigation of patterns of change with development is still a worthwhile pursuit.

Inflection points.—The patterns described indicate major inflection points at approximately 6 and 10 years (and perhaps also at 14). Undoubtedly, some of these reversals are at least enhanced by psychometric factors. For example, the variability of the Binet is lowest in the 5–7-year period, and thus minor variations in performance will be amplified by applying Pinneau's correction which equates the variability at different ages. Moreover, some have claimed that there is a dearth of items of intermediate difficulty in middle childhood (McNemar 1942) which might also distort trends in the preschool period and at 6 and 10 years. Further, some of the leveling off of profiles after ages 6 and 10 may reflect ceiling or floor effects. Finally, the Tucker procedure, used as an intermediate step before clustering subjects, would tend to emphasize shifts in the middle ages (ages 6–7, since this procedure did not take into account the unequal spacing of tests). In the face of these psychometric and statistical considerations, one might conclude that the profiles developed in this paper are so unique to these procedures and to the Binet test that they have nearly no meaning or generality.

On the other hand, whatever statistical procedures are employed, they do not alter the raw data pictured in figures 4 and 5. While different statistical techniques might return clusters with somewhat different general patterns, there is no escaping the fact that groups of subjects can be observed (fig. 4) for whom the general trends in figure 5 are faithful reflections of their actual performance. Moreover, the clusters of IQ change were associated with global assessments of parental behavior which could be sensibly interpreted and were reasonably consistent with other literature. For example, one of the two parental behaviors found to distinguish between increasing versus decreasing preschool trends was the accelerational attempt of parents for intellectual achievements. Freeberg and Payne (1967), in a review of parental correlates of single-assessment mental performance, came to a similar conclusion: "Children of superior intellectual ability come from homes where parental interest in their intellectual development is evidenced by pressures to succeed and assistance in doing so, particularly in the development of the child's verbal skills" (p. 71). Thus, the existence of relatively interpretable parental correlates of IQ change which are consistent with other literature argues for the potential meaningfulness of the IQ profiles developed here.

Although the timing of inflections in IQ patterns could be influenced by psychometric factors, they are not without interpretation and empirical precedent. Age 6 is the usual beginning of school, and throughout history societies have marked the 5–7 age period as a major transition in the child's mental and interpersonal development (White 1965). By the time the child reaches 10 years, the intellectual excitement of learning basic reading and mathematical skills may have dissipated. On Freudian grounds, the 6- and 10-year inflections roughly coincide with the beginning and end of the latency period.

The most compelling empirical evidence focuses on the 6-year point. This age marks the transition from parental dominance of mental development to the beginning of the influence of school, and shifts in a variety of mental performances have been catalogued for the 5–7-year period. White (1965) has summarized these transitions and hypothesizes that prior to this point the child's learning behavior has a strong associative character but that sometime between 5 and 7 years of age a more cognitive orientation may overlay this early foundation. Of course, it is not possible to determine whether some environmental circumstance (e.g., school) produces these changes in learning and mental performance or whether maturational transitions govern the time when a child is ready for new educational experiences. Nevertheless, a major inflection in IQ performance at approximately age 6 appears quite reasonable.

Regardless of the causes of these inflections, their apparent existence has implications for the practical application of IQ tests. One of the contemporary issues surrounding the use of the IQ test as a selection tool is that too much confidence is often placed in the subsequent stability of performance. A child may become locked into a given educational level or program on the basis of a single assessment early in childhood. The fact that age 6 signifies a potentially important inflection in performance suggests great caution in using the Binet as a one-shot screening technique in the 4–6-year range. Since inflections appear to occur approximately every 4 years, placement decisions should be reviewed frequently during childhood.

Generality.—Given these data, there is no adequate method of determining whether shifts in Binet performance over age actually reflect changes in mental ability or changes in the sensitivity of the test to different types of mental performance. What meager and imperfect data exist on the Binet during childhood and the Wechsler tests during early adulthood suggest that the general trends observed over age are not complete artifacts of averaging over many different mental abilities, each of which possesses a unique developmental pattern. Moreover, it does not appear that the Binet changes its relative content (within broad definition) after the 4-year level. However, while this implies that the developmental trends described herein may not be artifacts of changing test content, it does not imply that the observed patterns will be characteristic of other mental abilities which are not strongly represented in the Binet item pool. Within Guilford's model, the Binet is most strongly influenced by cognitive and convergent operations on semantic content with units, systems, and relations as products. It is least reflective of divergent operations on symbolic content with classes and transformations as products. Thus, although some psychologists have been prone to regard IQ as representative of all mental performance potential, such an inference is not called for by previous data or those presented in this *Monograph*. Indeed, one heuristic value of the present results is to encourage research on the generality and qualifications of such patterns with respect to specific mental abilities, sample characteristics, etc. While the Binet does not appear to represent a polyglot of independent abilities, and the IQ trends presented here are not likely to be artifacts of the emergence and decline of the influence of different abilities on test performance over age, there are probably other skills which are not well represented on the Binet that may have contrasting patterns of developmental change (see literature review in section on Specific Abilities, pp. 10–12).

Parent behavior.—It cannot be stressed too strongly that the parental behavior correlates found for these IQ patterns are just that—correlates. Thus, causality cannot be inferred from their presence. Third factors may operate which are related to both parental behavior and child IQ pattern, and/or the direction of causality might be from child to parent. The latter is quite possible since the parental variables are weighted more heavily with behavior occurring during late childhood, while the major differences in IQ pattern transpired earlier in development.

