A Theory of Primary and Secondary Familial Mental Retardation

ARTHUR R. JENSEN

UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA

I. DIAGNOSIS AND TAXONOMY OF MENTAL RETARDATION

Recent evidence derived from experimental studies of learning in mentally retarded children and adults leads to a hypothesis of a hierarchy of mental abilities. The hypothesis has important implications for the taxonomy and diagnosis of mental retardation. This paper explicates the hypothesis and reviews some of the relevant experimental evidence. The implications of the hypothesis for the education of the retarded are also indicated.

A. Established Diagnostic Categories

Two broad categories of mental retardation are now generally recognized. The first category is diagnostically the most obvious; it is the variety of severe mental defects resulting in IQs for the most part below 50 and accompanied by physical abnormalities or clear signs of neurological damage. This category of mental deficiency forms a distribution of ability which, in a sense, stands apart from the normal distribution of mental abilities in the general population. Most of these severe defects appear to be due to (*a*) single mutant genes, often labeled "major gene" defects, (*b*) chromosomal defects, and (*c*) brain damage. Examples of *a* are recessive genetic defects such as phenylketonuria, galactosemia, amaurotic family idiocy, microcephaly, and hypertelorism, to name but a few. Examples of *b* are Down's syndrome (mongolism), due to triplication of chromosome 21, giving the child 47 rather than the normal 46 chromosomes; Kleinfelter's syndrome, due to an extra female sex chromosome in the male (XXY); and Turner's syndrome, a marked deficiency in spatial ability due to a missing sex chromosome in the female (XO instead of the normal XX). Examples of *c* are birth trauma, kernicterus due to prematurity or to rhesus incompatibility, and brain damaging diseases such as maternal rubella (German measles), neonatal septicemia, meningitis, and encephalitis.

The majority of persons with IQs below 50 are included in these diagnostic categories. Studies in England have found that among individuals in this severely subnormal range of IQ no specific causal factor was identifiable in about 30% of the cases (Kushlick, 1966, p. 130).

In the IQ range from 50 to 70, on the other hand, at least 75% of the individuals included therein appear clinically normal, evincing no signs of neurological damage, sensory defects, or physical stigmata. In fact, a report of the National Institute of Neurological Diseases and Blindness states that in 75 to 80% of *all* cases of mental retardation there is no specific identifiable cause such as those found in the categories outlined above (Research Profile No. 11, 1965).

These cases of retardation with no clinically identifiable cause are now commonly labeled cultural-familial retardation. The vast majority bearing this designation fall in the IQ range from 50 to 70. The evidence seems quite clear that these clinically normal persons are a part of the normal distribution of intelligence in the population, a distribution which is determined mainly by polygenic inheritance—that is, the influence of a large number of genes each of which contributes a small increment to mental ability (Gottesman, 1963). Familial retardation represents the bottom 2 to 3% of the lower tail of this normal distribution.

Some 70 to 80% of all persons identified as retarded at some point in their lives are in the familial category (Heber, Dever, & Conry, 1968).

The well-known excess or bulge at the lower end of the IQ distribution is attributable to major gene defects and brain damage which override normal polygenic determinants of intelligence. A study in England based on a complete sample of 3361 children showed actual frequencies not in excess of the frequencies expected from the normal or Gaussian distribution *above* IQs of 45. But the frequency of IQs *below* 45 was almost 18 times greater than would be expected (Roberts, 1952).

The most convincing evidence that the severely subnormal and the mildly subnormal familial retardates are different distributions and not different parts of a single underlying continuum of causal factors is the differences in amount of regression toward the mean IQ of the general population seen in the siblings of two types of retarded children. The siblings of familial retardates, on the average, have an IQ about half-way between the IQ of their retarded sib and the mean of the general population, an amount of regression that is rather precisely predictable from a polygenic model of the inheritance of intelligence. The very same amount of regression toward the mean is found in siblings of gifted children. On the other hand, the siblings of retardates with extremely low intelligence (IQs below 45 or 50) have an average IQ which is the same as the mean for the general population. In other words, the mental defect of the retarded sibling is superimposed upon and overrides the normal polygenic basis for intellectual development. Presumably the majority of the severely retarded would have been of normal or superior intelligence were it not for the devastating effect of a mutant gene, an abnormal chromosome, or brain damage (Shields & Slater, 1961).

It is still uncertain whether the normal distribution of polygenically determined intelligence extends below IQ 50 or thereabouts. The determination of this is made extremely difficult by the very small proportion of all retardates below IQ 50 that would be expected at this extreme of the normal curve. It is entirely possible, however, that some proportion of the 30% of the severely subnormal for whom no clinically identifiable etiology can be found are actually the lowest extreme of the normal distribution.

B. Cultural-Familial Retardation

Having now made this basic distinction between subnormality due to major genetic defects and neurologic damage, on the one hand, and cultural-familial retardation, on the other, the remainder of this paper is concerned with taking a diagnostically more analytic look at the cultural-familial category of mental retardation. This is not a sharply defined category. Traditionally the criteria for the diagnosis of cultural-familial includes IQs in the range from 50 to 70 or 75 and to this criterion is generally added some assessment of social competence. Persons not deficient in social competence are seldom regarded as retarded, despite a low IQ, except within the traditional school setting. From an educational standpoint and in terms of the scholastic requirements for entry into an ever increasing proportion of today's occupations, IQs below 85 are usually associated with educational retardation within the context of ordinary schooling, and consequently also with limited occupational opportunities. In preliterate and preindustrial societies most persons in the IQ range from 70 to 85 would not be perceived as retarded or occupationally disadvantaged, but in today's technological society they are at a marked disadvantage. More occupations today call for a higher level of developed skills than was true for past generations. Largely for this reason the American Association on Mental Deficiency has changed the intelligence test part of the criterion for retardation from two standard deviations (IQ 70) below the population mean to only one standard deviation (IQ 85) below the mean.

Edgerton (1968), an anthropologist who has studied mental retardation in primitive tribes, has expressed the doubt that the persons he has observed in industrial societies with the diagnosis of retardation in the IQ range 50 to 70 would be competent even in simpler, preliterate societies. Edgerton claims that the demands of life in African tribal society, for example, involve an amount of learning of customs, knowledge, and skills that is more than could be coped with by most persons regarded as mildly retarded by the usual IQ criterion. This is an important observation in the light of the major hypothesis put forth in this paper, for it falls in line with the observations that initially led to the studies which form the basis for our hypothesis, namely, the observation that some, perhaps many, of the children found to be retarded in school performance and on IQ tests appear to be normal and even bright in terms of a variety of criteria that clearly lie outside the scholastic realm.

The most likely reason that students of mental retardation have in the past failed to note or to emphasize this observation is that the criterion of social incompetence, as well as low IQ and poor scholastic performance, has determined the diagnosis of retardation and, even more than the intelligence test or scholastic criteria, has been the chief basis for admission to institutions for the retarded. A much broader spectrum of mental retardation is to be found in the public schools than in special residential institutions, and it would be difficult, if not impossible, to observe in institutions one type of retardation we have seen frequently in public schools—a "bright" child with a presumably valid low IQ (i.e., 50–75) which, in addition to his low scholastic performance, often results in his being placed in a special class for the retarded or for "slow learners."

A reformulation of the classification of cultural-familial retardation would therefore seem to be in order. A monolithic conception of this category, for example, has led to disputes over the claim that many persons are retarded only during their school years and once they leave school they become non-retarded. Mental retardation is thus viewed as a condition that results largely from the imposition of middle class standards and values by the schools. However, Heber *et al.* (1968) have noted that this interpretation fails to consider that the opportunities and criteria for evaluating mental retardation are very different for the preschool and postschool populations. Assessment based on clinical psychological tests have shown approximately the same incidence of retardation in the pre- and post-school population as are found in school, which only means that the criteria used in the psychological clinic are much the same as those used in schools. In the pre- and post-school years the IQ is less important and behavioral maturity and social competence are more important criteria in the assessment of retardation. Despite the general stability of the IQ throughout and beyond the school years, there are marked differences among children classed by the school as retarded. They differ in their social and occupational competence after leaving school, and these differences are only slightly correlated with IQ and scholastic performance. Some other important dimensions of ability, not assessed by the usual IQ tests nor highly correlated with scholastic performance, would seem to be involved in this phenomenon. We are concerned to find the nature of these non-IQ abilities and their educational and social implications.

II. MENTAL RETARDATION AND SOCIAL CLASS

Kushlick (1966, p. 130) has pointed out the fact that parents of severely subnormal children are evenly distributed among all the social strata of industrial society. Cultural-familial retardation, on the other hand, is predominantly concentrated in the lower social classes. On the basis of a number of surveys made largely in England, Kushlick concludes that "mild subnormality in the absence of abnormal neurological signs (epilepsy, electroencephalographic abnormalities, biochemical abnormalities, chromosomal abnormalities, or sensory defects) is virtually confined to the lower social classes." He goes on to say "there is evidence that almost no children of higher social class parents have IQ scores of less than 80, unless they have one of the pathological processes mentioned above." The same conclusion has been drawn by other investigators (e.g., Hardy, 1965) and is entirely consistent with the writer's experience gained in conducting studies in schools in lower class and middle class neighborhoods. The incidence of mild retardation is undoubtedly strongly associated with socioeconomic status (SES). Anyone who has attempted to do research on the relationship between retardation and SES knows the extreme difficulty in finding subjects in the IQ range from about 50 or 60 up to about 80 or 85 in the middle and especially upper-middle class segment of the population. Conversely, it has been our experience that it is not nearly as difficult to find gifted children (IQs above 130) in the lower classes as it is to find mildly retarded children in the upper classes. The Scottish National Survey established on a large scale that *high* intellectual ability is more widely distributed over different social environments than is low intellectual ability (Maxwell, 1953). This finding, of course, reflects the increasing *range* of mental test scores that we find as we move from the upper to the lower levels of occupational status. The upper bound of the IQ range changes relatively little going down the occupational scale, while the lower bound of the IQ range decreases markedly in going downward from the professions to unskilled labor (Tyler, 1965, pp. 338–339).

The association of the incidence of retardation with SES is also entirely consistent with the results of research on the relationship of SES to intelligence over the entire range of IQs. Correlations between the occupational status of adults and their IQs range between .50 and .70 (Tyler, 1965, p. 343) and between parents' occupation and children's IQ the correlations are, of course, lower than this—half of all such correlations reported in the literature are between .25 and .50 (Jensen, 1968c).

A. Genetic and Environmental Factors

The correlation between IQ and SES has led some writers to attribute the cause of this association strictly to environmental factors associated with SES. Neff (1938), for example, concluded from his extensive review of the evidence that environmental factors alone were sufficient to account for the observed relationship between SES and IQ. This conclusion, however, is decisively contradicted by evidence found in Neff's own review. If Neff accepts as valid the correlations he cites between the IQs of pairs of identical and fraternal twins, he must acknowledge the conclusion derived from these correlations, namely, that individual differences in intelligence have a genetic component. Once this is accepted, Neff's argument collapses unless it could be shown that there is no correlation whatsoever between the genetic component of intelligence variance and persons' occupational and educational status, which are the chief indices of SES. Similarly, a recent textbook states: "Inborn or biological differences in intelligence exist, but between individuals, not between large social or racial groups [Havighurst & Neugarten, 1967, p. 159]." For this statement to be true it would have to mean that all the factors involved in social mobility, educational attainments, and the selection of persons into various occupations have managed scrupulously to screen out all variance associated with genetic factors among individuals in various occupational strata. The possibility that the selection processes lead to there being only environmental variance among various socioeconomic groups and occupations—a result that could probably not be accomplished even by making an explicit effort toward this goal—is so unlikely that the argument amounts to a *reductio ad absurdum.* If individual differences in intelligence are due largely to genetic factors, then it is virtually impossible that average intelligence differences between social classes (based on educational and occupational criteria) do not include a genetic component.

This argument goes as follows. Twin studies and other methods for estimating the heritability of intelligence have yielded heritability values for the most part in the range from .70 to .90, with a mean value of about .80 (Jensen, 1967). Heritability (H) is a technical concept in quantitative genetics, referring to the proportion of variance in a metric characteristic, such as height and intelligence, that is attributable to genetic factors. $1 - H = E$, the proportion of variance due to non-genetic or environmental factors, which of course include prenatal as well as postnatal influences. The correlation between phenotypes (the measureable characteristic) and genotypes (the genetic basis of the phenotypes) is the square root of the heritability, i.e., $\sqrt{H}$. An average estimate of $\sqrt{H}$ for intelligence is .90, which is the correlation between phenotype and genotype. An average estimate of the correlation between occupational status and IQ (i.e., phenotypic intelligence) is .50. What Neff (1938) and Havighurst and Neugarten (1967) are saying, essentially, is that the correlation between IQ and occupation (or SES) is due entirely to the environmental component of IQ variance. In other words, their hypothesis requires that the correlation between the genotypes and SES be zero. So we have three correlations between three sets of variables: (*a*) between phenotype and genotype, $r_{pg} = .90$; (*b*) between phenotype and status, $r_{ps} = .50$; and (*c*) the hypothesized correlation between genotype and status, $r_{gs} = 0$. The first two correlations (r_{pg} and r_{ps}) are determined empirically, and are represented here by average values reported in the research literature. The third correlation (r_{gs}) is hypothesized to be zero by those who, like Neff and Havighurst and Neugarten, believe genetic factors play a part in individual differences but not in group differences. The question then becomes: is this set of correlations possible? The first two correlations we know are possible, because they are empirically obtained values. The correlation seriously in question is the hypothesized $r_{gs} = 0$. We know that mathematically the true correlations among a set of variables, 1,2,3, must meet the following general requirement: $r^2_{12} + r^2_{13} + r^2_{23} - 2r_{12}r_{13}r_{23}$ cannot have a value greater than 1.00. The fact is that when the values of $r_{ps} = .50$ and $r_{gs} = 0$ are inserted in the above formula, they yield a value greater than 1. This means that r_{gs} must in fact be greater than zero.

Perhaps an even simpler way of regarding this problem is as follows: if only the *E* (environmental) component determined IQ differences between status groups, then the *H* component of IQs would be regarded as random variation with respect to status. Thus, in correlating IQ with status, the IQ test in effect is like a test with a reliability of $1 - H = 1 - .80 = .20$. That is to say, only the *E* component of variance is not random with respect to indices of SES. Therefore the theoretical maximum correlation that IQ could have with SES would be $\sqrt{.20} = .45$. This value is very close to the obtained correlations between IQ and SES. So if we admit no genetic component in SES differences, we are forced to conclude that persons have been fitted to their socioeconomic status (meaning largely educational attainments and occupational status) almost *perfectly* in terms of their environmental advantages or disadvantages. In other words, it must be concluded that persons' innate abilities, talents, and proclivities play no part in their educational and occupational placement. This seems a preposterous conclusion. The only way one can reject the conclusion that there are genetic intelligence differences between SES groups is to reject the evidence on the heritability of individual differences in intelligence. But the evidence for a substantial genetic component in intellectual differences, is among the most consistent and firmly established research findings known in the fields of psychology and behavioral genetics. Much of the relevant evidence has been reviewed in detail elsewhere (Burt, 1955, 1958, 1959, 1961a, 1966; Eckland, 1967; Erlenmeyer-Kimling & Jarvik, 1963; Fuller & Thompson, 1960; Gottesman, 1963, 1968; Huntley, 1966; Jensen, 1967, 1968a, 1969; Jones, 1954).

More direct lines of evidence for SES genetic intelligence differences are also available. For example, the weak effect of SES as a causal factor in intellectual differences is seen in studies of identical twins separated shortly after birth and reared in different homes. The most valuable of these studies is by Sir Cyril Burt (1966), since the 53 pairs of identical twins in his study were separated at birth or within the first 6 months after birth and were reared apart in families that ranged across all the SES categories of the British census. Furthermore, there was a slightly negative but nonsignificant correlation between co-twins with respect to the SES of the homes in which they were reared. Yet the correlation between the Stanford-Binet IQs of co-twins at about 10 years of age was .87, which corresponds to an average difference of about 6 points on the IQ scale. (Corrected for attenuation, i.e., test unreliability, the difference is about 4 points.) Not all of even this small difference is due to social environmental factors; some of the difference, perhaps as much as half, is probably attributable to prenatal factors. Co-twins are not equally advantaged with respect to intrauterine space and prenatal nutrition; this is reflected in inequalities in their birth weights, inequalities which are correlated (positively) with their later IQs (Willerman & Churchill, 1967).

Another line of evidence is from studies of adopted children. The correlation between their IQs and the educational level of their biological parents is about the same as for children reared by their biological parents, while the correlation beween the adopted children and the education of the adopting parents is close to zero (Honzik, 1957). Children reared from infancy in an orphanage, and with no knowledge of their biological parents, show nearly the same correlation (about .25) between IQ and father's occupational status (graded into five categories) as is found for children reared by their parents (Lawrence, 1931). Also, adopted children show a smaller dispersion of mean IQ level as a function of SES of the adopting parents than do children reared by their own parents. Leahy (1935) matched two sets of parents on a number of SES indices—parents rearing their own children and foster parents of adopted children. Table I shows the mean IQs of the adopted and control children as a function of the father's or foster father's occupation.

TABLE I
IQs of Adopted and Control (own) Children in Homes of Different Occupational Categories[a]

	Adopted children			Control (own) children		
Occupation of father	*N*	Mean IQ	SD	*N*	Mean IQ	SD
Professional	43	112.6	11.8	40	118.6	12.6
Business manager	38	111.6	10.9	42	117.6	15.6
Skilled trades	44	110.6	14.2	43	106.9	14.3
Farmers	—	—	—	—	—	—
Semi-skilled	45	109.4	11.8	46	101.1	12.5
Slightly skilled / Day labor	24	107.8	13.6	23	102.1	11.0
General mean	194	110.6		194	109.7	

[a]Taken from Leahy (1935).

The variance among the occupational means for the control children's IQs is 15 times greater than among the mean IQs for adopted children (56.24 vs. 3.72).

Siblings have on the average only half of their genes in common, and show an average correlation of .5 for intelligence and other highly heritable traits. The average absolute intelligence difference between sibs reared together is about 12 IQ points on the Stanford-Binet. Most of the intelligence difference between siblings reared together is attributable to their genetic differences. There is evidence that when siblings reared in the same family move into different social strata, the sibs with IQs above the family average are more likely to move to a SES above that of their family and sibs with IQs below the family average are more likely to move down in SES (Young & Gibson, 1965). This condition would, of course, cause the gene pools for intelligence to differ among SES levels.

Since the mean IQ differences between SES categories reflect some combination of genetic and environmental determinants of intelligence, and since there is a broad spread of IQs about each category mean, as shown by the standard deviations of 10 to 12 points *within* SES categories, there should be increasing proportions of children falling below IQ 75, the borderline of mental retardation, in the IQ distribution of each SES category from the highest to the lowest. If genetic factors are predominant, the increasing proportion of IQs below 75 as we move down the scale of SES, should be in evidence throughout the scale, even between the higher SES categories in which there is no environmental disadvantage or deprivation in the usual sense of the term. Even the most disadvantaged environments found in industrial society, short of rare cases of almost total social isolation, do not produce IQs below 75 in the majority of children reared in such deprived environments. Thus genetic factors are almost certainly implicated in this degree of retardation, even when it occurs at the lowest end of the SES continuum. On the basis of large normative studies of the Stanford-Binet, Heber *et al.* (1968) have estimated the prevalence of IQs below 75 as a function of SES and race, as shown in Table II. It should be kept in mind that the estimates in Table II are based on Stanford-Binet IQs. We now have good reason to believe that on some other tests of mental ability, to be described shortly, the percentages for whites and Negroes would be much more similar than those in Table II, and SES differences would be very much smaller.

All this is quite consistent with what is known about polygenic inheritance. If we accept the polygenic theory of the inheritance of intelligence, which is strongly supported by the evidence, it follows that a certain proportion of the population will have relatively low intelligence. Furthermore, if we recognize the fact of what geneticists call assortative mat-

TABLE II
ESTIMATED PREVALENCE OF CHILDREN WITH IQS BELOW 75, BY SOCIOECONOMIC STATUS (SES) AND RACE GIVEN AS PERCENTAGES[a]

SES	White	Negro
High 1	0.5	3.1
2	0.8	14.5
3	2.1	22.8
4	3.1	37.8
Low 5	7.8	42.9

[a]Taken from Heber *et al.* (1968).

ing—the tendency for like to marry like—we should expect that the frequency of genes for intelligence would become unequally assorted in different families and groups in the population. If persons were mated on a purely random basis, the average absolute difference in IQ between husbands and wives would be about 18 IQ points.[1] The degree of assortative mating in our society, however, is such that the average absolute difference between husbands and wives is actually between 10 to 13 IQ points, according to various studies. Thus, in terms of the polygenic theory the binomial expansion of $(\frac{1}{2} A + \frac{1}{2} a)^{2n}$ (where A and a represent intelligence enhancing and non-enhancing genes, respectively, and n is the number of gene loci) must be regarded as representing only the relative frequencies of these genes in the population. On the average, the frequencies of A and a genes in the population are assumed to be equal. Within a group selected for intelligence, however, the relative frequencies of A and a genes may be quite different, say, 20% A and 80% a, so that the binomial expansion of $(.2A + .8a)^{2n}$ will yield a skewed distribution of values, in this case having a preponderance of low values. The normal distribution of phenotypes in the total population should be thought of as the average of many differently skewed distributions for various "breeding groups." A variety of social, ethnic, educational, and economic factors in our society insures a high degree of assortative mating with respect to intelligence.

Given this polygenic model, plus the fact of assortative mating, we should predict that mental retardation would not occur in all families with equal probability. From this model it would be estimated that at least 25% of retarded persons would have one or both parents retarded. A corollary of this is that if none of the retarded reproduced, there

[1]The mean absolute difference between all possible pairs of scores in a normal distribution is equal to $2\sigma/\sqrt{\pi}$. For the Stanford-Binet test $\sigma = 16.4$.

would be a substantial reduction in the frequency of retardation in the next generation.

The most monumental study of this matter has been carried out by two geneticists, Elizabeth and Sheldon Reed, and their colleagues, at the University of Minnesota (Reed & Reed, 1965). They began with 289 retarded persons (IQ below 70) who were resident in a state institution for the retarded at some time during the years 1911 to 1918. From this nucleus of 289 retardates, the investigation branched out to include the study of 82,217 of their relatives. Practically all the descendants of the grandparents of the probands (i.e. the originally selected retardates) were included. Family pedigrees were traced over as many as seven generations, the primary aim being to determine as accurately as possible the mental status of all persons in the study. This involved searching school records for the subjects' grades and IQ scores and following their occupational histories. Analysis of these massive data lead to some clear conclusions.

First, it should be pointed out that in the following discussion of the Reeds' study the term "retarded" always means an IQ below 70. Since such individuals constitute about 3% of the white population, it means there are close to 6 million retardates in the white population of the United States.

The Reeds found that only 0.5% of children of normal parents (i.e. IQs above 70) with normal siblings were retarded.[2] The remaining 2.5% of the population who are retarded, therefore, have at least one parent or an aunt or uncle who is retarded. In other words, some 5 million of the 6 million retardates in the United States have a retarded parent or a normal parent who has a retarded sibling. Among 15,000 unselected retardates 48.3% had one or both parents retarded. The belief that the retarded of one generation contribute only a negligible proportion of the retarded of the next generation is therefore patently false.

Assortative mating occurred to a very high degree in families with a high incidence of retardation; retardates rarely marry anyone much above their own level. However, it is of some interest that 30% of illegitimate children born to the 289 probands were retarded, while only 11% of legitimate children were retarded. One might expect just the oppo-

[2]It is of interest that this is close to the percentage of retarded found among the offspring of Terman's gifted group. These were 1528 school children selected for IQs over 135 (mean IQ of entire group = 152). Their development has been followed into adulthood (most of them are now in their fifties). Among the 2452 children born to gifted parents, only 13 or 0.53% were retarded. Most of these cases were probably due either to major gene defects or brain damage rather than to polygenic inheritance. The average IQ of all the offspring of the gifted group was 132.7 when they were last tested (Terman & Oden, 1959, p. 404).

site. The explanation is that a high percentage of illegitimate children in this group were the product of incestuous relationships which would, of course, increase the probability of producing genotypes in the retarded range.

It is certainly true that the children of retarded parents are often subjected to a culturally and intellectually impoverished environment that would tend to depress their mental development. Yet, it is most important to note that of the children of retarded parents fewer than half are retarded. This would be difficult to explain strictly in terms of environmental influence. But it is what we should expect in terms of the polygenic theory. Although nearly all the children born into subnormal homes are presumably subjected to influences unfavorable to intellectual development, the fact that more than half of such children are not mentally retarded suggests that the more intelligent children must have received more desirable gene combinations.

Another striking finding is that retardation was extremely rare in some families. For example, in 37 of the families of the 289 cases, the only retardate was the proband. In some large families comprising over 2400 persons there were less than 1% retarded.

It is instructive from the standpoint of genetics to note the frequency of retardation among relatives of the probands as the distance of relationship increases. The results of such an analysis are shown in Table III. The probands were classified on the basis of case histories into one of four categories describing the most likely cause of retardation. The percentage of retarded relatives for three degrees of relationship was also determined, as shown in Table III. First degree relationships are those with whom the proband has one-fourth of his genes in common: mother, brothers, sisters, and children. Second degree relationships are those with whom the proband has one-fourth of his genes in common: grandparents, uncles, aunts, half-siblings, nephews, nieces, and grandchildren. Relatives of the third degree are those with whom the proband has one-eighth of his genes in common: half-uncles and aunts, half-nephews and nieces, great-nephews and nieces, and first cousins.

The point of primary interest in Table III is the rapid drop in the incidence of retardation as we go from first to second to third degree relatives. (Recall that the incidence of retardation in the general population is about 3%.) Note also that the etiological categories differ in the percentage of retarded relatives and in the rate of decline as the degree of relationship becomes more distant. Why should the category "primarily genetic" have fewer retarded relatives than the "probably genetic" category? First, because the "primarily genetic" category included some probands with major gene defects about which there was no doubt concerning genetic origin (and, as was pointed out earlier, these defects are very

TABLE III
THE PERCENTAGES OF RETARDATION IN THE RELATIVES OF THE PROBANDS ACCORDING TO DEGREE OF RELATIONSHIP AND CATEGORY OF CLASSIFICATION[a]

Category	First degree	Second degree	Third degree	Average percentage retarded
Primarily genetic	33.6	9.2	3.7	8.8 (452 of 5149)
Probably genetic	50.7	16.8	5.3	13.2 (496 of 3759)
Environmental	21.4	2.0	1.1	3.3 (60 of 1831)
Unknown	15.6	2.6	2.1	3.7 (275 of 7327)
All categories				
Percentages	28.0	7.1	3.1	7.1
Totals	(532 of 1897)	(434 of 6070)	(317 of 10,099)	(1283 of 18,066)

[a]Taken from Reed and Reed (1965).

rare); second, because the chief criterion for classification into the category "probably genetic" was that the proband have retarded relatives in the first degree of relationship.

Table IV indicates the IQ frequency distributions of children resulting from various matings in which either one or both parents were retarded. It is most interesting that a number of bright (IQs 111–130) and definitely superior (131+) children resulted from such matings, despite

TABLE IV
IQ RANGE OF TESTED CHILDREN OF RETARDATE UNIONS[a]

Type of union	IQ range: 0–49	50–69	70–89	90–100	111–130	131+	Total	Average IQ	Percent retarded
Retardate × retardate	6	29	36	17	1	0	89	74	39.4
Male retardate × normal	0	12	41	75	24	1	153	95	7.8
Female retardate × normal	6	15	32	43	10	1	107	87	19.6
Male retardate × unknown	3	16	68	80	20	1	188	90	10.1
Female retardate × unknown	10	29	64	79	22	2	206	87	19.0
Total	25	101	241	294	77	5	743	86	17.0

[a]Taken from Reed and Reed (1965).

the fact that some of these children came from what the Reeds described as "extremely impoverished environment." The largest number (294) of children from retardate unions was found in the average range of IQs from 90 to 110, again despite impoverished environment. Note, however, the skew of the overall distribution (i.e. the bottom "Total" line).

Another interesting feature of these data is that the mating of male retardate × normal female results in a significantly lower percentage of retarded offspring than the mating of a female retardate × normal male. Two hypotheses are suggested by this: (*a*) When the mother is retarded, the child's early environment may be more severely lacking in the kinds of mother-child interaction that promote mental development; (*b*) the retardate mothers may provide a poor *prenatal* environment for the developing fetus. Adverse intrauterine conditions could also have a genetic basis.

Table V shows the results of various retardate matings in more precise terms, made possible by having IQ scores on both parents.

Like low IQs, high IQs tend to cluster in particular families, rather than occurring in random distribution among families. In one family where the parents had IQs of 157 and 151, the three children had IQs of 132, 134, and 149. An unusual union in which one parent had an IQ of 135 and the other an IQ of 67 resulted in five children with IQs of 112, 115, 113, 97, 131 (average IQ of parents = 101, average IQ of children = 114).

All these findings taken together would seem to provide a more than adequate answer to the view expressed in a well-known book on mental subnormality by Masland, Sarason, and Gladwin (1958, p. 196): "We do not propose to *deny* that heredity is a factor, particularly in mental deficiency, but rather that we should leave it out of our accounting until it is supported by more than speculation and bias." The hereditary aspect of mental retardation is obviously now supported by more than "speculation and bias."

Furthermore, there would seem to be some eugenic implication in the Reeds' conclusion that

> . . . the one to two percent of our population composed of fertile retardates produced 36.1 percent of the retardates of the next generation, while the other 98 to 99 percent of the population produced only 63.9 percent of the retarded persons in the next generation [p. 48].

The fact that the majority of the mildly retarded (IQs 50–70) are found in the lowest socioeconomic classes means that the majority of the mildly retarded children are born to parents who have the least to offer their children. The Reeds do not believe that social deprivation is a primary cause of retardation in the IQ range below 70. They state:

TABLE V

IQ Range of Tested Children of Retardate Unions in Which Both Parents Had Been Tested[a]

Type of union	IQ range 0–49	50–69	70–89	90–100	111–130	131+	Total	Average IQ of children	Percent retarded
Both parents IQ 60 or below; average IQ 60 (12)	5	23	12	6	0	0	46	67	60.9
Father IQ 69 or below, average IQ 62; mother IQ 70 or above, average IQ 92 (26)	3	3	20	43	12	1	82	94	7.3
Mother IQ 69 or below, average IQ 63; father IQ 98 (15)	0	9	18	20	2	0	49	86	18.4
Total (53)	8	35	50	69	14	1	177	82	24.3

[a]Taken from Reed and Reed (1965).

We must assume that some cases of mental retardation are due primarily to social deprivation, but we don't find a large proportion of our probands who are available for this classification after an allocation has been completed for the causes which appear to have been present [p. 75].

They proceed to say: "One inescapable conclusion is that the transmission of mental retardation from parent to child is by far the most important *single* factor in the persistence of this social misfortune [p. 48]." The problem is how to prevent the approximately 6 million retarded persons in the United States from transmitting it genetically or environmentally. The Reeds conclude:

The transmission of mental retardation from one generation to the next, should, therefore, receive much more critical attention than it has in the past. It seems fair to state that this problem has been largely ignored on the assumption that if our social agencies function better, that if everyone's environment were improved sufficiently, then mental retardation would cease to be a major problem. Unfortunately, mental retardation will never disappear, but it can be reduced by manipulating the genetic and environmental factors involved . . . When voluntary sterilization for the retarded becomes a part of the culture of the United States, we should expect a decrease of about 50 percent per generation in the number of retarded persons, as a result of all methods combined to reduce retardation [p. 77].

An important point, in terms of the theory of primary and secondary retardation proposed in this paper, must be made concerning the interpretation and conclusions of the Reeds' study of familial retardation. It should especially be noted that all the retardates in this study were found by tracing down the more than 82,000 "blood" relationships of the 289 *institutionalized* probands. As will be shown in a later section, there is good reason to believe that institutionalized retardates differ in important ways from many individuals with IQs in the 50 to 70 range who do not become institutionalized. It seems very likely that a high proportion of the institutionalized retarded are the result of different genetic factors than those involved in the majority of noninstitutionalized persons with IQs below 70–75. Study of the relatives of institutionalized persons is also likely to give a much stronger weight to hereditary than to environmental and educational factors in the causation of retardation. We have found that there are some psychologically fundamental differences in the patterns of mental abilities between (*a*) institutionalized retardates, (*b*) non-institutionalized retardates from socially deprived backgrounds, and (*c*) retardates from non-deprived or middle-class backgrounds.