However, the existence of the correlated parent behavior is nevertheless an empirical fact that begs interpretation, albeit circumspect. Children who evidenced early increases in IQ and had the highest level of performance had parents who displayed the most attempts to accelerate their children in intellectual tasks. The opposite was true for parents of children who declined in IQ during the preschool period and who had the lowest levels of performance in this sample. Moreover, parents who accelerated their children least also employed either the most or the least severe penalties. Among subjects who had declining patterns and lower levels of IQ performance, children whose parents were most severe in their penalties had the most depressed IQ record, while children whose parents were most lax in their penalties showed some recovery in IQ during the middle school years. In contrast, the accelerating parents seemed to adopt a middle-of-the-road policy on the severity of discipline. Thus, the environment associated with optimum IQ profiles seemed to be one in which the parent encouraged and attempted to accelerate intellectual behavior, but in a context of moderate structure and discipline.

Unfortunately, the pattern of IQ over age and the general level of IQ performance were not independent in these data. Consequently, it is not clear whether these parental behaviors are correlates of general level, IQ pattern over age, or both. Although there was overlap in general level across IQ pattern clusters and differences existed in parental behavior even after general level of child IQ was covaried, there is no conclusive method of isolating the correlates of IQ change per se from these data.

Because children in this sample were born between 1930 and 1938 but the parental assessments were not begun until 1938, the serial assessments of parental behavior had to be averaged over the entire available childhood period. Not only did this prevent looking at contemporary changes in parent and child behavior but it also meant that the averaged parental variables for some subjects were only representative of parental behavior assessed later in childhood since earlier assessments were not made. Interestingly, Sontag et al. (1958) found accelerational attempt and a rational approach to discipline to characterize the parents of IQ increasers during the elementary school period, but there were no parental correlates of IQ shift during the preschool period. Regrettably, even though some of the same data are involved in this report, procedural differences make direct comparisons difficult.

From a scientific standpoint, the interpretation of these IQ profiles should terminate here. However, the performance of Cluster 4, and to some extent Cluster 5, invites speculation. Cluster 4 (fig. 5) displayed the highest levels of IQ performance of any group at every age and showed an inverted-*U* profile with the peak at 10 years (IQ at 10 years = 159). The children of Cluster 5 showed a similar rise during the preschool period, but they did not score as high or decline as much as Cluster 4 after age 10. While the pattern of IQ over age was different for these two groups, the behavior of the parents of these children could not be differentiated. The parents of Cluster 4 had the highest ratings of adjustment, clarity of policies, accelerational attempt, protectiveness, education, and parental Otis IQ, and they were relatively rewarding and minimally coercive. The pattern was similar but less extreme for Cluster 5. In short, these parents apparently provided the presumably ideal intellectual home. The interpretive issue is why these children, particularly those in Cluster 4, lost the impressive gains in IQ made during the preschool years?

A first consideration is that perhaps they reached a ceiling on the Binet. A ceiling effect is not an adequate explanation for two reasons. First, Cluster 4 did not plateau after reaching the apparent limit: rather it declined and by age 17 lost all of the impressive gains made between 2½ and 10 years. Second, while Cluster 5 did evidence a pattern which tended to level off after age 10 (though some decline was evident), this leveling off could not be due to a test ceiling since Cluster 4 was scoring higher by 22 points at the same age that Cluster 5 was at its peak.

Two of many possible speculations will be offered. Accelerating parents may be successful in influencing IQ when they have relatively pervasive control over the intellectual and social circumstances and experiences of their child. However, when school starts parents relinquish a great deal of control over educational activities. Their continued accelerational attempts coupled with less direct control over such development may translate into a "pushiness" that has the opposite motivational effect. The actual decrease in test performance may be delayed until age 10 partially because basic reading and mathematics is initially challenging and partially because the onset of serious strivings for independence from parental authority and values may not occur until somewhat later.

A second orientation is that following mastery of basic reading and mathematics these children may have found the public school of the 1930s and 1940s relatively uninteresting. Given their high IQ and the rewarding, personalized, and clear instructional experience provided by their parents, the public school environment may have appeared cold, impersonal, nonrewarding, restrictive, and regimented by comparison. As a result these children might have been relatively "turned-off" toward intellectual activities. If this analysis has validity, it would imply that the tacit axiom of "teaching to the median or below since the bright child will learn it anyway" may not always be true. Such an educational offering may constitute an "equalizer" for diverse groups of children but at the expense of failing to help the most capable to flower.

A synthesis of these views has been offered by Virginia Crandall (personal communication, 1972). The pushiness of highly accelerating parents probably results in greater skill acquisition during the preschool years because the child is still in what Veroff (1969) calls the "autonomous achievement period," in which he simply compares his achievements and strivings toward parental standards with his own past performance. Later, when basic intellectual achievements are performed in the social context of the school, the child takes on some values and standards of his classmates and evaluates his achievement performance against that of his peers. Parental acceleration and pushiness at this time can easily become threatening, frustrating, and disruptive. To the extent that he still wishes to please (and appease) his parents (which is likely, given the parents of children in Clusters 4 and 5), his inability to meet parental standards under these pressuring circumstances, plus the relatively uninteresting school environment, might make dropping out of the academic race the easiest solution. Of course, not all such parents are disruptively "pushy" and not all high-IQ children lose their status (witness Terman's geniuses).

VI. SUMMARY AND IMPLICATIONS

Summary

When the results reported above are integrated with the previous literature, the following trends emerge.

1. Normal home-reared middle-class children change in IQ performance during childhood, some a substantial amount. In the present sample, the average individual's range of IQ between 2½ and 17 years of age was 28.5 IQ points, one of every three children displayed a progressive change of more than 30 points, and one in seven shifted more than 40 points. Rare individuals may alter their performance as much as 74 points. High-IQ children are likely to show greater amounts of change than low-IQ children. Shifts in IQ are not random fluctuations about a constant value but represent relatively progressive, simple linear, quadratic, and cubic trends over childhood. There is little evidence that these profiles are straightforward products of repeated testing.

2. There is some indication (though not emphatic in the present data) that boys are more likely to show increases in IQ over age than girls and that girls who are relatively more favorably disposed to traditional masculine roles tend to increase in IQ more than girls who are less so. Since these data were collected between 1930 and 1960, contemporary changes in the female role in society may modify this conclusion.