B. Motoric Precocity and Later Intelligence

Another interesting and important fact in terms of its diagnostic implications in the light of our theory of primary and secondary retardation is the low but significant negative correlation generally found between performance on infant mental tests, such as the Bayley Scales, and later IQ. Infant tests for children under 2 years of age yield a Developmental Quotient (DQ), as distinguished from the IQ, which can be obtained beyond 2 years of age by means of tests such as the Stanford-Binet. Bayley (1965b) has shown that it is the motor subtests rather than the perceptual-attentional subtests that largely account for the slightly negative correlation between DQ and IQ. Furthermore, up to about 1 year of age, the DQ—largely due to the motoric items—has a negative correlation with the SES level of the infants' parents. This inverse relationship between DQ and parental SES is much more marked in boys than in girls, for whom the correlation is close to zero. Bayley believes that genetic factors are involved in these relationships, and the pronounced sex difference at this early age would support this view. Beyond 2 years of age, on the other hand, boys and girls both show an increasingly positive correlation between IQ and SES. Bayley's results are shown in Fig. 1. Bayley (1965a) has also found that Negro infants up to 15 months of age perform better on the Bayley Scales, especially on the motor items, than white infants of comparable age. The highest mean scores on the Bayley Scales for any sizeable group that I have found reported in the literature were obtained on Negro infants of about 6 months of age living in the poorest neighborhoods of Durham, North Carolina (Durham Education Improvement Program, 1966–1967a,

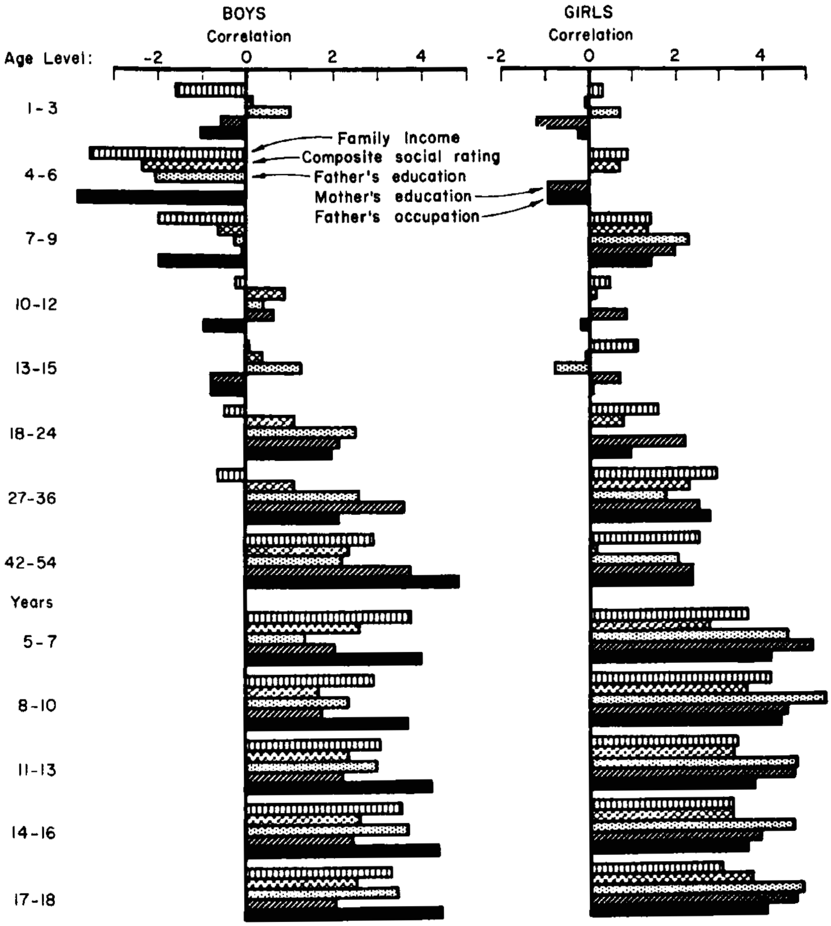

FIG. 1. Correlations between children's mental test scores, at 1 month to 18 years, and five indicators of parents' socioeconomic status at the time the children were born (from Bayley, 1966).

1966–1967b). These infants obtained Developmental Quotients on the motor items of the Bayley Scale averaging about 1 standard deviation above white norms. (On non-motor items they averaged half of a standard deviation above white norms.) The older siblings of these infants, by contrast, had IQs averaging about 1.3 standard deviations below white norms. Thus the negative correlation between DQ and IQ appears very marked in this segment of the Negro population. Similar findings have been reported in at least five other studies (Bayley, 1965a; Curti, Marshall, Steggerda, & Henderson, 1935; Geber & Dean, 1966; Knoblock & Pasamanick, 1958; Walters, 1967).

When the test employed involves strictly cognitive rather than motoric aspects of development, negative correlations between performance and SES are found in children even below 12 months of age. For example, Kagan (1966) reports that on certain laboratory tests of cognitive functioning

> lower-class children, as early as 8 to 12 months of age, show slower rates of information processing than middle-class children of the same ordinal position. Lower-class children show less rapid habituation, less clear differentiation among visual stimuli, and, in a play situation, show a high threshold for satiation. The latter measure is obtained by placing the child in a standard playroom with a standard set of toys (quoits on a shaft, blocks, pail, mallet, peg board, toy lawn mower, and toy animals) and by noting the time involved in each activity. Some children play with the blocks for 10 seconds and then skip to the quoits or the lawn mower, playing only 10–20 seconds with each individual activity before shifting to another. A second group of children, called "high threshold for satiation infants," spends 1 or 2 minutes with an activity without interruption before changing. We do not believe the latter group of infants is taking more from the activity; rather it seems that they are taking longer to satiate on this action. It is important to note that the observation that lower-class infants show high threshold for satiation contrasts sharply with the observation that 4-year-old lower-class children are distractable and hyperkinetic. We believe both descriptions. The paradox to be explained is why these lower-class children are pokey and lethargic and nondistractible at 12 months of age, yet display polar-opposite behavior at 48 months of age.

III. THEORY OF PRIMARY AND SECONDARY RETARDATION

The empirical findings on which our hypothesis of primary and secondary retardation is based can be more easily summarized and their relevance more readily indicated if the hypothesis is described first in general terms.

A. A Hierarchy of Abilities

There is much evidence that mental abilities stand in some hierarchical relationship to one another. A number of factor analytic models have yielded results consistent with a hierarchical hypothesis (Vernon, 1950), but, as pointed out by Guilford (1967), the hierarchical factor model is as much a product of the particular method of factor analysis as of the raw data that go into it, and other models than hierarchical ones are possible. However, there are other lines of support for a hierarchical view of abilities which stem from experimental studies of the learning process, such as Gagné's (1962, 1968) work on learning hierarchies, and from studies of the developmental aspects of cognitive processes, such as those reviewed by White (1965). Both lines of evidence indicate that for many abilities there is a natural order of acquisition or emergence, such that when ability B is found, ability A will always be found, but not the reverse. Deficiencies in a lower level ability almost always imply deficiency in some higher level ability, but the reverse need not be the case. Some aspects of the ability hierarchy are attributable to the learning of specific subskills which stand in some hierarchical relationship to one another; these aspects are usually more closely related to the individual's grade in school and to the nature of the instruction he has received up to that point. Learning various operations and concepts in arithmetic is a good example. Other abilities are of a more maturational or developmental nature and are practically impossible to explain in terms of previous learning of specific subskills. The emergence of such abilities is apparently more dependent upon the growth of brain structures than upon learning and experience. Experience may be necessary but it is far from sufficient for certain abilities to become manifest in performance. Abilities that depend upon the maturation of neural structures can also be hierarchical, in the sense that normal maturation of a lower level does not necessarily insure maturation of higher levels in the hierarchy. Failure of maturation at lower levels, on the other hand, will result in some deficiency or impairment of the emergence of higher level functions in behavior, even if their neural substrate is normal.

The essential characteristic that most generally describes the levels of this mental maturation hierarchy is the degree of correspondence between "input" and "output." Lower levels of the hierarchy involve relatively little processing or transformation of the informational input; the stimulus-response correspondence is relatively simple and direct. Higher levels of the mental ability hierarchy depend upon elaborations and transformations of informational input, and upon comparisons of the informational input with previously stored information. Various cognitive tasks can be hypothetically placed along this continuum, from low to high: simple reaction time, Pavlovian conditioning, instrumental conditioning, complex reaction time, pursuit-rotor learning, discrimination learning, immediate memory span for digits (forward), immediate memory span for digits (backward), memory span for digits after a brief delay (i.e., 5–15 seconds) between presentation and recall, serial rote learning, free-recall of uncategorized word lists, paired-associate learning, free-recall of categorized word lists, complex concept learning and problem solving (e.g., verbal analogies, arithmetic "thought" problems, Raven's *Progressive Matrices*). It should be noted that this continuum is not one of increasing task difficulty per se. A digit span test can be made more difficult than a *Progressive Matrices* problem in terms of percentage of the population "passing" the items. Neither does the continuum necessarily represent one of increasing stimulus (input) complexity. The continuum seems to be best described in terms of the amount of transformation of the input—the amount and complexity of "mental" activity—called forth in the subject in the process of his responding to the stimulus in order to learn, retain, recall, or produce the correct response to a problem.

1. Level I and Level II Abilities

Although up to now we have regarded these tasks as ranging along a single continuum, our hypothesis, for reasons that will become apparent, holds that the continuum is the resultant of at least two types of ability, which we shall call Level I or "associative ability" and Level II or "cognitive" ability.

Levels I and II are viewed as being qualitatively different, as existing in parallel, but as having quite different developmental rates. Individual differences in Levels I and II may in fact be correlated, but not because they are different manifestations of the same underlying structures or processes. That the underlying processes are essentially different and are not inherently correlated could be shown by obtaining groups of persons in whom the correlations are zero or even negative between tests that are highly loaded on Level I and tests loaded on Level II functions, such tests, for example, as digit span (Level I) and the *Progressive Matrices* (Level II). Probably no test on the behavioral level is completely free of both Levels I and II, but different tests can have markedly different loadings on each Level.

Correlation between tests of Level I and tests of Level II can occur in a given population mainly for three reasons:

(a) The essentially independent genetic factors determining individual differences in Level I and Level II may become associated through assortative mating. That is to say, persons who are below average in, say, scholastic ability, whether because they are below average in Level I or in Level II, or in both, have a greater probability of marrying one another than of marrying someone who is markedly different in ability. This tends to bring together in their offspring poor genetic potential for both Level I and Level II abilities. In the previous section in the review of the research of Reed and Reed (1965) on the genetic transmission of mental retardation, it was shown in Table IV that more retarded children resulted from matings of a retarded mother with a normal father than from a retarded father and a normal mother. While the explanation in terms of quality of the maternal environment offered by the Reeds is quite possibly sufficient, it is not the only possible explanation. A possible explanation in terms of the theory here proposed is that more of the retarded mothers than of the retarded fathers in the Reeds' sample had genotypes for deficiency in Level I abilities. Because of the demands of earning a living, mentally deficient men are less apt to be able to marry than retarded women, especially if the man's deficiency is in basic Level I processes, which would be a handicap in almost any line of work. Most standard intelligence tests are heavily loaded on Level II ability, and because of the hierarchical dependence of Level II on Level I for its manifestation in performance, a person who is deficient in Level I will also show some deficiency in behavioral indices of Level II. If Level I and Level II are under independent genetic control, and granting the hierarchical relationship between Levels I and II, one would predict that a normal person (i.e., average or above on Levels I and II) mated with a person genetically deficient in Level I would produce a higher proportion of phenotypically retarded children than a normal person mated with a person who is genetically deficient only in Level II abilities.

(b) The second basis for correlation between Levels I and II is already evident from the preceding discussion, viz., the functional dependence of the behavioral expression of Level II processes on Level I. The degree of this dependence is not yet completely known, but the evidence suggests that the degree of dependence may become increasingly weak above some "threshold" value of Level I; higher correlations between Level I and Level II tests would therefore be expected in the average to below average range of the distributions than in the above average ranges.

(c) Some of the information processing skills involved in Level II tests depend not only on the normal functioning of the neural substrate of Level I but also upon the prior learning of certain skills. The speed and thoroughness of acquisition of these skills depend also upon Level I associative learning ability. Thus there comes about a correlation between measures of Levels I and II.

2. Intelligence Tests

Most standard intelligence tests are made up of items that are a mixture of Level I and Level II functions. Partly for this reason, it has been difficult to infer the two types of processes from total scores on these tests; the scores are too much an amalgam of Level I and Level II functions. Most intelligence tests that are heavily loaded with what Spearman characterized as the *g* factor—a capacity for abstract reasoning—are mainly indices of Level II functioning. Among standardized tests, Raven's *Progressive Matrices* and Cattell's *Culture-Fair Tests* are perhaps the purest measures of Level II ability. The Stanford-Binet and Wechsler tests have slightly lower *g* loadings than the Raven and Cattell tests and also contain subtests which are relatively pure measures of Level I abilities, such as the digit span and digit symbol tests of the Wechsler. Moreover, these conventional IQ tests contain informational items, such as vocabulary and general information, which depend upon previous learning. The low conceptual quality of the definitions required for passing, especially for the easier, more concrete words, and the simple factual content of the general information items, would involve Level I ability as well as Level II. The net effect is that these tests order individuals along a general, crude continuum of intellectual ability, somewhat more heavily weighted with Level II ability, but without making any clear distinction between individuals' relative strength or weakness in Level I and in Level II.

Some children who obtain seemingly valid low IQs in the range 50 to 80 on these tests appear to be socially bright and do not seem in the least retarded in learning the names of classmates, in acquiring playground skills and the practical knowledge of getting along with their neighborhood playmates. For many such children, who usually come from the lower classes, the IQ test is commonly presumed to be invalid because of the cultural loading of its item content. While some of the items in such tests as the Stanford-Binet and Wechsler have an obvious cultural element, as have also many of the group tests used in schools, it has been found that these items are not necessarily those on which lower-class children with low IQs do the most poorly. These children generally do no better, and often they do worse, on the less culturally loaded subtests such as block designs, and on tests like Raven's *Progressive Matrices* and the *Culture-Fair Tests* of Cattell (see Jensen, 1968c). Something besides cultural bias of test items is clearly involved. Eells *et al.* (1951), in their famous study of cultural bias in standard intelligence tests, found that the one characteristic that distinguished most between items showing a large social class difference in the probability of giving the correct answer was the degree of *abstractness* of the test items. This attribute of test items is a more important factor in determining disparity of test scores between upper and lower classes than the factor of cultural content per se. Examination of items in standard tests, moreover, supports the conclusion that the more culturally loaded items in tests are also among the least abstract. "Who wrote *Faust*?" (an item in the Wechsler-Bellevue), for example, is more culturally biased, but also less abstract or conceptual, than some other less cultural items from the same test, such as "In what way are an *egg* and a *seed* the same?" and "If seven pounds of sugar cost twenty-five cents, how many pounds can you get for a dollar?" Probably it was largely because of this inverse relationship between the cultural loading and the abstractness of intelligence test items that it was possible for McGurk (1967) to show that Negro children performed better (relative to whites) on the more culturally loaded items than on the less cultural questions of an intelligence test.

The cultural loading of test items is best regarded as essentially orthogonal to the Level I–Level II dimension along which various tests may range. The writer has argued the point elsewhere that the most objective index of a test's culture-fairness is its heritability coefficient (h^2) in the normative population (Jensen, 1968c). The two-dimensional space which must be hypothesized in order to comprehend the facts of SES differences in measured intelligence is shown in Fig. 2. The hypothetical positions of various mental tests in this space are indicated.

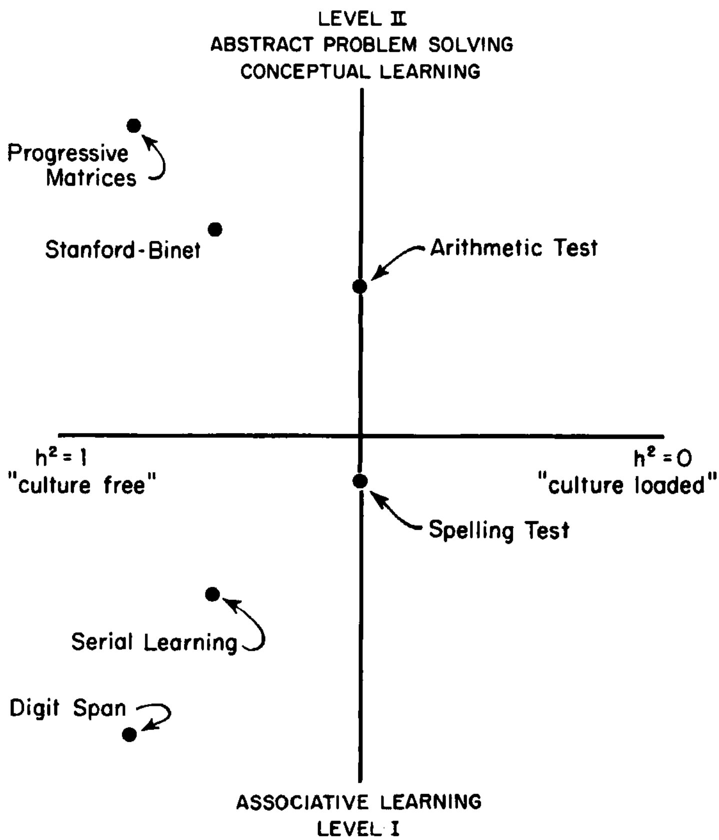

FIG. 2. The two-dimensional space required for comprehending social class differences in performance on tests of intelligence, learning ability, and scholastic achievement. The locations of the various "tests" are hypothetical.

Although various tests and forms of learning may differ in the extent to which they actually *require* Level II processes, there is little way to prevent Level II processes from entering into a subject's performance on tasks that require no more than Level I. Subjects tend to use whatever abilities they have at their command in approaching a learning or problem solving situation. Some tasks, however, minimize the usefulness of Level II processes. Mnemonic elaboration, coding, or other mediational processes are more often likely to hinder than to aid digit span memory, for example, and therefore digit span tests tap mostly Level I processes. Paired-associate (PA) learning, on the other hand, can be accomplished with Level I abilities, but Level II can also play a large role in PA learning. Thus, for individuals who are well endowed with Level II ability, such as college students, individual differences in PA learning may be determined largely by Level II, which will largely override individual differences in Level I. In young children, in whom Level II processes are still rudimentary, on the other hand, PA learning would be more a manifestation of Level I ability. Consequently, the correlation among tasks that can potentially involve both Level I and Level II but for which only Level I is essential should decrease with increasing age of the subjects from preschool to adolescence.

3. Relationship of Level I and II to "Fluid" and "Crystallized" Intelligence

Cattell (1963) has proposed a distinction between what he calls *fluid* and *crystallized* general intelligence.

Fluid intelligence is a basic capacity for learning and problem solving, a general "brightness" that is manifested in new learning, novel problem solving, and general intellectual adaptibility. It is independent of education and experience but is invested in the particular opportunities for learning afforded by the circumstances of the individual's life. Tests designed to minimize the importance of cultural and educational advantages, such as Cattell's *Culture-Fair Tests* and Raven's *Progressive Matrices*, are the best measures of fluid intelligence. Fluid intelligence reaches the peak of its growth curve in late adolescence, and thereafter reaches a plateau and begins gradually to decline in middle age, thus paralleling physical structures and functions such as brain weight and vital capacity.

Crystallized intelligence consists of learned knowledge and skills. It has been characterized as a "precipitate out of experience" – the resultant of the interaction of the individual's fluid intelligence and his culture. It increases throughout most of a person's life, depending upon the amount of his fluid intelligence and his opportunities for learning and new experience. From an operational standpoint, the difference between fluid and crystallized intelligence really amounts to the difference between culture-fair and culture-loaded tests.

Levels I and II are seen as being essentially orthogonal to fluid and crystallized intelligence. While many of the tests that characterize Level I processes, such as digit span, are also those that characterize tests of fluid intelligence, not all tests of fluid intelligence are confined to Level I functions. The *Progressive Matrices* and *Culture-Fair Tests*, for instance, are tests of fluid intelligence and are also among the best measures of Level II ability.

4. Relationship of Socioeconomic Status to Levels I and II

As shown in Fig. 3, individual differences in Level I and Level II abilities are hypothesized as having different distributions as a function of SES. The distribution of Level I abilities is shown as independent of SES. This may or may not, in fact, be true, but so far we have found little or no evidence that would contradict this simple assumption. When large, truly random samples of the population are tested however, it should not be surprising to find some difference between SES groups in the distribution of Level I abilities, especially in adults and in children beyond 8 to 10 years of age, for two reasons: (*a*) because of the hierarchical (but not complete) dependence of Level II on Level I ability we should expect assortative mating to affect gene pools for Level I in a manner similar to Level II, though to a much lesser degree, and (*b*) beyond 8 or

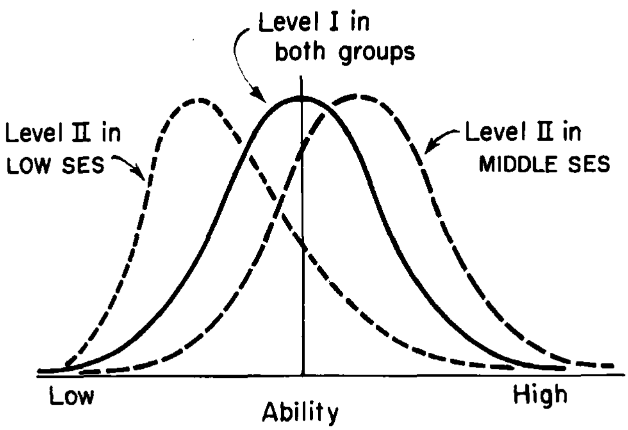

Fig. 3. Hypothetical distributions of Level I (solid line) and Level II (dashed line) abilities in middle-class and lower-class populations.

10 years of age, when both Level I and Level II processes are already clearly established in children's intellectual performance, it seems doubtful that Level II functions would not enter into performance on tasks that are intended as predominantly Level I, especially for children who are well-endowed in Level II ability. When performance on a Level I task is further facilitated by bringing Level II processes to bear upon it, upper SES children will show an advantage over lower SES children even in Level I tasks. Provided a sufficient number of different Level I and Level II tasks have been administered, factor analysis can aid in distinguishing the extent of involvement of Level I and Level II processes in the various tests, and factor scores representing Level I and Level II should show greater differences between lower and upper SES groups for the Level II factor and smaller differences for the Level I factor.

Why should Level II ability be different in upper and lower classes, while Level I is hypothesized as having little if any relationship to SES? One of the main factors determining an individual's SES is occupation or the occupation of the spouse. Occupation in turn is related to the individual's ability and educational attainments. Scholastic performance under traditional methods of instruction is heavily dependent upon Level II abilities. This is mainly why IQ tests, which were expressly devised to predict scholastic performance, are largely measures of Level II ability. Since individuals select mates of similar education and occupational status, the genetic component of Level II becomes segregated in the population. The greater the social mobility that is permitted by the society, the greater will be the segregation of genetic factors associated with social mobility, the chief factors in which are educational and occupational attainments in modern industrial society. In the course of generations there will be a gradual elimination of genetic factors making for poor Level II ability in the upper classes. Also, since there is some dependence of Level II upon Level I ability, low grades of Level I ability would also tend to be eliminated from the upper classes. In lower SES groups, on the other hand, education is not the chief means of succeeding, and small demands are made on abstract, conceptual ability, that is, the Level II processes. Level I abilities, however, are required to succeed in many manual occupations, and others' perception of the individual's intelligence or "wits" is based largely on his Level I ability when indices of scholastic attainments are lacking, are not valued, or are more or less uniformly meager among members of the group. In such cases, assortative mating will take place in terms of practical intelligence, "wits," cleverness, shrewdness, and the like. The Negro vernacular has its own term for this kind of intelligence: "mother wit."

High Level I ability is of value in any society or walk of life, and in primitive cultures it is probably of much more importance to survival than Level II ability. When there is little or no division of labor, except by sex role, every individual needs the ability to learn a large variety of facts and practical skills in order to fulfill his adult role in the society. Therefore there should be positive selection for Level I ability in all strata of all societies. The only condition under which one might expect a diminution of selection against low Level I ability is under circumstances in which no significant economic disadvantage is attached to relative inability to compete and in which vocational ineptitude is no barrier to mating, as might be the case when a society assumes complete support of its least able members and takes no measures to reduce their fecundity.

5. Levels I and II and the Focus of Attention

Rimland (1964), in his book on *Infantile Autism*, proposed a two-factor theory of mental functioning which bears considerable resemblance to the present distinction between Levels I and II. Rimland conceives of this difference as having to do largely with the focus of attention. He postulates that the brain contains a mechanism which focuses attention in a manner analogous to the operation of certain kinds of electronic equipment. His information-theory model of this aspect of brain function states, simply, that there is ordinarily a trade-off between fidelity and bandwidth in human attention. According to Rimland, the bandwidth aspect of mental functioning corresponds to Level II. It permits the individual to view, attend to, and recall specific experiences with respect to a larger context of associations, generalizations, and broad transfer from other experiences, to see differences and similarities between situations, and therefore to be able to deal with abstractions. "Fidelity," corresponding to Level I, permits an individual to deal in detail with the immediately given physical attributes of stimuli. Rimland believes that persons are capable of trading-off fidelity for bandwidth in their cognitive contact with the world, but each person has his own modal configuration of these capacities which characterizes his cognitive style and his pattern of mental capabilities. Rimland believes that persons whose main strength is Level I, or fidelity-reproductive processes, have a focus of attention that is largely *extracerebral*, that is, focused on real-world events taking place in the here and now of the person's environment. Such persons learn mainly by looking and doing. Unless they are also high in Level II, they are at a disadvantage in the traditional academic realm, which depends heavily upon learning from symbolic or abstract representations in the form of lectures and books. The person whose major strength is Level II, in contrast, directs more of his attention to intracerebral events a good part of the time. In the extreme, such individuals can become "lost in thought," which can at times put the individual at a disadvantage in dealing with many of the immediate exigencies of practical life. For example, it was said of Ernest O. Lawrence, the Nobel Prize-winning inventor of the cyclotron, that his tendency to become "lost in thought" while driving his car made him an unsafe driver to such an extent that he found it necessary to employ a chauffeur to drive him to and from work.

An important feature of Rimland's (1964) formulation of a two-process theory of cognitive functioning is that he cites cases in which Level II is almost entirely lacking despite apparently very superior Level I functioning, as found in some autistic children and so-called idiot savants. These observations support the notion that quite distinct brain processes are involved in these two types of ability, and thus they cannot be conceived of as simply different parts of a single underlying continuum of general mental ability. Just the opposite condition is found in Korsakoff's syndrome, in which some but not all Level I functions, such as the consolidation of short-term memory traces, are markedly deficient, although the victim retains the ability for normal performance on Level II tests (Talland, 1965).

B. Correlation between Level I and Level II

At present our hypothesis regards individual differences in Level I and Level II abilities as uncorrelated genotypically (i.e., in terms of their underlying mechanisms) but correlated phenotypically, because Level II functions have some degree of hierarchical dependence on Level I. [For example, solving an orally presented "thought problem" in arithmetic involves Level II, but also requires that the subject have sufficient short-term memory (Level I) to retain the elements of the problem in mind long enough to solve it. It is possible to retain the problem in mind without being able to solve it, but the reverse cannot be true.]

Tests of Level I and Level II, should, according to our hypothesis, produce correlation scatter diagrams like those shown in an exaggerated clear-cut form in Fig. 4. Level I is represented by tests of associative learning ability and Level II by intelligence tests with a high *g* loading. Because low Level II ability is not a crucial disadvantage in the lower SES groups, there is not much selection against it, while it tends to be eliminated from the upper SES groups. Thus the scatter diagrams for lower and upper SES groups differ mostly in the proportion of persons falling into the upper left quadrant. Because of the dependence of Level II on Level I in actual test performance, few if any authentic cases should be found in the lower right quadrant of either SES group. But if there is some fairly low threshold value of Level I above which any amount of

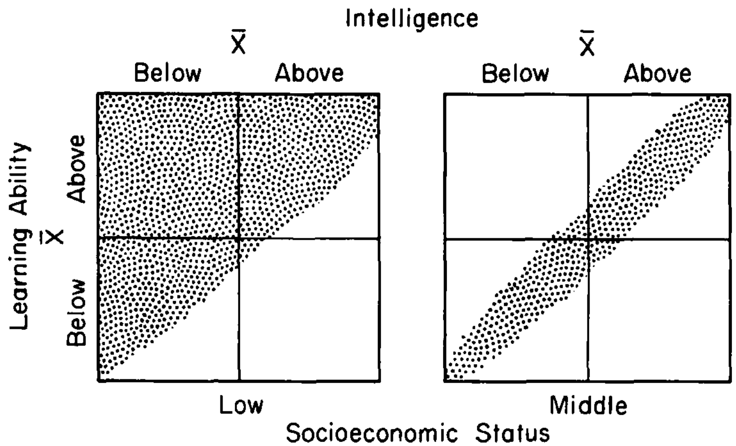

FIG. 4. Schematic illustration of the form of the correlation scatter-diagram for the relationship between associative learning ability and IQ in low SES and Upper-Middle SES groups.

Level II can be fully manifested, there may be more cases in the lower right quadrant than is depicted in Fig. 4. So far we have not found individuals who are superior in Level II tests and are also authentically deficient in Level I abilities. A few pseudo-deficient Level I cases with high IQs seem to be due to some fluke in the Level I testing, such as failure to understand instructions, excessive anxiety in the laboratory testing situation, etc. However, older brain-damaged and senile subjects could very probably be found in the lower right quadrant of the scatter diagram.

The hypothesized characteristics of the scatter diagram for lower and for upper SES groups implies much higher correlations between tests of Level I and Level II in high than in low SES groups. In fact, it was the finding of this difference in correlations between learning tests and IQ tests for lower and upper SES groups that initially prompted the formulation of this dual-process hypothesis of cognitive functioning.

1. Hypothetical Growth Curves of Levels I and II as a Function of SES

These are shown in Fig. 5. Since most of the child's behavioral development up to about 4 years of age is attributable, according to this hypothesis, to the growth of Level I, and since SES groups do not differ appreciably in Level I, there should be little or no differences between SES groups in early childhood. Children who appear retarded during this early stage of development are regarded as very probably retarded in Level I ability. If the degree of retardation is only slight, and if the child possesses normal or superior Level II ability, he will appear to be a "late bloomer" and during the early school years will come up to par intellectually. Thus, there is a near zero correlation (in fact, a low *negative* correlation for boys) between indices of early development and later IQ.

Figure 5 also illustrates a possible basis for the so-called "cumulative deficit" generally found in low SES children, that is, the fact that scholastically they tend to lag further and further behind their middle-class age mates as they go through school. As the content of the school's curriculum becomes increasingly abstract and conceptual with advancing grades, the child with below-average Level II ability, regardless of his status on Level I, will be at an increasing disadvantage. The cumulative deficit effect will then snowball because of the child's discouraging experience of diminishing returns from his efforts in school. The most important reinforcement in school learning is probably the student's perception of his own success and progress in learning, and when this reinforcement diminishes, the child is, in effect, on an extinction schedule with respect to the behaviors involved in classroom learning. This results in some children's appearing to be unable to learn even the simplest things taught in the classroom, despite the fact that outside the classroom they may learn more difficult things quite readily. Such extinction of school learning behavior could probably be prevented by conducting instruction in the basic school subjects more in accord with Level I processes rather than by means of techniques that maximize the role of Level II abilities in classroom instruction.