3. While most samples studied showed increases in IQ over age, not all subjects nor groups of subjects follow that general trend. Specifically, the predominant pattern is one of no change or decline in IQ over age for low-income and culturally isolated groups.

4. Changes in performance on standardized mental tests continue throughout life, and the nature of these developmental patterns may be different for different abilities.

5. People who show increases in IQ over age differ in personality from those who show decreases, and these personality correlates depend in part upon age and sex. Preschool children who gain in IQ are described as independent and competitive in a social context. Elementary grade children who

gain in IQ are independent, scholastically competitive, self-initiating, and problem solving, while IQ increases in adulthood are related to the characteristics of interpersonal distance, coldness, and introversion. These shifts may reflect the changing sources of educational experiences and motivation for intellectual achievement beginning with the family, then the competition with peers at school, and finally the self-education and intrinsic motivations that characterize maturity.

6. Parents of children who show gains in IQ provide their children with acceleration and encouragement for intellectual tasks and take a moderate, rationally structured approach to discipline.

7. When the specific nature of the pattern of IQ change over age is considered, there appear to be major inflection points at ages 6 and 10 (and possibly 14). The 6-year point coincides with major transitions in a variety of learning behaviors and may represent a shift from a relatively associative learning strategy to a more cognitive orientation.

8. Related individuals are more similar in the general level of their performance on infant tests as well as on childhood IQ assessments. However, while there may be some degree of similarity among related individuals for profile contour during the first year of life, there is little evidence that the pattern of IQ over age thereafter possesses heritability.

9. The degree of within-pair similarity in IQ profile is considerably greater for DZ twins than for sibling pairs during infancy and childhood, yet there is the same amount of genetic overlap within pairs. This difference presumably derives from the fact that environmental events (e.g., home atmosphere, specific intellectual determinants, testing procedures, etc.) occur at the same age for twins but at different ages for siblings.

10. Correlations between parent and child IQ when parent and child assessments were made at the same chronological age are somewhat lower than parent-child correlations when the parent is assessed as an adult. There is no parent-child correlation of either kind for these test scores before the child is 2–3 years. While adult-parent and child relationships are relatively stable after ages 3–4, there may be some increase over age in the size of same-age parent-child correlations. The correlations for same-sexed parent-child pairs tend to be slightly higher than for cross-sex pairs.

11. Multiple births not only score somewhat lower on standardized mental tests but they also have less variability as a group within an age and less intraindividual variability over age than singletons. This apparently is true during both infancy and childhood, and may limit generalizations from twin to singleton populations, especially when profile contour is considered.

12. The general level of IQ and the specific profile contour of IQ over age are probably not independent of one another, and this may characterize twins more than singletons. Some check on this possible lack of independence is necessary to assert within-pair profile similarity apart from similarity in general level.

Implications

The discussions at the ends of previous chapters have considered the present data in light of relevant focused and technical topics. The following is a brief consideration of the implications of the literature review and present data for certain broader issues.

The nature of "intelligence."—The authors believe that much confusion, misunderstanding, and needless rancor has derived from failing to distinguish between the concept of intelligence and the measurements we have come to call tests of intelligence. The cause is not furthered by denoting the concept to be synonymous with the test score, since there would then be as many concepts of intelligence as tests (Guilford 1967).

Although most serious students of mental performance have argued against a pervasive, unitary, and constant conception of intelligence, the term "intelligence" has certainly come to have that connotation among the laity as well as many professionals. Sometimes individuals will deny allegiance to a unitary and constant intelligence but vehemently protest that IQ tests are biased indicators of intelligence and are unfair to certain subgroups in society, childhood IQ cannot be predicted from infancy because the infant is incapable of overtly demonstrating his intellectual potential, or that brain-wave tests of intelligence hold great promise because they are truly culture free. While there is a certain amount of validity to these assertions, often they reflect a tacit belief in that insidious unitary and constant intelligence and a blurring of the distinction between intelligence and IQ.

The Binet test was developed for a highly specific applied purpose, and its principal inventor did not bother to define the term "intelligence." It was later that Wilhelm Stern devised the IQ score, and still later that conferences were held to attempt to define intelligence. With the increasing potency of the Binet score to predict scholastic and occupational success, the unitary, pervasive, and constant connotations evolved over the protests of many psychometricians and others.

Even if IQ scores are accepted as a reflection of intelligence, the data are hardly consistent with such a concept. The failure to predict childhood IQ from infancy (McCall et al. 1972) plus the present data on the extent of systematic IQ change during childhood argue against constancy. As early as the turn of the century people found that performance on some mental tests was independent of performance on other tests (Guilford 1967), and different mental skills have evidenced contrasting developmental patterns during adulthood (Bayley 1968a). Such data should have quieted claims of the unity of intelligence. Despite the fact that some of this discordant evidence was available from the beginning of the modern testing movement, a constant and unitary intelligence was literally reified after the Binet test was constructed despite Spearman's (1927) reiteration of John Stuart Mill's warning that "the tendency has always been strong to believe that whatever receives a name must be an entity of being, having an independent existence of its own. And if no real entity answering to the name could be found, men did not for that reason suppose none existed, but imagined that it was something peculiarly abstruse and mysterious" (p. 14).

The authors interpret the results reported in this *Monograph* as applying to IQ scores, not to the colloquially defined concept of intelligence. The authors view IQ as a sampling of behavior that lives or dies as a function of its practical utility in correlating with other behaviors for a sizable segment of the population. Deshackled from its implications for "intelligence," these scores do predict educational and occupational criteria rather well and apparently do so for majority and minority groups alike (Kennedy 1969; Kennedy et al. 1963). The tests are biased against certain groups in society, but such bias also resides in the criteria.