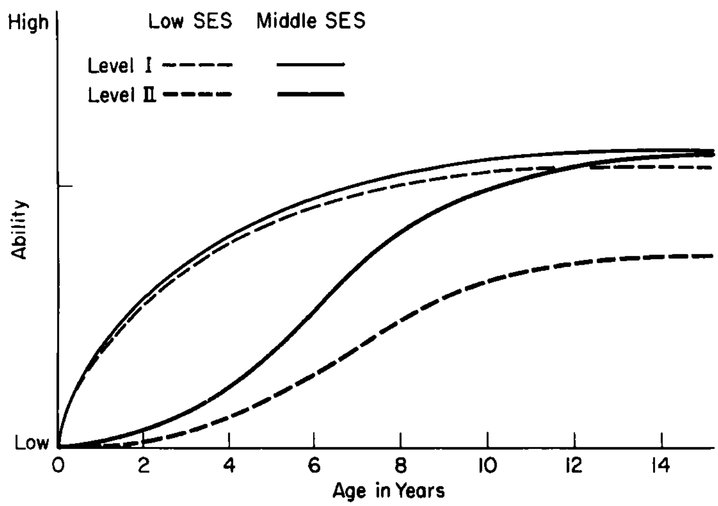

FIG. 5. Hypothetical growth curves for Level I and Level II abilities in middle SES and low SES populations.

2. The Heredity-Environment Aspects of Levels I and II

The previous review of the genetic aspect of mental retardation and of SES differences in intelligence bears directly on the question of the sources of individual differences in Levels I and II. Those who argue from the cultural deprivation hypothesis of SES intelligence differences would claim that Level I tests reflect more nearly the individual's genetic potential, and that tests of Level II reflect the individual's cultural acquisitions. According to this view, the basic source of individual differences in mental ability is seen as consisting of Level I processes, while Level II processes are regarded as the resultant of the interaction of the individual's Level I processes and the opportunities for learning afforded by his environment.

The present theory, on the other hand, postulates separate genetic mechanisms for Level I and Level II abilities. Although the development and manifestation in performance of Level II abilities doubtless depends upon experience and learning (the capability for much of which, in turn, depends upon Level I), experience and learning are regarded as *necessary but not sufficient* for the development of Level II. The idea that individual differences in Level II ability are largely determined by environmental factors, even granted a largely genetic determination of Level I, is contradicted by the evidence on the inheritance of intelligence, most of which is based upon tests that largely measure Level II functions. The purest Level II tests, such as Raven's *Progressive Matrices*, yield heritability estimates as high or higher than are found for omnibus intelligence tests like the Stanford-Binet (e.g., Shields, 1962). There have been no comparable studies of the heritability of Level I per se, but there is no reason to believe that Level I abilities are not fully as heritable as Level II. For example, pursuit-rotor learning—a form of perceptual-motor learning in which the subject practices keeping a stylus on a moving metal disc (or "target")—would seem to be a relatively pure type of Level I ability. Analysis of the correlations between sets of identical and fraternal twins for total time "on target" in the course of acquiring the pursuit-rotor skill yielded a heritability coefficient of .88, which is close to the heritability of physical stature (Bilodeau, 1966, Ch. 3).

C. Relationship of Levels I and II to Mental Retardation

Severe grades of mental defect due to mutant genes, chromosomal abnormalities, and brain damage probably always involve a marked deficiency in Level I; consequently Level II will also be deficient. Even in the severely retarded, however, the most elemental Level I functions are often prominent, such as high-fidelity transmission of stimulus inputs as commonly seen in the echolalia and echopraxia of imbecile children—in many cases these are their only signs of learned behavior, a high-fidelity "echoing" of what they see and hear (O'Connor & Hermelin, 1961). But here we are not primarily concerned with this category of severe mental deficiency. Rather, our present concern is with the milder forms of mental retardation associated with normal polygenic inheritance and due to the fact that polygenic characteristics assume a "normal" distribution of values in the population and such a distribution has a lower "tail." We have postulated two such distributions representing different genetically conditioned aspects of mental development: Level I and Level II. Because there are two underlying distributions, there are theoretically three ways that an individual can be retarded, but phenotypically two of these three "types" may look much alike from the standpoint of diagnosis. An individual may be diagnosed as retarded because (*a*) he is low on Level I but not on Level II; or because (*b*) he is low on Level II but not on Level I; or because (*c*) he is low on both Level I and Level II. Individuals in the categories *a* and *c* are probably the least distinguishable in performance and at present we do not know any means for clearly differentiating these groups, since normal Level II ability seems not to be manifested when Level I is very low.

Primary retardation here refers to a deficiency in Level I. *Secondary retardation* refers to a deficiency in Level II. This diagnostic distinction, we believe, has important implications for education and for occupational selection and training. While retardation generally refers to individuals who are more than two standard deviations below the general population mean on conventional IQ tests, there is a substantial segment of the population, largely among the groups now called culturally disadvantaged, who fall in the IQ range from 70 to 85 and might be regarded as of "borderline" intellectual ability in terms of conventional test scores and scholastic performance. The *primary* versus *secondary* distinction would seem especially important with respect to this group. Approximately half the Negro population of the United States, for example, is below IQ 85 on standardized tests, and approximately six times as many Negroes as whites are classified as mentally retarded by traditional criteria (Shuey, 1966). We do not know what proportions are below the average range in the *primary* or in the *secondary* sense, but from the evidence we have gathered so far, it appears that comparatively little of the intellectual retardation found in low SES groups is of the primary type. It is unfortunate that the label "retarded" is ever used in connection with individuals who are of average ability in Level I processes although they are quite far below average in Level II. Most such individuals are not perceived as retarded once they leave school, and, unless they show emotional instability or other severe behavior problems, they do not become institutionalized. Accurately speaking, they are not "slow-learners." Neither is their particular pattern of abilities primarily the result of cultural deprivation, in the majority of individuals. Some children with exceptionally high Level II ability come from a culturally deprived background (for some striking examples, see Burt, 1961b). Barrett, a student of mental retardation, has stated that "Perhaps the major obstacle to analysis and habilitation of retarded behavior is the paucity of measurement methods that amplify rather than homogenize the parameters of individual behavior [Barrett, undated, p. 16]." The differential assessment of Level I and Level II abilities is a step toward the more refined diagnosis of familial retardation, and it is a diagnostic approach based on a theoretical conception of the development and structure of mental abilities.

IV. EVIDENCE FOR THE LEVEL I–LEVEL II HYPOTHESIS

A. General Observations

The observations that initially gave rise to the studies that led to the dual-process hypothesis proposed here were brought to the writer's attention by school psychologists and teachers in classes for the educable mentally retarded (EMR, with Stanford-Binet IQs between 50 and 75) in schools that contained a large proportion of children called culturally disadvantaged. It was the teacher's impression, confirmed by the writer's own observations made in the classroom, on the playgrounds, and in laboratory testing, that low SES children in the EMR groups appeared in many ways to be much less retarded, and in fact usually appeared quite normal, as compared with middle-class children of the same IQ, even excluding those with sensorimotor disabilities or signs of neurological impairment. The same held true in observations of children not in EMR classes but in the "slow learner" category of IQs from 75 to 85 or 90. The low SES children, whether white, Negro, or Mexican-American, appeared more mature and capable in social interactions and in activities on the playground than middle SES children, despite very similar scores on a variety of intelligence tests, both verbal and nonverbal, and very similar performance in school subjects such as reading and arithmetic.

We found it possible to devise special tests, which we call "direct learning tests," that measure how fast the child could learn something new right in the test situation itself. Such tests are much less tests of achievement than the ordinary intelligence tests. Direct learning tests depended relatively little on knowledge or specific skills that have been acquired prior to being tested. The "direct learning tests" consist of measures of short-term memory and rote associative learning; they minimized conceptual learning. In brief, it was found that the low SES children in EMR classes and in the IQ range from 75 to 85 performed on the average much better on these learning tests than their middle-class counterparts of similar IQ. Low SES children of average or above average IQ, however, were found not to perform any differently on the learning tests than middle SES children of the same IQ. This finding suggested that the low SES versus middle SES difference was not simply due to the IQ tests' being more culturally loaded than our learning tests, such that the IQ underestimated the intelligence of the low SES group. It appeared that two different kinds of ability were being assessed—associative learning abilities, to which we later gave the general label of Level I, and conceptual or cognitive abilities, which we have labeled Level II. The typical results of several of these studies are summarized by Fig. 6.

It later became apparent that selecting subjects only from EMR classes actually biased our experimental results *against* the hypothesis. In many schools in low SES neighborhoods, it was found that the majority of children with IQs in the 50 to 75 range are not found in EMR classes but are in the regular classes, although their scholastic achievement is usually commensurate with their low IQs. The low SES children who are placed

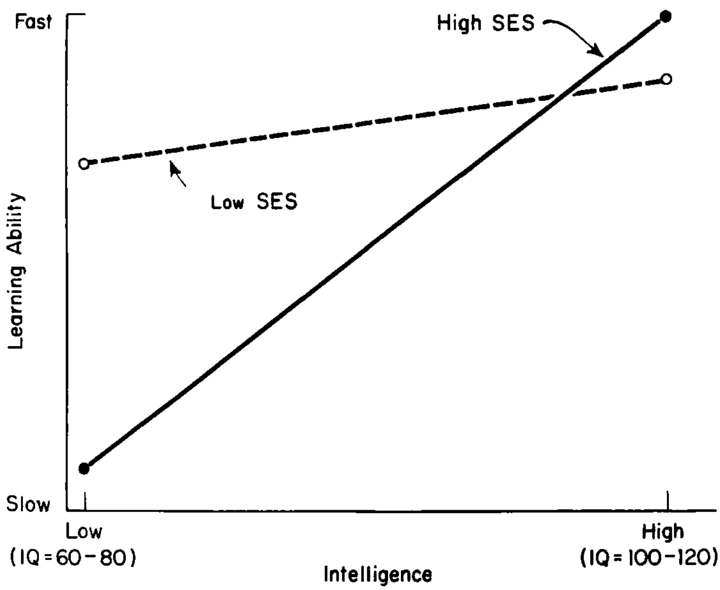

FIG. 6. Summary graph of a number of studies showing relationship between learning ability (free recall, serial and paired-associate learning) and IQ as a function of socioeconomic status (SES).

in EMR classes are more likely to resemble middle SES children of the same IQ than are low SES children in the regular classes despite IQs in the EMR range. On the other hand, we have found no middle SES children with IQs between 50 and 75 in regular classes. When such children are found, they are in the special EMR classes. The great majority of low SES children in regular classes but with low IQs and with scholastic achievement 2 or 3 years below grade level perform in the same average range as the majority of average IQ middle SES children on our Level I learning tests.

The literature on mental retardation frequently notes that many retardates are regarded as retarded only during their school years and make a normal social and vocational adjustment once they are out of school. From then on most are rarely perceived as retarded (Robinson & Robinson, 1965; Tyler, 1965, pp. 370–377). Only a small minority of individuals diagnosed as retarded while in school are ever placed in institutions or sheltered workshops for the retarded.

We have tested institutionalized familial retardates, as well as those in sheltered workshops, on some of our direct learning tests. We find that almost without exception these individuals are as deficient on our learning tests as on conventional IQ tests, and this is true even when we rule out individuals with any suspicion of organic impairment. (Retardation due to single gene and chromosomal defects has never formally entered into our research, but the several such cases that have been tested showed marked deficiency on the Level I tests.) It seems clear that among groups diagnosed as familial retarded, especially when social incompetence is part of the diagnostic criterion, there is a preponderance of primary retardation.

There is an indication that primary and secondary retardation can exist in different siblings reared together in the same family. Barnett (undated) studied four brothers, 8 to 14 years of age, diagnosed as familial retarded, with both parents also retarded, in an instrumental discrimination learning situation. Instrumental learning clearly qualifies as a Level I process. Two of the brothers (IQs 72 and 55) were grossly superior to the other two (IQs 63 and 48) in instrumental learning. One of the brothers (IQ 72), in fact, performed like a normal adult. All were markedly retarded in school work, although the two showing the better instrumental conditioning were also somewhat better in scholastic performance.

B. Psychometric Evidence

1. MA, IQ, and Cognitive Development and Learning Rate

As illustrated in Fig. 5, different developmental curves are hypothesized for Level I and Level II processes, with Level II becoming increasingly prominent beyond the preschool years. Mental Age (MA), as derived from tests such as the Stanford-Binet, is an index of the individual's status in this form of cognitive development. But it is also an index of the amount of learning, as represented by the acquisition of knowledge and skills, that has taken place up to the chronological age at which the child is tested. Some part of this knowledge acquisition depends mainly on the child's associative learning ability, which is a Level I process. Thus, MA is a composite index representing both cognitive developmental status and amount of learning. The IQ, being a ratio of MA/CA, is an index of the rate of cognitive development and of the rate of learning. Culture-fair tests tap cognitive development more than learning.

2. Heterogeneity of Familial Retardation

If the relationship between Level I and Level II performance is as shown in the correlation scatter diagrams in Fig. 4, we should expect to find greater heterogeneity in associative learning abilities among a group of retarded than among average or gifted children, even though all three groups have much the same variance on the IQ (or Level II) measure. Jensen (1963) tested all the children in EMR classes (IQs 50–75) in an urban junior high school on a trial-and-error selective learning task and compared their performance with representative samples of average (IQs 90–110) and gifted children (IQs 135 and above) in the same school. The groups all differed significantly from one another, in the expected direction. But the most striking finding was the extreme hetergeneity of the EMR group on the learning task. Although the standard deviation of their IQs was 7.13 as compared with 8.06 for the average and 4.94 for the gifted, the EMR's variance on various trial and error selective learning tests was from 2 to 5 times greater than the variance of the average group, and from 10 to 25 times greater than the variance of the gifted group. Several of the EMR children performed above the mean level of the gifted group. Interestingly enough, the two fastest learners in the study had IQs of 147 and 65! On the other hand, none of the average or gifted subjects had scores as low as the mean for the retarded. None of the gifted, in fact, was below the mean of the average group. These results are highly consistent with our dual process formulation. Virtually the full range of Level I ability was found among the EMR, though all were deficient in Level II. Also, the lowest part of the range of Level I ability was not found in the average and gifted IQ groups.

If (*a*) there are two underlying ability distributions, Level I and Level II, and if (*b*) omnibus intelligence tests like the Stanford-Binet contain items that measure both Levels to some extent, and if (*c*) one distribution (Level II) but not the other (Level I) is correlated with SES, then we should predict an *increase* in the population variance and an increase in the mean SES difference on tests which are more pure measures of Level II. This is exactly what Cattell (1934) found with a "culture-fair" measure of *g*, a test which taps Level II almost exclusively. When IQ is derived from Cattell's test in the same fashion that it is derived from the Stanford-Binet, by taking MA/CA, the standard deviation of the Cattell test is 50% greater than for the Stanford-Binet (i.e., 24 vs. 16), and SES IQ differences are greatly magnified by the Cattell test, despite the fact that it contains much less cultural content than the Stanford-Binet. This would be expected from our hypothesis.

A similar finding is that of Higgins and Sivers (1958), who found that large groups of 7- to 9-year-old low SES Negro and white children who did not differ on Stanford-Binet IQ showed a significant difference, with Negroes scoring lower, on Raven's *Colored Progressive Matrices*, a relatively pure test of *g* or abstract reasoning. Sperrazzo and Wilkins (1958, 1959) (also see Jensen, 1959) found similar Negro-white differences in each of three subgroups on the SES scale.

The Porteus mazes test, often regarded as one of the most culture-free tests and recognized for its sensitivity to brain damage, appears to be more a test of Level I processes than of *g* or Level II. The test apparently correlates with other intelligence tests because of their partial dependence on Level I functions, not because it measures Level II functions directly. Its lack of loading on Level II makes it particularly suited to distinguishing primary and secondary familial retardation, as shown in a study by Cooper, York, Daston, and Adams (1967). They were led to the use of the Porteus test by their impression that the Wechsler and Stanford-Binet tests often result in misleading and erroneous decisions when applied to a population of lower-class Southern Negro adolescents. They state:

> We were first led to question these procedures through observations of Southern Negro adolescents committed to a state institution for the mentally retarded. In the judgment of their teachers, nurses, social workers, and attendants a substantial number of these adolescents were functioning socially and vocationally at levels far above those to be expected of persons mentally retarded.

They point out that "extended retesting [on Wechsler and Stanford-Binet] failed to produce any reliable discrimination between the adolescents who appeared behaviorally nonretarded and those who were grossly deficient in effective and adaptive social behavior." Here, then, appears to be a clear-cut example of the failure of IQ tests, which tap mainly Level II, to discriminate between primary and secondary retardation. The Porteus test apparently made this discrimination. Subjects were divided into 2 groups—those for whom judges gave the answer "yes" to 6 or more of the following questions and those for whom they answered "no" to 6 or more:

Is he socially alert?
Is he socially effective?
Is his general activity level high?
Is he mentioned more often?
Is his vocational ability high?
Does he have sports ability?
Is his physical appearance good?
Is his social judgment accurate?