Separating IQ tests from the issue of intelligence does not condone all applications of IQ tests for prediction or selection. At the very least, there needs to be some empirical demonstration that the tests indeed do correlate with the specific criterion of interest. Moreover, even if such correlations are high, the tests could be unfairly employed to eliminate people from being selected who could *learn* the criterion skills or behavior if given a chance. Finally, performance is variable within an individual, and testing should be done periodically since intraindividual changes can be substantial and major inflections in developmental profiles occur at approximately 6 and 10 years of age.

Heritability and IQ.—There is nothing about the assertion that IQ has a heritability of .80 that implies anything about the changeability of IQ. Even the most celebrated proponents of a genetic emphasis warn against making this inference (e.g., Jensen 1969a). The IQ heritability as usually calculated reflects the percentage of variability in IQ that can be attributed to differences in genotypes within the sample investigated. Therefore, contemporary or future environments not represented in the sample could have a profound impact upon IQ, and many examples of such cases can be cited. The undesirable attributes of phenylketonuria, a monogenetic trait which produces mental retardation if left untreated (i.e., heritability = 1.00), can be eliminated by the introduction of a special diet in infancy. The average height of people in the United States and Japan has increased over the last century by several inches despite a heritability of approximately .90. Tuberculosis once had high heritability, but now the bacillus is so rare that contracting T.B. depends less upon one's genetic susceptibility than one's likelihood of exposure. The heritability is now low. Therefore, a heritability of .80 for IQ in the broad sense of the term says nothing about our future ability to influence such performance.

However, heritability does indicate the relative impact current differences in the environments represented in the sample are having on IQ. It is not the purpose here to evaluate the tenability of that estimate, although it should be noted that through all the furor over Jensen's (1969a) controversial monograph there is almost no criticism of this value for middle-class, white samples. Two implications can be drawn from this estimate.

First, allowing the heritability of .80, the phenotypic variability (i.e., standard deviation) of IQ for a single genotype has been estimated to be between 6 and 9 IQ points. This means that, if a group of individuals all somehow had genotypes for an IQ of 100, the standard deviation of the distribution of their scores would be between 6 and 9 points (Jensen 1969b). Thus, the range of scores would be from below 80 to above 120 as a function of differences in between-family environmental circumstances.

Second, almost all of the heritability information on IQ as well as the basis for the above analysis derives from single-age assessments of IQ. As indicated in Chapter I, such data provide very little information about developmental changes. Moreover, it is conceptually important to notice that the environmental factors that transpire during an individual's childhood which influence IQ may be quite different in nature than those that contribute to between-family IQ differences assessed at a single age. Environmentalists are quick to point out the massive differences between the family environments of various segments of society in the form of the availability of books, experience with contingent interpersonal behavior, encouragement and expectancy of intellectual success, educational opportunity, etc. However, many of these factors are relatively constant (at least in a correlational sense) throughout the child's life and could not account in a straightforward manner for intraindividual variability in performance with development. Therefore, not only is the amount of change in IQ over age potentially independent of general level of performance but the environmental factors influencing developmental variation may be different than the between-family circumstances that operate on general level.

Intraindividual variability and environment.—How much do individuals change during their childhood years? The average intraindividual developmental standard deviation for the sibling sample between 3 and 12 years of age was 0.47 standardized IQ units. Since the average standard deviation for the IQ tests represented in this sample between these ages was 16.5, the intraindividual standard deviation is roughly comparable to 7.8 IQ points.

However, this may be an underestimate of the actual value. The IQ scores were standardized within each test within each age. Therefore, all group age trends were eliminated prior to calculating a subject's developmental variability, and other analyses presented in Chapter IV demonstrated that group trends did exist.

The important point is that an individual subject possesses a constant genotype throughout development; yet he changes almost as much in IQ during his childhood as different people vary in IQ at a single age as a function of contrasting between-family environments.

Unfortunately, change in IQ over age may reflect the influence of one or more factors: (1) genetically based developmental trends, (2) the emergence with development of greater genetic control over IQ, (3) change in what is being measured by the IQ test, (4) random temporary environmental fluctuations and test error, (5) constant between-family environmental differences, and (6) idiosyncratic interactions between environmental events and the skills and motivational dispositions of the child at a particular time.

First, if the developmental trends reported in this *Monograph* were obviously gene based, one would have expected siblings to have been more similar in their IQ profiles than unrelated children. This was not the case.

Second, if the patterns of change in IQ over age reflected the gradually increasing influence of a genetic factor, one would expect some degree of sibling similarity in profile (which was not present) and a corresponding increase in the correlation between siblings over age. While there is certainly literature that suggests a gradually increasing genetic effect across childhood (e.g., Honzik 1957), the correlations among siblings for this particular sample do not show an increasing trend over age (McCall 1970a). Moreover, if the increasing genetic influence manifested itself in a general group trend in IQ, such a manifestation would have been eliminated by the use of standardized scores. On balance, this interpretation does not gain clear support for this sample.

Third, the evidence is less decisive with respect to possible changes with age in what is being measured by IQ tests. The Guilford-Meeker breakdown of items on the Binet implies that the relative contribution of Guilford's facets to the IQ score is constant after the fourth-year level (approximate chronological age 3½ for this sample), and IQ patterns were not obviously a function of one or another of the Guilford-Meeker attributes. Unfortunately, these subtest scores were not mutually independent, and the definition of any single attribute might differ from one age to another in systematic ways. Thus, while the available data do not support this possibility, neither are they decisive.

Fourth, it is unlikely that the observed trends are simply products of temporary environmental fluctuations or test error and unreliability since the developmental patterns described in Chapter IV were relatively simple, consistent, gradual trends and not random variations about a constant value.

Fifth, changes in IQ are not likely to be strongly determined by constant environmental characteristics of the home (e.g., opportunity for education, reward for intellectual pursuits, etc.) as is often emphasized. If this were the case, the general intellectual climate of the home should have been shared by siblings, and they should have been similar in their pattern of IQ change. This was not the case. On the other hand, IQ change groups were differentiated by the general parental attributes of accelerational attempt and severity of penalties. Unfortunately, it is not clear whether these correlates apply to the general level or the developmental profile of IQ. Thus, while such constant between-family differences undoubtedly contribute to IQ, their impact on change in IQ may be much less than supposed.

This leaves the idiosyncratic interaction of specific environmental events with the skills and motivational dispositions of the child at the time of the environmental circumstance as a possible major contributor to IQ change. The estimates from other studies of a 6–9-point environmental variation include both between-family and individual developmental variation. If the average individual developmental variation is 7.8 points and much of it derives from idiosyncratic interactions between environmental events and the skills and interests of the child at that particular time, the current emphasis placed on gross between-family differences may have to give way to more complicated and more idiosyncratic environmental dynamics. Examples of such circumstances might be: a teacher whose special interest catches a pupil "at the right time" and fosters an enduring interest in a particular intellectual pursuit, or a youngster whose older brother becomes an athletic star devotes his energies to scholarly activities as a means of independent self-fulfillment and parental approval. These environmental events might have had different effects if they happened at another time or to another child.

If such specific, individual dynamics characterize a significant portion of the environmental contribution to IQ, then enrichment programs and public schools might attempt to develop curricula and learning processes that are not only sensitive to but capitalize on the individual interests, motivations, and skills of each child. Moreover, although only a few enrichment programs have been successful at raising the IQs of impoverished children substantially past 100 (see Heber 1969), the present data suggest that considerable change is possible above 100, and the factors that govern the achievement of superior performance may be even more idiosyncratic than those determining shifts at other levels of the scale. At the very least, it is time to permit our theoretical orientations and our practical efforts at providing "stimulating" environments to reflect the complexity of the phenomenon.

REFERENCES

Anastasi, A., & Foley, J. P., Jr. *Differential psychology: individual and group differences in behavior.* (2d ed.) New York: Macmillan, 1949.

Asher, E. J. The inadequacy of current intelligence tests for testing Kentucky mountain children. *Journal of Genetic Psychology,* 1935, **46**, 480–486.

Bajema, C. J. Relation of fertility to occupational status, IQ, educational attainment, and size of family of origin: a follow-up study of a male Kalamazoo public school population. *Eugenics Quarterly,* 1968, **15**, 198–203.

Baker, C. T.; Sontag, L. W.; & Nelson, V. L. Specific ability in I.Q. change. *Journal of Consulting Psychology,* 1955, **19**, 307–310.

Baldwin, A. L. The relative difficulty of Stanford-Binet items and their relationship to I.Q. *Journal of Personality,* 1948, **16**, 417–430.

Baldwin, A. L.; Kalhorn, J.; & Breese, F. H. The appraisal of parent behavior. *Psychological Monographs,* 1949, **63**(4, Whole No. 299).

Baldwin, B. T., & Stecher, L. I. Additional data from consecutive Stanford-Binet tests. *Journal of Educational Psychology,* 1922, **13**, 556–560.

Bayley, N. Mental growth in young children. In G. M. Whipple (Ed.), *Thirty-ninth Yearbook, National Society for the Study of Education. Part II. Intelligence: its nature and nurture.* Bloomington, Ill.: Public School Publishing Co., 1940. Pp. 11–47.

Bayley, N. Consistency and variability in the growth of intelligence from birth to eighteen years. *Journal of Genetic Psychology,* 1949, **75**, 165–196.

Bayley, N. On the growth of intelligence. *American Psychologist,* 1955, **10**, 805–818.

Bayley, N. Data on the growth of intelligence between 16 and 21 years as measured by the Wechsler-Bellevue Scale. *Journal of Genetic Psychology,* 1957, **90**, 3–15.

Bayley, N. Comparisons of mental and motor test scores for ages 1–15 months by sex, birth order, race, geographical location, and education of parents. *Child Development,* 1965, **36**, 379–411.

Bayley, N. Learning in adulthood: the role of intelligence. In H. J. Klausmeier and C. W. Harris (Eds.), *Analyses of concept learning.* New York: Academic Press, 1966. Pp. 117–138.

Bayley, N. Behavioral correlates of mental growth: birth to thirty-six years. *American Psychologist,* 1968, **23**, 1–17. (a)

Bayley, N. Cognition in aging. In K. W. Schaie (Ed.), *Theory and methods of research on aging.* Morgantown: West Virginia University Library, 1968. Pp. 97–119. (b)

Bayley, N. *Bayley scales of infant development.* New York: Psychological Corp., 1969.

Bayley, N. Development of mental abilities. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology.* Vol. 1. New York: Wiley, 1970. Pp. 1163–1209.

Bayley, N., & Jones, H. E. Environmental correlates of mental and motor development: a cumulative study from infancy to six years. *Child Development,* 1937, **8**, 329–341.

Bayley, N., & Oden, M. H. The maintenance of intellectual ability in gifted adults. *Journal of Gerontology,* 1955, **10**, 91–107.

Bayley, N., & Schaefer, E. S. Correlations of maternal and child behaviors with the development of mental abilities: data from the Berkeley Growth Study. *Monographs of the Society for Research in Child Development,* 1964, **29** (6, Serial No. 97).

Bell, R. Q.; Weller, G. M.; & Waldrop, M. F. Newborn and preschooler: organization of behavior and relations between periods. *Monographs of the Society for Research in Child Development,* 1971, **36**(1–2, Serial No. 142).

Birren, J. E. Psychological aspects of aging: intellectual functioning. *Gerontologist,* 1968, **8**, 16–19.

Bloom, B. S. *Stability and change in human characteristics.* New York: Wiley, 1964.