Although these 2 groups had mean Wechsler IQs of 56.0 and 63.1, respectively, their mean IQs on the Porteus were 63.6 and 121.7. None of the primary retardates scored above 84 on the Porteus and none of the secondary retardates scored below 102; the highest scored 132.

C. Memory Span

Tests of immediate memory span are among the best indices of Level I ability.

Memory span for digits has been underrated as a psychometric test by most clinical psychologists. The main reasons for the depreciation of the digit span test as it is generally used by clinicians are (*a*) its relatively low reliability as compared with most other subtests, and (*b*) the fact that in some cases it yields results that are highly discrepant from other subtests, as when a person with a very low IQ obtains an average or superior score on digit span. Poor performance on digit span, however, is rarely found in persons of average or superior IQ, unless there is evidence of extreme anxiety, an organic brain condition, or other pathologic disturbance. Wechsler (1958) has stated that "Except in cases of special defects or organic disease, adults who *cannot* retain 5 digits forward and 3 backward will be found, in 9 cases out of 10, to be feeble-minded or mentally disturbed [p. 71]." He adds, "Rote memory more than any other capacity seems to be one of those abilities of which a certain absolute minimum is required, but excesses of which seemingly contribute relatively little to the capacities of the individual as a whole." This view probably underrates the importance of individual differences in the ability assessed by digit span in the region above the minimum requirement Wechsler speaks of.

The relationship of memory span to general intelligence is actually greater than is generally believed. Memory span for digits formed a part of the original Binet intelligence scale and has been included in all the revisions of the test. It is also among the subtests of the Wechsler Adult Intelligence Scale (WAIS) and the Wechsler Intelligence Scale for Children (WISC). The low reliability of the very brief digit span (DS) test as used in these batteries is probably what misled Wechsler to state that ". . . as a test of general intelligence it [digit span] is among the poorest [Wechsler, 1958, p. 70]." This statement, however, is belied by the massive normative data presented in Wechsler's own book.

First of all, it must be noted that the reliability of the DS test of the WAIS is between .66 and .71 for various age groups. The WISC Manual reports DS reliabilities between .50 and .60 for various age groups (Wechsler, 1949). By comparison, the reliability of the Full Scale IQ on both the WAIS and the WISC is between .92 and .97. Vocabulary has the highest reliability (.95) of any of the single scales. But low reliability is no real problem with the DS test. Its reliability can be boosted to any desired level simply by increasing the number of series presented. It also helps to standardize the procedure as much as possible, by presenting the digits at a metronomic 1 second rate by means of a tape recording for auditory digit span or an automatic projector for visual digit span. We obtain reliabilities above .90 under these conditions, and a reliability as high as .96 has been obtained even among a relatively homogeneous group of university students.

The correlation between DS and Full Scale IQ (minus DS) on the WISC, after correction for attention, ranges between .60 and .70, and for the WAIS it is .75. These correlations compare favorably with those of other individual scales after they are corrected for attenuation. The ability to repeat two digits at age 2½ correlates .62 with Stanford-Binet IQ at that age (Terman & Merrill, 1960, p. 342).

Of further interest is Wechsler's claim that DS correlates very little with *g*, the general factor common to all the WAIS subtests. Yet Wechsler (1958, p. 122) presents a factor analysis (Holzinger's bi-factor method) of the WAIS in which a large *g* factor, accounting for some 50% of the total variance, was extracted. The DS test has a loading of .63 on *g* in the age group 18–19, which is the peak age for DS performance. Corrected for attenuation, this factor loading becomes approximately .80, which is a very substantial loading as compared with the *g* loadings of other subscales. Wechsler's notion that DS ceases to correlate significantly with other measures of intelligence once DS exceeds a certain minimal threshold would seem to be further belied by the correlation of .60 (.73 corrected for attenuation) between the DS and Vocabulary subtests of the WAIS in the normative population. It appears that seemingly small individual differences in immediate memory span, when multiplied over a lifetime of experiences, make for highly significant differences in such acquired indices of intelligence as vocabulary. A person with good short-term memory span plus rapid consolidation of the memory traces would learn more per unit of time from his experience than a person with a shorter span or slower trace consolidation. This seems a reasonable explanation for the substantial correlation between DS and Vocabulary in Wechsler's normative population. Another line of evidence that rote memory abilities do not cease to be important above a minimal threshold was obtained by Jensen (1965b), who derived 12 factor scores from a battery of memory span and serial rote learning tasks administered to university students. The multiple correlation between the 12 factors and students' college grade point average was .76 (.68 after correction for shrinkage).

The reader should not gain the impression that memory span is a unitary ability. There is ample evidence, for example, that the abilities to repeat digits forward and backward are not entirely the same. Korsakoff patients, for instance, show far greater than the normal discrepancy between forward and backward digit span (Wechsler, 1958, p. 71). And factor analyses of the intercorrelations among a variety of tests including forward and backward span have shown that they have different factorial compositions (Jensen, 1965b; Osborne, 1966). From these analyses repeating digits forward can be interpreted as an almost pure measure of Level I ability, while repeating digits backward involves some Level II ability. This is in line with the fact that backward span calls for a transformation of the input, which brings some Level II elements into play. Forward digit span, for example, correlates more with the WISC Information subtest than with Arithmetic "thought" problems, while backward digit span is just the opposite. Also, backward digit span is more highly correlated with Block Design than is forward digit span, and Block Design is the best measure of *g* among the Performance tests.

Other procedural variations of the digit span task, such as requiring a 10-second delay between presentation and recall of the digit series, introduce further individual differences factors. Subjects do not remain in the same rank-order of ability on immediate and delayed recall (Jensen, 1965b).

The argument that digit span is positively correlated with IQ mainly because more intelligent subjects are capable of more sophisticated strategies for encoding strings of digits is not very convincing. For one thing, digit span correlates at least as highly with IQ at 2½ years of age as at any later ages. Furthermore, digit span reaches a peak at around 19–20 years of age and shows a relatively early gradual decline, following much the same curve as brain weight and vital capacity. This seems hard to account for in terms of conscious strategies for remembering digits. It is more likely that digit span is closely tied to very basic brain functions. Intensive training of digit span ability has been shown not to produce any permanent increase in children's digit span over what would be normal for their mental age (Gates & Taylor, 1925).

1. Short-term Memory and Retardation

Ellis (1963) has proposed the hypothesis that the mentally retarded are essentially characterized by a deficit in short-term memory (STM). He has postulated that the retardate is deficient in both the strength and duration of the stimulus trace. There is considerable support for this theory, most of it based on studies of institutionalized retardates. The position of the present paper is that Ellis' theory applies only to primary retardation as here defined. It is hypothesized that secondary retardation does not involve a STM deficit but depends upon a specific deficiency in Level II, i.e., abstract and conceptual processes. We also believe that the majority of low SES children with IQs in the range from 50 to 85 are intellectually retarded only in the secondary sense and do not evince a STM deficit.

2. Interaction of Digit Span, IQ, and SES

We have found that the substantial correlation between DS and IQ in the normative population of the Wechsler and Stanford-Binet intelligence tests breaks down completely in low SES segments of the population (Jensen, 1968b). The reason for the low or negligible correlation between DS and IQ in low SES groups is attributable, according to our theory, to a deficiency in Level II mechanisms. We hypothesize that there is too little variance in Level II potential in low SES groups for even quite large individual differences in Level I to make any substantial difference in tests of Level II.

If digit span correlated as highly with IQ in the low SES population as it does in the middle-class population, we could claim to have a culture-free test of general intelligence in the form of digit span. But we have found that DS and IQ are much less correlated in low than in middle SES groups. The fact that the low correlation in the low SES group is found even for the most status-fair tests, such as the *Progressive Matrices*, indicates that the phenomenon we are observing is not a result of DS and IQ differing in culture-fairness, but rather is a result of their measuring quite different mental abilities.

In one study (Jensen, 1968b), children from grades 4 to 6 in an all-Negro school in a low SES neighborhood and children in an all-white school in an upper-middle-class suburban neighborhood were given an auditory digit span test and Raven's *Colored Progressive Matrices.* (The mean IQ difference between the two schools is approximately 2 standard deviations.) The nonparametric correlation (phi coefficient) between digit span and *Progressive Matrices* was 0.33 for the low SES ($N = 60$) and 0.73 for the upper-middle SES ($N = 60$). The idea that STM as indexed by DS may be necessary but is certainly not sufficient for performance on a highly *g*-loaded test such as the *Progressive Matrices* is supported by a comparison of the 30 *highest*-scoring children on DS in the Negro ghetto school (the upper 7.9% in DS in grades 4, 5, 6) with the 30 *lowest*-scoring children on DS in the white suburban school (the lower 6.1% in DS in grades 4, 5, 6). The mean DS scores (expressed as percent of the maximum possible score) were 65.3 for the ghetto group and 38.7 for the suburban group. Yet the corresponding *Progressive Matrices* scores (expressed as percent of possible maximum score) were 64.7 and 72.6, respectively.

A more detailed analysis of auditory digit memory in relation to IQ in low and high SES groups was performed on groups of preschool children between 3 and 5 years of age. The low SES group ($N = 100$) was predominantly Negro children attending day-care centers; in all cases their parents were receiving public welfare assistance. The upper-middle SES group ($N = 100$) was composed of white children in private nursery schools. The mean ages of the high and low SES groups were 50 and 52 months, respectively. All the children were administered a battery of tests composed of auditory digit series of from 2 to 9 digits, the Binet and Wechsler digit span tests, serial and paired-associate learning of pictures of common objects, and the Peabody Picture Vocabulary Test (PPVT). The various tests yielded 26 variables in all. The intercorrelations among the variables were factor analyzed (i.e., a varimax rotation of the 5 principal components having Eigenvalues greater than 1) separately for the low and high SES groups. The results of the factor analysis were quite different for the two groups. Although the groups differed by 19 points in PPVT IQ (an average mental age difference of 16 months), they showed no appreciable differences in the digit span and serial and paired-associate learning tests. The pattern of intercorrelations among tests differed, however, in the low and the high SES groups, and these differences were, of course, reflected in the factor analyses. In the high SES group a single factor accounted for most of the variance on all the tests; the intelligence test and the digit series and learning tests were all substantially intercorrelated, yielding a large general factor common to all. In the low SES group, on the other hand, there was a clear separation of the intelligence factor from the factor representing the digit series and learning tests.

The results are shown in Table VI. It is especially instructive to examine the intelligence factor in detail. The intelligence factor is so defined because it is the only factor with a high loading on PPVT mental age. Digit span on both the Binet and Wechsler is defined as the longest series of digits the subject can recall perfectly (after a single auditory presentation at a rate of 1 second per digit) on 50% of the trials. As shown in Fig. 6, the low and high SES groups do not differ significantly in means or standard deviations on either the Binet or the Wechsler digit span tests, despite a 16 months difference between the mean mental ages of the groups. Also note that DS has nonsignificant loadings on the intelligence factor in the low SES group and very substantial loadings in the high SES group.

The digit series test, comprised of series of from 2 to 9 digits, were administered in the same manner as the DS test from the Binet and Wechsler, but they are scored differently. Two different scores were obtained. The *position* (Pos.) score is the number of digits recalled in the correct absolute position. The *sequence* (Seq.) score is the number of digits correct in forward adjacent sequence, regardless of absolute position. Since the maximum possible sequence score is necessarily 1 less than the maximum possible position score for a given series length, +1 is added to

TABLE VI

MEANS, STANDARD DEVIATIONS, AND CORRELATIONS WITH INTELLIGENCE FACTOR IN LOW AND HIGH SOCIOECONOMIC GROUPS ($N = 100$ IN EACH GROUP)[a]

	Mean				Standard deviation				Factor loadings			
Variable	low SES		high SES		low SES		high SES		low SES		high SES	
Mental age (mo.)	48.41		64.46		22.67		19.16		.504		.512	
Binet digit span	3.72		3.63		1.05		1.07		.047		.482	
WISC digit span	3.99		4.12		1.02		1.12		.073		.613	
	Pos.	Seq.	Pos.	Seq.	Pos.	Seq.	Pos.	Seq.	Pos.	Seq.	Pos.	Seq.
Digit series 2	1.99	1.99	1.99	1.99	.05	.05	.09	.05	.032	.032	.023	.023
3	2.82	2.85	2.88	2.91	.40	.31	.38	.29	.138	.181	.214	.210
4	3.06	3.20	3.02	3.13	1.13	.88	1.15	.95	.023	.010	.877	.870
5	2.00	2.46	1.83	2.42	1.32	.98	1.58	1.21	.157	.156	.563	.511
6	1.02	2.01	1.05	1.95	1.03	.83	1.03	.90	.340	.478	.372	.273
7	.54	1.53	.56	1.63	.65	.63	.84	.88	.325	.534	.072	.017
8	.41	1.66	.38	1.46	.49	.71	.60	.65	.138	.698	.057	.020
9	.26	1.71	.28	1.71	.37	.83	.49	.91	.148	.760	.133	.194

[a]Factor loadings significant beyond .001 level are underscored.

TABLE VII
CORRELATION BETWEEN POSITION AND SEQUENCE SCORING OF DIGIT SERIES TEST

	Series length							
SES	2	3	4	5	6	7	8	9
High	1.00	.98	.93	.93	.85	.60	.47	.39
Low	1.00	.95	.91	.90	.83	.29	.16	−.01

the sequence score to make it equivalent to the position score. The reason that the two types of scores were used is that it had been found in a previous study of digit memory in college students that in supraspan series (i.e., series lengths beyond the subject's memory span) the two scores cease to be highly correlated and apparently measure different factors (Jensen, 1965b). In supraspan series the subject seems to retain pair-wise associations between adjacent digits in the series rather than some mental representation of the series as a whole, in which absolute position is retained. Table VII shows the correlations between position and sequence scores for different series lengths.

Note again in Table VI that the low and high SES groups do not differ significantly in means or standard deviations on any series by either form of scoring. The loadings on the intelligence factor, however, are entirely different for the low and high SES groups. The low SES group has no appreciable loadings on any series for position scoring. The high SES group has very high loadings for series of 4 and 5 digits, which are the series lengths near the threshold of subjects' memory span at this age. For the high SES group the loadings are approximately the same for position and sequence scores. This is not so for the low SES group, which has its only sizeable digit series of lengths 7, 8, and 9, the clearly supraspan series which more or less force subjects to learn only adjacent associations. This strongly suggests that the intelligence test (PPVT) is measuring different mental processes in the high and low SES groups. It is hard to characterize psychologically the processes of the high SES group, but those of the low SES group appear to be of an associative nature, since their sequence scores are the only ones that correlate with the intelligence factor. These different patterns of correlations within the digit series tests would be most difficult to account for in terms of culture influences, especially in view of the fact that the distributions of scores in the low and high SES groups are indistinguishable. The different correlation patterns more likely reflect fundamental differences in neurological organization.

D. Associative Learning and Intelligence

Some of the most puzzling research in all of psychology is concerned with the relationship between psychometric intelligence and learning ability. An enormous range of correlations between various learning measures and intelligence test scores has been found, leading to a diversity of conclusions and disputes about the relationship between learning ability and intelligence (Rapier, 1962). Reviews of studies of learning ability in the mentally retarded show that this field is also characterized by similar conflicting findings (Goulet, 1968; Prehm, 1968; Zeaman & House, 1967).

Much of the puzzlement in the research findings is probably due to the failure, first, to distinguish between subjects on the basis of primary and secondary retardation and, second, to pay sufficient attention to the properties of the learning task with respect to its position on the Level I–Level II continuum. If one makes some judgment about whether the subjects of the study were predominantly primary or secondary retardates, and about whether the learning tasks were most heavily dependent on Level I or Level II processes, a considerable degree of order emerges from the various findings. For example, there is no disagreement among various researches that persons called retarded by any criteria are deficient on tests involving abstract and conceptual abilities. This characterizes both primary and secondary retardates. But as we get into the realm of associative learning tasks, the findings appear confusing, because it is in this type of learning that primary and secondary retardates show divergent abilities. The results will depend largely upon the proportions of primary and secondary retardates in the investigator's sample. If the subjects have IQs below 50, they will almost always be primary retardates, and the evidence is quite clear that these subjects are markedly below average in associative learning. If the subjects have IQs in the range 50 to 75 and are institutionalized, the chances are great that most of them are primary retardates, for we know that the vast majority of persons in this IQ range are never institutionalized. Thus, institutionalized subjects usually show a severe deficiency in learning ability. When the subjects are school children with IQs between 50 and 75 and are in special classes for the educable mentally retarded, there will be a considerable mixture of primary and secondary types of retardation, so that great variance will be found on rote learning tasks, and often the group's mean on such tasks will differ little from that of children with average IQs. When the subjects are children of low SES with IQs between 50 and 80, and are in regular classes, there will be little or no evidence of deficiency in associative learning as compared with the performance of middle-class children of average IQ.