Bock, R. D. Multivariate analysis of variance of repeated measurements. In C. W. Harris (Ed.), *Problems of measuring change.* Madison: University of Wisconsin Press, 1963. Pp. 85–103.

Bock, R. D., & Vandenberg, S. G. Components of heritable variation in mental test scores. In S. G. Vandenberg (Ed.), *Progress in human behavior genetics.* Baltimore: Johns Hopkins Press, 1968. Pp. 233–260.

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. *Annals of Mathematical Statistics,* 1954, **25**, 484–498.

Burt, C. The genetic determination of differences in intelligence: a study of monozygotic twins reared together and apart. *British Journal of Psychology,* 1966, **57**(1–2), 137–153.

Cattell, P. Constant changes in the Stanford-Binet I.Q. *Journal of Educational Psychology,* 1931, **22**, 544–550.

Cattell, R. B. Theory of fluid and crystallized intelligence: a critical experiment. *Journal of Educational Psychology,* 1963, **54**, 1–22.

Champney, H. The measurement of parent behavior. *Child Development,* 1941, **12**, 131–166.

Conrad, H. S.; Jones, H. E.; & Hsiao, H. H. Sex differences in mental growth and decline. *Journal of Educational Psychology,* 1933, **24,** 161–169.

Crandall, V. C., & Battle, E. S. The antecedents and adult correlates of academic and intellectual achievement effort. In J. P. Hill (Ed.), *Minnesota Symposium on Child Psychology.* Vol. **4.** Minneapolis: University of Minnesota Press, 1970. Pp. 36–93.

Crandall, V. J., & Preston, A. Patterns and levels of maternal behavior. *Child Development,* 1955, **26,** 267–277.

Cronbach, L. J. Proposals leading to analytic treatment of social perception scores. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior.* Stanford, Calif.: Stanford University Press, 1958. Pp. 353–379.

Cronbach, L. J. Year-to-year correlations of mental tests: a review of the Hofstaetter analysis. *Child Development,* 1967, **38,** 283–290.

Cronbach, L. J., & Gleser, G. C. Assessing similarity between profiles. *Psychological Bulletin,* 1953, **50,** 456–473.

Davidson, M. L. Univariate versus multivariate tests in repeated-measures experiments. *Psychological Bulletin,* 1972, **77,** 446–452.

Dearborn, W. F., & Rothney, J. *Predicting the child's development.* Cambridge, Mass.: Science-Art Publications, 1941.

Deutsch, M.; Katz, I.; & Jensen, A. R. *Social class, race, and psychological development.* New York: Holt, Rinehart & Winston, 1968.

Ebert, E., & Simmons, K. The Brush Foundation study of child growth and development, I: Psychometric tests. *Monographs of the Society for Research in Child Development,* 1943, **8**(2, Serial No. 35).

Eichorn, D. H. Developmental parallels in the growth of parents and their children. Presidential address, Division 7, American Psychological Association, Washington, D.C., September 1969.

Erlenmeyer-Kimling, L., & Jarvik, L. F. Genetics and intelligence. *Science,* 1963, **142,** 1477–1479.

Freeberg, N. E., & Payne, D. T. Parental influence on cognitive development in early childhood: a review. *Child Development,* 1967, **38,** 65–88.

Freeman, F. N., & Flory, C. D. Growth in intellectual ability as measured by repeated tests. *Monographs of the Society for Research in Child Development,* 1937, **2**(2, Serial No. 9).

Geisser, S., & Greenhouse, S. W. An extension of Box's results on the use of the *F* distribution in multivariate analysis. *Annals of Mathematical Statistics,* 1958, **29,** 885–891.

Goodenough, R. L. The relation of the intelligence of preschool children to the education of their parents. *School and Society,* 1927, **26,** 54–56.

Gray, S. W., & Miller, J. O. Early experience in relation to cognitive development. *Review of Educational Research,* 1967, **37**(5), 475.

Guertin, W. H.; Rabin, A. I.; Frank, G. H.; & Ladd, C. E. Research with the Wechsler Intelligence Scales for Adults: 1955–1960. *Psychological Bulletin,* 1962, **59,** 1–26.

Guilford, J. P. *The nature of human intelligence.* New York: McGraw-Hill, 1967.

Härnqvist, K. Relative changes in intelligence from 13 to 18, II: Results. *Scandinavian Journal of Psychology,* 1968, **9**, 65–82.

Heber, R. *Rehabilitation of families at risk for mental retardation.* Madison: Regional Rehabilitation Center, University of Wisconsin, 1969.

Heilman, J. D. Sex differences in intellectual abilities. *Journal of Educational Psychology,* 1933, **24**, 47–62.

Hellmuth, J. (Ed.) *Disadvantaged child.* Vol. **2**. New York: Brunner/Mazel, 1968.

Herzberg, F., & Lapkin, M. A study of sex differences on the Primary Mental Abilities Test. *Educational Psychological Measurement,* 1954, **14**, 687–689.

Higgins, J. V.; Reed, E. W.; & Reed, S. C. Intelligence and family size: a paradox resolved. *Eugenics Quarterly,* 1962, **9**, 84–90.

Hilden, A. H. A longitudinal study of intellectual development. *Journal of Psychology,* 1949, **28**, 187–214.

Hindley, C. B. Social class influences on the development of ability in the first 5 years. *Proceedings of the 14th International Congress of Applied Psychology,* 1961, 29–41.

Hindley, C. B. Stability and change in abilities up to five years: group trends. *Journal of Child Psychology and Psychiatry,* 1965, **6**, 85–99.

Hirsch, N. D. M. An experimental study upon three hundred school children over a six-year period. *Genetic Psychology Monographs,* 1930, **7**(6), 487–544.

Hobson, J. R. Sex differences in primary mental abilities. *Journal of Educational Research,* 1947, **41**, 126–132.

Honzik, M. P. Developmental studies of parent-child resemblance in intelligence. *Child Development,* 1957, **28**, 215–228.

Honzik, M. P. A sex difference in the age of onset of the parent-child resemblance in intelligence. *Journal of Educational Psychology,* 1963, **54**, 231–237.

Honzik, M. P. Environmental correlates of mental growth: prediction from the family setting at 21 months. *Child Development,* 1967, **38**, 337–364. (a)

Honzik, M. P. Prediction of differential abilities at age 18 from the early family environment. Paper presented at the meeting of the American Psychological Association, Washington, D.C., September 1967. (b)

Honzik, M. P., & Macfarlane, J. W. Personality development and intellectual functioning from 21 months to 40 years. Paper presented at the meeting of the American Psychological Association, Miami, September 1970.

Honzik, M. P.; Macfarlane, J. W.; & Allen, L. The stability of mental test performance between two and eighteen years. *Journal of Experimental Education,* 1948, **18**, 309–324.

Horn, J. L. Organization of abilities and the development of intelligence. *Psychological Review,* 1968, **75**, 242–259.

Horn, J. L., & Cattell, R. B. Refinement and test of the theory of fluid and crystallized general intelligence. *Journal of Educational Psychology,* 1966, **57**, 253–270.

Hunt, J. McV. *Intelligence and experience.* New York: Ronald, 1961.

Hunt, J. McV. *The challenge of incompetence and poverty.* Urbana: University of Illinois Press, 1969.

Jensen, A. R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review,* 1969, **39,** 1–123. (a)

Jensen, A. R. Reducing the heredity-environment uncertainty: a reply. *Harvard Educational Review,* 1969, **39,** 449–483. (b)

Jones, H. E. The environment and mental development. In L. Carmichael (Ed.), *Manual of child psychology.* (2d ed.) New York: Wiley, 1954. Pp. 631–696.

Jones, H. E., & Conrad, H. S. The growth and decline of intelligence: a study of a homogeneous group between the ages of ten and sixty. *Genetic Psychology Monographs,* 1933, **13,** 223–298.

Kagan, J., & Moss, H. A. Parental correlates of child's IQ and height: a cross-validation of the Berkeley Growth Study results. *Child Development,* 1959, **30,** 325–332.

Kagan, J., & Moss, H. A. *Birth to maturity.* New York: Wiley, 1962.

Kennedy, W. A. A follow-up normative study of Negro intelligence and achievement. *Monographs of the Society for Research in Child Development,* 1969, **34**(2, Serial No. 126).

Kennedy, W. A.; Van de Riet, V.; & White, J. C., Jr. A normative sample of intelligence and achievement of Negro elementary school children in the southeastern United States. *Monographs of the Society for Research in Child Development,* 1963, **28**(6, Serial No. 90).

Littell, W. M. The Wechsler Intelligence Scale for Children: review of a decade of research. *Psychological Bulletin,* 1960, **57,** 132–156.

Lorr, M., & Jenkins, R. L. Three factors in parent behavior. *Journal of Consulting Psychology,* 1953, **17,** 306–308.

McCall, R. B. IQ pattern over age: comparisons among siblings and parent-child pairs. *Science,* 1970, **170,** 644–648. (a)

McCall, R. B. The use of multivariate procedures in developmental psychology. In P. Mussen (Ed.), *Carmichael's manual of child psychology.* New York: Wiley, 1970. Pp. 1366–1378. (b)

McCall, R. B. *Fundamental statistics for psychology.* New York: Harcourt, Brace & World, 1970. (c)

McCall, R. B. Discussion: relations of newborn physiology and preschool verbal communication to intelligence, maturity, and pace of play in the school-age period. Paper presented at the meeting of the American Psychological Association, Washington, D.C., September 1971.

McCall, R. B. Similarity in developmental profile among related pairs of human infants. *Science,* 1972, **178,** 1004–1005. (a)

McCall, R. B. Similarity in IQ profile among related pairs: infancy and childhood. *Proceedings of the American Psychological Association Meeting,* Honolulu, 1972, 79–80. (b)

McCall, R. B., & Appelbaum, M. Bias in the analysis of repeated measures designs: some alternative approaches. *Child Development,* 1973, **44,** 401–415.

McCall, R. B.; Hogarty, P. S.; & Hurlburt, N. Transitions in infant sensorimotor development and the prediction of childhood IQ. *American Psychologist,* 1972, **27,** 728–748.

McHugh, G. Changes in I.Q. at the public school kindergarten level. *Psychological Monographs*, 1943, **55**(2), 1–34.

McNemar, Q. *The revision of the Stanford-Binet Scale: an analysis of the standardization data.* Boston: Houghton-Mifflin, 1942.

Meeker, M. N. *The structure of intellect: its interpretation and uses.* Columbus, Ohio: Merrill, 1969.

Miles, C. C., & Miles, W. R. The correlation of intelligence scores and chronological age from early to late maturity. *American Journal of Psychology,* 1932, **44**, 44–78.

Moore, T. Language and intelligence: a longitudinal study of the first eight years, I: Patterns of development in boys and girls. *Human Development,* 1967, **10**, 88–106.

Moore, T. Language and intelligence: a longitudinal study of the first eight years, II: Environmental correlates of mental growth. *Human Development,* 1968, **11**, 1–24.

Moriarty, A. E. *Constancy and I.Q. change: a clinical view of relationships between tested intelligence and personality.* Springfield, Ill.: Thomas, 1966.

Morrison, D. F. *Multivariate statistical methods.* New York: McGraw-Hill, 1967.

Overall, J. E. Note on multivariate methods for profile analysis. *Psychological Bulletin,* 1964, **61**(3), 195–198.

Overall, J. E., & Klett, C. J. *Applied multivariate analysis.* New York: McGraw-Hill, 1972.

Owens, W. A. Age and mental abilities: a longitudinal study. *Genetic Psychology Monographs,* 1953, **48**, 3–54.