Extremely simple forms of learning, which require no discriminations and involve no competition among multiple response alternatives—for example, classical conditioning—do not distinguish even between primary and secondary retardates or between retardates and persons of average or superior IQ. It is only when discriminative features enter the conditioning procedures that some correlation with intelligence is manifested (Zeaman & House, 1967, pp. 195–197).

In general, the evidence leads to the conclusion that there is a moderate correlation between IQ and learning ability for simple discrimination learning, for paired-associate and serial learning, and in learning-set formation (Zeaman & House, 1967). Our theory would predict that these correlations should be higher in groups containing fewer secondary retardates. A test of this hypothesis that does not require the diagnosis of primary and secondary retardation would be to obtain the correlation between IQ and associative learning ability (or any Level I test) in random samples of school children, one group with IQs from 60 to 95, the other group with IQs from 105 to 140. All the instances of secondary retardation could be presumed to be in the 60 to 95 IQ range. The correlation between associative learning and IQ in this range should be lower than in the range 105 to 140. This test of the hypothesis has not yet been made, although some evidence to be reviewed shortly comes very close to it and is consistent with the hypothesis.

Prehm (1968, pp. 37–38), in reviewing the research on rote verbal learning in the retarded, has drawn 12 conclusions from the evidence:

[1] The rote verbal learning performance of the retarded is considerably more variable than that of *S*s of normal intelligence.

This is what should be expected when the retardate groups are a mixture of primary and secondary types.

[2] The rote learning performance of the retarded is inferior to that of normal *S*s. *This is most true when the materials are more abstract than pictures of common objects.*

We would expect that more abstract items would depend more upon Level II processes.

[3] The serial learning performance of the retarded seems to be subject to the same principles (invariance of the serial position curve, isolation effects, etc.) governing the serial performance of *S*s of normal intelligence.

In a later section we will mention some important exceptions to this generalization which are predictable from our theory.

[4] When compared to massed practice, distributed practice enhances the learning performance of the retarded to a greater extent than it does for normal *S*s.

This conclusion supports the hypothesis that primary retardates have a slower rate of consolidation of short-term memory traces, which, prior to consolidation, are easily interfered with or "erased" by new input; distributed practice allows more time for consolidation and freedom from input and output interference, to the relatively greater advantage of retardates than of normals. It is hypothesized from the present theory that this generalization applies only to primary retardates.

[5] Retardates learn a list of paired associates more readily when the stimulus and response items are the actual objects rather than a picture of that object and when they can pronounce a CVC trigram as a word as opposed to spelling the response.

Paired associate learning tasks can differ in their relative dependence on Level I and Level II processes. Less abstract materials depend less upon Level II processes.

[6] The exposure of stimulus items for longer (four to seven seconds) intervals enhances the learning performance of the retarded.

Again, more consolidation time is of relatively greater advantage to the primary retardate.

[7] The retarded use high level mediational strategies in paired-associate learning to a lesser degree than do *Ss* of normal intelligence.

This conclusion should hold for both primary and secondary retardates, since mediational strategies are examples of Level II processes.

[8] When non-meaningful and meaningful materials are equated for degree of difficulty, retardates exhibit a learning deficit on both types of material.

[9] The retarded exhibit both a short- and a long-term retention deficit.

This, again, theoretically applies only to primary retardates. There is no question of their STM deficit. Long-term deficit is more difficult to prove, since it depends upon equating groups for degree of original learning, which is rarely accomplished. Zeaman (1965) has concluded on the basis of the present evidence, such as it is, that long-term retention is good even in primary retardates.

[10] The retention deficit of the retarded can be minimized by instituting overlearning procedures. The relationship between amount of overlearning and the amount of retention loss is, however, unclear.

[11] Although associative clustering [in free recall of verbal materials] occurs in the retarded, their performance on tasks of this type is inferior to that of the normal *Ss*.

Recent experiments from our laboratory, to be reported in a following section, indicate that free recall per se is a Level I ability and that clustering is a Level II process. Our theory thus mediates certain predictions about the relationships among the variables of age, IQ, free recall, and clustering tendency.

[12] The retention performance of the retarded is impaired as a function of both proactive and retroactive inhibition, with the unlearning of OL [original learning] associations accounting for the effects of retroactive inhibition (RI). Overlearning during OL significantly reduces the effects of RI.

Conflicting Evidence

So far in his search of the literature the writer has found only one experimental result which is unequivocally in conflict with the major hypothesis set forth here. Pursuit-rotor learning would seem to be an even purer form of Level I ability than digit span, serial, and paired-associate learning. So we should expect pursuit-rotor learning to show little if any difference between groups of school children who presumably differ in IQ but not in Level I ability. In fact, in one study of the relationship between pursuit rotor learning ability and MA, the correlation was only .17 (McNemar, 1933). Wright and Hearn (1964) found a large, significant difference in pursuit rotor learning between a group of 20 institutionalized mental defectives and a group of 20 high-school and college students, which is consistent with the idea that institutionalized retardates are usually deficient in Level I. The evidence that appears to be in direct conflict with our theory is from a recent experiment by Noble (1968, pp. 230–232), who found highly significant differences among a sample of 500 rural school children of white (W) and Negro (N) ancestry. The groups were matched for age, sex, and conditions of practice (L vs. R hand). The outcome was WR > WL > NR > NL. When whites, mulattoes (M), and Negroes, similarly matched on age and sex, were compared, the results were W > M > N. As Noble points out, it is hard to know how to interpret these results. Since we have found no difference between Negro and white children on such Level I measures as digit span and serial learning, though they differ by 15 to 20 points in IQ (mostly Level II), it is puzzling why Negro children should perform less well than white children on pursuit rotor learning, which seems to be a purely Level I task. One likely hypothesis is that pursuit rotor learning involves a form of work inhibition ("reactive inhibition" in Hullian terminology) which is absent in STM and verbal learning tasks. There could well be racial differences in rates of build-up and dissipation of reactive inhibition, just as there are highly reliable individual differences within races. Pursuit rotor experiments manipulating distribution of practice, the measure of reminiscence, and other measurements of reactive inhibition such as those described by Jensen (1966), should provide the means for testing this hypothesis.

Goulet (1968) has reviewed the research on serial rote learning in the retarded and concluded that these studies show "unequivocal findings of superior learning for normal *S*s." He goes on to state that these studies, however, "have not provided insight into the specific process or factor responsible for the retardate deficit."

According to our theory, the serial learning deficit should be found only in primary retardates, since serial learning is a Level I ability closely related to memory span. All of the studies of serial learning reviewed by Goulet were based on groups of retardates among whom could be expected a preponderance of primary retardates. The one study which probably had a relatively smaller proportion of primary retardates was one by Cassell (1957). Cassell selected from a population of 152 retardates the 52 subjects who could read; non-readers were excluded. The 52 retarded *S*s who could read showed only a marginal difference from a group of normal children in serial learning ability. Among the retardates, the readers did not differ from the non-readers in IQ. We conjecture that while all were more or less equally deficient in Level II ability, more of the readers were not deficient in Level I ability (i.e., they were secondary retardates) and therefore were of normal ability in serial learning. There can be little doubt that authentic primary familial retardates are markedly deficient in serial learning ability. A study by Jensen (1965a), for example, showed that institutionalized young adult familial retardates were markedly inferior in serial rote learning compared with normal children matched for Stanford-Binet mental age.

Two main types of evidence support the contention that serial learning is essentially a Level I ability. In the first place, normal subjects, when questioned after a serial learning experiment, claim not to resort to the use of strategies, mnemonic devices, mediational techniques, or other "higher level" mental processes in serial rote learning. Their subjective reports of how they learned the serial list are in marked contrast to their reports on paired-associate learning, in which verbal mediational processes play a prominant role in normal adult subjects. Furthermore, neither normals nor retardates show an improvement in serial learning when given special instructions to use verbal mediators in learning the serial list. The same type of instructions, however, greatly facilitate paired-associate learning, relatively more in retardates than in normals (Jensen & Rohwer, 1963a, 1963b). Paired-associates can be learned by means of Level I associative processes, but they also permit the greater play of Level II elaborative processes for subjects who possess these abilities.

Second, Jensen (1965b) has found that individual differences in serial learning are highly correlated with STM for digit series. When a battery of 14 different memory span tests and 17 serial learning measures were factor analyzed together, the loadings of both the memory span and serial learning measures were of approximately the same magnitude on the general factor common to all tests in the battery. Between 67 and 78% of the variance in the various serial tasks and between 67 and 82% of the variance on the memory span tasks was accounted for by the communalities (i.e., the common factor variance).

A series of experiments by Jensen and Roden (1963) showed a relationship between memory span and the degree of skewness of the serial position curve in normal subjects. Subjects with longer memory spans made relatively fewer errors in the first half of the serial position curve than did subjects with shorter memory spans. Since the degree of skewness (i.e., the piling up of errors more toward the end of the serial list during the learning trials prior to mastery) is related to memory span, we should expect from our theory that primary retardates should not only be slower in learning a serial list, but should produce a less skewed serial position curve. Consistent with this prediction, Barnett, Ellis, and Pryer (1960) found a tendency for normal high school students to make relatively more errors for middle items and fewer errors for the beginning items than retarded subjects. The writer tested this hypothesis further by administering an 8-item serial list composed of pictures of familiar objects (i.e., comb, spoon, house, dog, shoe, etc.) to a group of 20 familial mentally retarded (Stanford-Binet IQs between 50 and 70, with a mean of 58) young adults in a state institution for the retarded. No subjects with sensorimotor handicaps or a history or signs of neurological abnormality were included in this sample. Subjects learned by the usual anticipation method. Since the absolute speed of learning was not the essential point of the study, in order to maximize the number who would attain the criterion of mastery (one errorless trial), the serial presentation was subject-paced and subjects were encouraged to guess rather than fail to respond in anticipating each item. Four of the 20 *S*s had to be dropped for failure to attain criterion; their repeated failures and mounting frustration after a reasonable length of time made it inadvisable to continue the task. The serial position curve for the remaining 16 *S*s who attained criterion, plotted as the mean percentage of total errors occurring at each position, is shown in Fig. 7. This serial position curve is extremely atypical from that of normal subjects. It is quite unlike

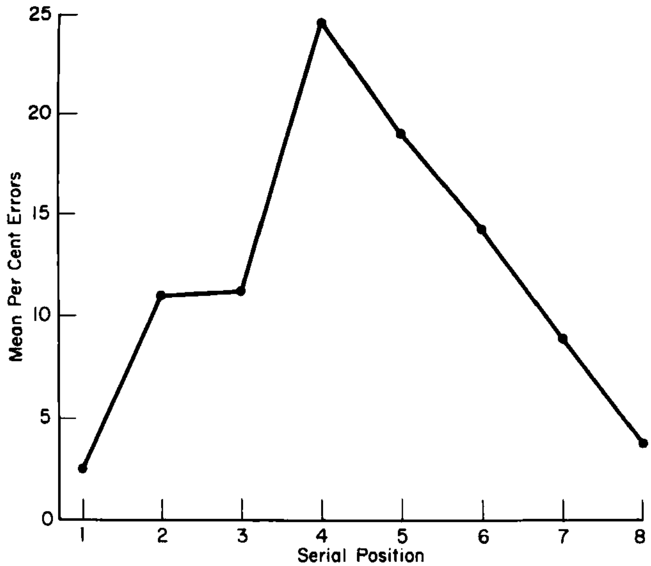

FIG. 7. Serial position curve for 16 primary mentally retarded young adults (IQ 50–70). Note the lack of skewness typically found in the serial position curve of normal subjects.

any the writer has seen in his serial learning experiments with normal subjects or any of the 70 serial position curves he has found in the literature and which closely fit the idealized serial position curve predicted by a theoretical model of serial learning (Jensen, 1962). The serial position curve of the retardates shows none of the skewness of normal serial position curves; the peak of errors comes before the middle of the series rather than just past the middle (i.e., position 4 rather than position 5). It is interesting to note that the best-fitting model of the serial position curve predicts a relative decrease in skewness as the length of the list increases even for normal *S*s (Jensen, 1962). An 8-item list for primary retardates is probably the equivalent of a list of 20 or more items for normal *S*s. For lists of this length the skewness of the serial position curve even for normal subjects would be hardly perceptible.

One serial learning experiment with retardates used the von Restorff effect (also called the isolation effect) to introduce a Level II factor into the serial learning. It is a well-established phenomenon that causing one item in the middle of a serial list to stand out from the others by making it distinctive in some way results in fewer errors on this distinctive item than if it had not been made distinctive. McManis (1966) made an item distinctive by printing it in red, while the remaining items in the serial list were printed in black. Both retarded and normal subjects showed a reduction of errors on the item isolated by this means. When the item in the same serial position was isolated by making it distinctly different in meaningfulness (inserting a low-meaningful item in a list of high-meaningful items), however, only the normal subjects showed the isolation effect – the retardates did not. The registration of the item's meaningfulness is mainly a Level II process, involving the arousal of the subjects' network of verbal associations. Since these spontaneous associative processes are notably deficient in retardates, this form of item distinctiveness in serial learning did not affect their performance.

E. Paired-Associate Learning

Paired associate (PA) learning apparently differs from serial learning mainly in benefiting to a larger degree from past verbal experience. PA learning can be more influenced by verbal mediational processes than serial learning (Jensen & Rohwer, 1963a). Also, the developmental growth curves for serial and PA learning appear to be markedly different. Serial learning ability reaches an asymptote much earlier in life than PA learning. Jensen and Rohwer (1965), in comparing serial and PA learning in children from kindergarten to twelfth grade, found little improvement in serial learning ability beyond 8 or 9 years of age, while PA learning ability showed improvement up to 18 years of age. Beyond 7 or 8 years of age serial learning is more highly correlated with IQ than with mental age, while the reverse is true for PA learning, which suggests that PA learning benefits more from cumulated past verbal experience. Four out of 7 studies of PA learning in which retardates were compared with normals of the *same mental age* showed no significant difference in learning rate; and 4 out of 9 studies in which the retarded and normal groups were of equal chronological age (and therefore differed both in IQ and MA) showed no significant difference in PA learning (Goulet, 1968). Furthermore, all but one of the studies showing retarded subjects to be inferior to normals in PA learning used institutionalized retardates. These findings support the notion that PA learning is largely a Level I function which is facilitated by amount of prior verbal experience largely associated with age, and may also involve Level II processes (mediational strategies, mnemonic elaboration, etc.) when the learning materials are of an abstract nature or are otherwise such as to evoke Level II processes in the learner. The evocation of Level II processes, however, can hinder as well as facilitate PA learning. Wallace and Underwood (1964) found, for example, that retardates do not suffer interference from *conceptual* similarity among items in the PA list, as do subjects of normal intelligence. This type of interference is clearly associated with Level II processes. Other things being equal, however, *abstractness* of the items in PA learning causes greater difficulty in learning for retardates relative to matched MA normals, for example, paired-pictures versus paired-objects (Iscoe & Semler, 1964; Semler & Iscoe, 1965).

F. Rote Learning, IQ, and Socioeconomic Status

A number of studies by the writer and some of his colleagues and graduate students at Berkeley are explicitly relevant to the theory outlined previously.