Pinneau, S. R. *Changes in intelligence quotient: infancy to maturity.* New York: Houghton-Mifflin, 1961.

Record, R. G.; McKeown, R.; & Edwards, J. H. An investigation of the difference in measured intelligence between twins and single births. *Annals of Human Genetics,* 1970, **34**, 11–20.

Reed, E. W., & Reed, S. C. *Mental retardation: a family study.* Philadelphia: Saunders, 1965.

Rees, A. H., & Palmer, F. H. Factors related to change in mental test performance. *Developmental Psychology Monograph,* 1970, 3(2).

Roberts, S. O.; Crump, E. P.; Dickerson, A. E.; & Horton, C. P. Longitudinal performance of Negro American children at 5 and 10 years on the Stanford-Binet. Paper presented at the meeting of the American Psychological Association, Chicago, September 1965.

Roff, M. A factorial study of the Fels Parent Behavior Scales. *Child Development,* 1949, **20**, 29–45.

Scarr-Salapatek, S. Unknowns in the IQ equation. *Science,* 1971, **174**, 1223–1228.

Schoenfeldt, L. F. Life history subgroups as moderators in the prediction of intellectual change. Unpublished manuscript, 1972.

Schoenfeldt, L. F., & Owens, W. A., Jr. Age and intellectual change: a cross-sectional view of longitudinal data. Paper presented at Midwestern Psychological Association, Chicago, April 1965.

Sherman, M., & Key, C. B. The "intelligence" of isolated mountain children. *Child Development,* 1932, 3, 279–290.

Skeels, H. M., & Fillmore, E. A. The mental development of children from underprivileged homes. *Journal of Genetic Psychology,* 1937, **50**, 427–439.

Skodak, M., & Skeels, H. M. A final follow-up study of 100 adopted children. *Journal of Genetic Psychology,* 1949, **75**, 8–125.

Sontag, L. W.; Baker, C. T.; & Nelson, V. L. Mental growth and personality development: a longitudinal study. *Monographs of the Society for Research in Child Development,* 1958, **23**(2, Serial No. 68).

Spearman, C. *The abilities of man.* New York: Macmillan, 1927.

Terman, L. M., & Merrill, M. *Measuring intelligence: a guide to the new revised Stanford-Binet test of intelligence.* Boston: Houghton Mifflin, 1937.

Thurstone, L. L. The differential growth of mental abilities. Psychometric Laboratory Report No. 14, University of North Carolina, Chapel Hill, 1955.

Tucker, L. R. Determination of generalized learning curves by factor analysis. Educational Testing Service Technical Report, 1960.

Tucker, L. R. Learning theory and multivariate experiment: illustration by determination of generalized learning curves. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology.* Chicago: Rand McNally, 1966. Pp. 476–501.

Veroff, J. Social comparison and the development of achievement motivation. In C. P. Smith (Ed.), *Achievement-related motives in children.* New York: Russell Sage, 1969. Pp. 46–101.

Wachs, R. D. Technical comment. *Science,* 1972, **178**, 1005.

Wechsler, D. *Measurement of adult intelligence.* (3d ed.) Baltimore: Williams & Wilkins, 1944.

Werner, E. E.; Honzik, M. P.; & Smith, R. S. Prediction of intelligence and achievement at ten years from twenty months pediatric and psychologic examinations. *Child Development,* 1968, 39, 1063–1075.

White, S. H. Evidence for a hierarchical arrangement of learning processes. In L. P. Lipsitt and C. C. Spiker (Eds.), *Advances in child development and behavior.* Vol. 2. New York: Academic Press, 1965. Pp. 187–220.

Wilson, R. S. Autonomic research with twins: methods of analysis. In S. G. Vandenberg (Ed.), *Progress in human behavior genetics.* Baltimore: Johns Hopkins Press, 1968. Pp. 287–302.

Wilson, R. S. Twins: early mental development. *Science,* 1972, **175**, 914–917. (a)

Wilson, R. S. Technical comment. *Science,* 1972, **178**, 1006–1007. (b)

Winer, B. J. *Statistical principles in experimental design.* (2d ed.) New York: McGraw-Hill, 1971.

ACKNOWLEDGMENTS

Portions of this research were supported by HD-04160 to R.B.M., by FR-05537, HD-00868, and FR-00222 to the Fels Research Institute, MH-10006 to the Psychometric Laboratory of the University of North Carolina, by the Grant Foundation of New York, and by the Fels Fund of Philadelphia. The data presented were available through the continuous effort over more than 40 years of the staff of the Fels Research Institute, then under the direction of Lester W. Sontag, and by the enduring faith of Mr. Samuel Fels and later the Fels Fund of Philadelphia. The labors of many people, professional and paraprofessional, who collected these data are gratefully acknowledged. Special appreciation is given to Virginia Nelson who administered more than 90% of the IQ tests used in this research.

This report would not be possible without the help of numerous individuals. Sontag served as a stimulus and consultant; and Richard Bell, Virginia Crandall, Dorothy Eichorn, Frank Falkner, Marjorie Honzik, Mary Meeker, and Alexander Roche generously contributed their time, knowledge, and facilities. Statistical advice was solicited from R. Darrell Bock, Elliot Cramer, Lloyd Humphreys, Arthur Jensen, John Loehlin, John Overall, and Ledyard Tucker. Several computer programs were written or modified by Hal Campbell, John Hawley, and Fred Lowe; and Darryl Bloom, Lynne Christensen, Nancy Hurlburt, and Katherine Pryor helped organize the data base. Finally, Crandall and Jerome Kagan criticized an early draft of the entire monograph; and Falkner, Jensen, and Loehlin commented on certain sections. The authors are gratefully indebted to these individuals who share in whatever virtues this *Monograph* possesses but who are free of any responsibility for its failings.