The first study in this series (Jensen, 1961) compared groups of Mexican-American and Anglo-American fourth and sixth grade school children of different levels of IQ ranging from 60 to 120 on a number of learning tasks consisting of immediate free recall of a dozen familiar objects, serial learning and paired-associate learning of familiar and abstract objects. On these measures of learning ability, Mexican-American children of low IQ (Mean IQ = 82.89, SD = 5.82) were much faster learners than Anglo-Americans of the same IQ (Mean IQ = 81.78, SD = 3.93). Bright Mexican-Americans (Mean IQ = 117.33, SD = 4.27), on the other hand, showed little difference in learning ability. The relationships for all learning tasks are essentially those summarized in Fig. 6. Teachers of the children in this study remarked that the low IQ Mexican-American children seemed much brighter on the playground than the Anglo-American children of similar IQ, although both low IQ groups performed equally poorly in scholastic subjects. Our interpretation is that most of the Mexican-American group in this range of IQs (73 to 89) are somewhat retarded only in Level II functions, while the Anglo-American group in this IQ range is retarded in both Level I and Level II. (The Level II retardation may be either direct or indirect, that is, due to the functional dependence of Level II processes on the more basic Level I processes.)

Rohwer and Lynch (1968) administered a paired-associate test consisting of 24 picture pairs presented 2 times at a rate of 3 seconds per pair to groups of low SES and middle SES children from kindergarten to sixth grade. More than 90% of the low SES children were Negro; all of the middle SES children were white. The low and middle SES groups have an average IQ difference at the various grade levels of between 15 and 20 points. The difference in their scholastic achievement is even more striking. Many children of the low SES group are described by their teachers as "nonlearners" in the classroom, and the majority of these children lag 2 or 3 grade levels behind middle SES children on standard achievement tests. The performances of these groups on PA learning are shown in Fig. 8. Analysis of variance showed no significant differences between the low and middle SES groups. (The difference between grade levels was significant.) The fact that these 2 groups which differ so markedly in IQ and scholastic performance do not differ on this paired-associate learning task leads to the interpretation that the groups differ in Level II but not in Level I abilities. To check this interpretation, Rohwer and Lynch administered the test under the same conditions to a group of retarded young adults in a state institution for the retarded. All were familial retardates without a history or signs of neurological impairment. The fact that they were in an institution is regarded as indicative that most, probably all, are primary retardates. Their average Stanford-Binet MA of 9.70 (IQ of 59) is equivalent to that of normal children

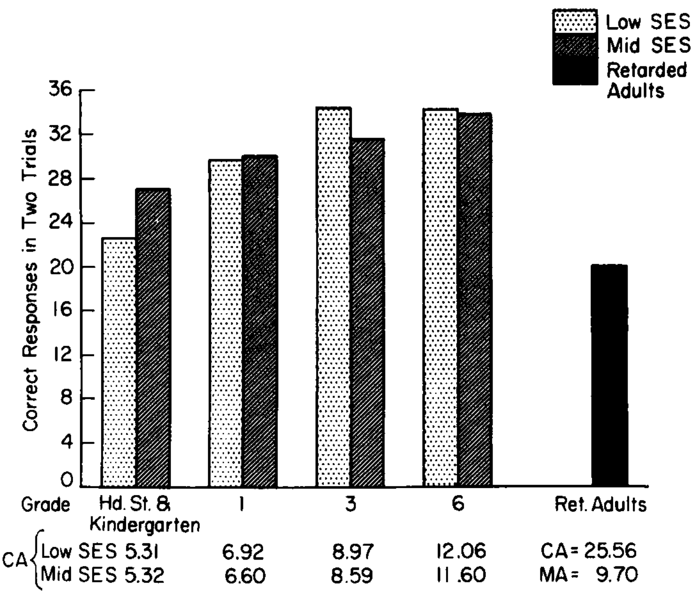

FIG. 8. Comparisons of low and middle socioeconomic groups at various ages with retarded adults on a paired-associate task (24 picture pairs presented two times at a rate of 3 seconds per pair) (from Rohwer, 1967).

in the fifth grade. Yet these retardates showed poorer paired-associate learning ability than the 5-year-old children in Head Start and kindergarten. Also consistent with our hypothesis is the fact that the correlation between PA learning scores and MA (with CA partialed out) is .51 for the middle SES group and .10 for the low SES group. The correlation scatter diagrams of the 2 SES groups show the characteristics depicted in Fig. 4.

In a more recent experiment, Rohwer (1968a) administered four 25-item PA tests (picture-pairs) to groups (total $N = 288$, with 48 in each group) of low SES Negro and upper-middle SES white children in grades K, 1, and 3. These SES groups at all grade levels differed by from about 1.5 to 2 standard deviations (20 to 30 IQ points) on the Peabody Picture Vocabulary IQ and on Raven's *Colored Progressive Matrices*. On the total PA learning score a significant difference between the lower and upper SES groups was found only for the kindergarten children. Rohwer comments

> . . . these results suggest that in the development of the kind of learning ability assessed by the PA test, the discrepancy between upper-strata white children and lower-strata Negro children progressively narrows with succeeding grade levels.

Rohwer goes on to note that this is in marked contrast with the results obtained with the PPVT and the Raven, which show increasing divergence between the SES groups from grades K to 3. This is just what would be predicted from the hypothesized growth curves for Level I and Level II processes (depicted in Fig. 5). This is the only study so far that has failed to show a significant SES difference in the correlations between associative learning ability and psychometric intelligence, although the differences are in the predicted direction. The MA correlated with total PA score .64 in the high SES and .52 in the low SES group; IQ correlated with PA .27 and .22 in high and low SES groups, and the corresponding correlations for Raven raw scores were .44 and .41.

A study by Rapier (1968) helps to establish the phenomenon described in Fig. 6 as a function mainly of social class rather than of race, as might be incorrectly interpreted from the fact that most of our experiments have confounded race and SES. When school children are retested on the basis of SES, there will be a preponderance of Negro and Mexican-American children, 8 to 12 years of age, in public schools. She compared low and middle SES children in special classes for the educable mentally retarded (mean Stanford-Binet IQs for low SES was 70.20, SD = 3.64, range = 63–68, and for middle SES 71.45, SD = 4.95, range = 63–78) and low and middle SES children of above-average intelligence in regular classes (IQ for SES 104.5, SD = 3.23, range 100–110, and for middle SES 105.1, SD = 3.70, range = 100–110). There were 20 *Ss* in each of the 4 groups. All children whose records indicated any sensorimotor, neurological, or emotional disabilities were excluded. (It is an interesting point that Rapier was able to obtain the 20 low SES retarded children from three special classes in one school district but had to canvass 10 special classes in 4 school districts to locate 20 middle SES retarded children.) Serial and PA learning tasks (using pictures of familiar objects) were given to all subjects: 1 serial list and 3 different PA lists administered on 3 different days. (Other experimental variables manipulated in this experiment, involving special instructions to prompt verbal mediation of PA learning, are not central to our present hypothesis.) Rapier's overall results reveal the same relationships as shown in Fig. 6, but, unlike the other studies in our series, the results were in the predicted direction but not significantly so on the first day's serial and PA learning tests. IQ showed a significant effect, but SES and the interaction of IQ × SES were non-significant. On the second day's tests, however, there was a significant IQ × SES interaction, with the low SES retardates and normals showing no appreciable difference in trials to criterion in PA learning (4.6 vs. 4.9) and the middle SES retardates and normals showing a large difference in PA learning trials to criterion 7.7 vs. 4.0). SES, IQ, and SES × IQ were all significant beyond the .01 level on the third day of testing. The normal subjects of the low and middle SES groups did not differ significantly in trials to criterion in PA learning (5.95 vs. 5.10), but the low and middle SES retarded groups differed markedly in learning trials (6.6 vs. 10.1). The learning-to-learn effects of 3 daily sessions on these rote-learning tasks mainly brought about a divergence of the middle and low SES retardates because the middle SES retarded group showed relatively little learning-to-learn (i.e., generalized practice effect).

Also consistent with our hypothesis were Rapier's findings concerning the difference in correlations between IQ and the learning scores for the middle and low SES groups. The average r between intelligence and the learning tests was .44 for the middle SES and .14 for the low SES group; in terms of variance in PA learning accounted for by the variance in the psychometric tests, this represents 19% vs. 2%.

Rohwer, Lynch, Levin, and Suzuki (1968) compared large groups (total N =432) of first, third, and sixth grade children from greatly contrasting high- and low-strata schools. The high-strata school's population was white; the low-strata school's population was Negro. The modal occupational category of fathers of the students in high-strata schools was professional whereas that of fathers of students of low-strata schools was semi-skilled or unskilled manual. The children in the two schools differed widely in psychometric intelligence and achievement. Yet total scores on a variety of PA learning tasks showed no significant difference ($F<1$) between school strata. Rohwer *et al.* state ". . . the average performance of children from low-strata schools was virtually the same as that of children from high-strata schools [p. 19]." This is especially interesting in view of the fact that the relatively low IQs of the low-strata children are commensurate with their generally poor scholastic performance as assessed by standardized tests and the fact that the teachers of these children describe them generally as being "slow to learn and difficult to teach." The PA learning task involves largely Level I ability while the schools' instructional methods apparently rely heavily on Level II abilities—those abilities measured by intelligence tests with a high *g* loading.

In a study by Jensen and Rohwer (1969), 100 low SES Negro preschool children in day care centers and 100 upper-middle SES white children in private nursery schools, all between 3 and 5 years of age, were given digit span tests, a serial learning test, and four paired-associate learning (both using pictures of familiar objects), along with the Peabody Picture Vocabulary Test as the measure of IQ. The correlation between MA and serial learning was .49 for the high SES and .27 for the low SES; the correlation between MA and the total of four PA tests was .58 for high SES and .20 for low SES. The multiple correlation was determined between MA, on the one hand, and CA, serial learning, PA learning, and digit span, on the other. Corrected for shrinkage, the multiple-R was .66 for the high SES and .42 for the low SES group. This corresponds to 44% and 18% of the variance, respectively. In other words, the Level I tests—learning and memory span (plus CA)—predict more than twice as much of the variance in MA for high SES as for low SES children.

G. Free Recall and Associative Clustering

The technique of free recall as a measure of learning and STM especially lends itself to the investigation of the Level I–Level II distinction. In the free recall of uncategorized lists (abbreviated as FR_u), the subject is presented briefly with a number of items (words, pictures, or objects) and then is asked to recall as many of the names as possible within some specified time limit. A number of experimental parameters can be varied

– the number of items, the types of items, the method and rate of presentation. Usually the items are presented in a new randomized order on each trial. Uncategorized lists are composed of items which are relatively unrelated to one another by any supraordinate concept or category labels. The procedures for free recall of categorized lists (FR_c) is the same as FR_u except that the list is composed of items which can be grouped into two or more perceptual or conceptual categories, usually categories that can be readily given a supraordinate category label, like furniture, musical instruments, food, etc. Perceptual categories are those based on resemblance among items on the basis of qualities that range along various dimensions of primary stimulus generalization, such as color, size, and shape. Conceptual categories are mediated by semantic associations, usually of a hierarchical type involving indirect associations among items via their supraordinate category labels.

Comparisons of the amounts of free recall of categorized and uncategorized lists are most valuable from the standpoint of our theory. It has been argued that the reason that low SES children perform so much better on our Level I learning tasks than would be predicted from their IQs and scholastic performance is that our Level I learning tasks (e.g., digit span, serial and PA learning) are less academic, more "interesting," more "relevant," and therefore more motivating to low SES children than are the usual intelligence tests. To rule out this motivational hypothesis as the explanation for our findings, we need two tasks that are essentially indistinguishable in general appearance and procedure, and thus will not elicit different motivational sets, but also which differ clearly in the extent to which performance on the tasks depends upon Level I and Level II abilities. Free recall of uncategorized and categorized lists meets these requirements. FR_u taps mainly Level I ability, or at least requires nothing more than Level I ability, involving simply the reproduction of the input. FR_c also requires nothing more than Level I ability, but it can also reflect Level II ability, i.e., the transformation of the random order of input into conceptual categories as reflected in the order of the subject's output of the items – the phenomenon known as "clustering." Thus, the random imput may be *chair, shoe, bed,* and *hat*; and if there is clustering according to the supraordinate categories of furniture and clothing, the output order will be *chair bed, shoe hat.* The rearrangement of the random input order on the basis of hierarchically arranged verbal mediators is clearly an abstract, conceptual process of the type that characterizes Level II. The amount of material recalled is increased when clustering is possible. Thus, more material is recalled from categorized than from uncategorized lists, and persons who are high on Level II ability should presumably have a relatively greater advantage over persons with low Level II ability in FR_c as compared with FR_u.

1. Associative Clustering in the Mentally Retarded

Studies of free recall and associative clustering in the retarded have been reviewed by Goulet (1968) and Prehm (1968). Three facts are well established both for normal and for retarded subjects: (*a*) perceptual and conceptual clustering both increase with age; (*b*) there is an increase both in the number of items recalled and in the degree of associative clustering over repeated trials; and (*c*) there is a positive correlation between individual differences in the amount of associative clustering and the number of items recalled.

A number of conclusions can be drawn from studies of the retarded. Retardates show less clustering and poorer recall than normals of the same CA. The results for comparisons of retardates and normals of equal MA are more ambiguous, but most studies indicate that MA is a chief source of variance in clustering; retardates and normals matched on MA show similar degrees of clustering (Goulet, 1968). One study, by Rossi (1963), suggests, however, that the level of MA at which retardate versus normal comparisons are made is an important factor, since clustering tendency increases with increasing MA at a faster rate in normals than in retardates. In general, we have claimed that above 5 or 6 years of age, MA, as measured by standard tests such as the Stanford-Binet, is essentially an index of the individual's developmental status in Level II functioning, and these results of equal-MA comparisons reflect just what we should expect according to this formulation.

Compared with normal persons of equal CA, retardates are found to show not only *quantitative* differences in clustering but also *qualitative* differences (Prehm, 1968). Normal subjects cluster items mainly by supraordinate categories; retardates show more pair-wise coordinate groupings, often of an idiosyncratic nature. For example, *bed* and *shoe* may be recalled together consistently on repeated trials. Other items in the list would usually lead to *bed* and *shoe* being separated by normal subjects into the clusters of *furniture* and *wearing apparel*. The retardates' basis for clustering is a coordinate association rather than hierarchical conceptual associations; for example, he will say *bed* and *shoe* go together because "you put your shoes under your bed."

2. Social Class Differences in Associative Clustering

How do groups of children differing markedly in Level II ability (e.g., IQ) but not differing appreciably in Level I (e.g., digit span and serial learning) compare in free recall and associative clustering? This question has been investigated in two studies in our laboratory, using subjects drawn from the same subject pool as that used in our other studies comparing low and middle SES groups in Level I and Level II performance. The prediction from our theory was that low and high SES children would differ little in FR_u but would differ markedly in FR_c, and that the SES difference between FR_u and FR_c would be greater with increasing age of the subjects. These predictions, of course, follow directly from the theory of the relationship between SES and Levels I and II.

Glasman (1968) used several 20-item lists of 4 categories each, with 5 items per category. The categories were: animals, foods, furniture, musical instruments, jobs, eating utensils, clothing, and vehicles. The items consisted of models, toys, or other three-dimensional representations of real objects. The 20 items were presented singly for 3 seconds each, in a random order, for 5 trials. After every trial subjects were allowed 2 minutes to verbally recall the items in any order; the *S*'s output was tape-recorded. There were 32 *S*s in each of the 4 groups formed by the 2 × 2 design; Kindergarten vs. 5th Grade and low SES vs. high SES. The low SES group was composed of Negro children from a school in a low SES neighborhood; the high SES group was drawn from an all white school in an upper-middle-class neighborhood. Thus social class and race are confounded in this experiment. The mean IQs (PPVT) of the groups were 90 for low SES and 120 for high SES. The grade levels were matched on IQ. The main results of the study are shown in Figs. 9 and 10. The measure of clustering (Fig. 10) is the one most commonly used in studies of clustering, and is described by Bousfield and Bousfield (1966). A cluster is defined as a sequence of two responses from the same category which are immediately adjacent. The Bousfield formula corrects this value by subtracting the expected value for a random sequence of the items recalled. The results shown in Tables X and XI clearly bear out our theoretical predictions. At Grade 5 the low SES and high SES groups differ by approximately 1 standard deviation, both in recall and in clustering. (The Grades × SES interaction is statistically significant beyond the .05 level for recall and beyond the .001 level for clustering.)

Since FR_c is essentially a Level II function, it should be correlated with MA about equally in both the low and high SES groups. This was what Glasman found. Correlation between MA and amount of *recall* was .62 for low SES and .72 for high SES; the correlation between MA and amount of *clustering* was .76 for low SES and .77 for high SES. The correlations are much higher for fifth Graders than for Kindergartners, who show very little clustering and are presumably still operating in this

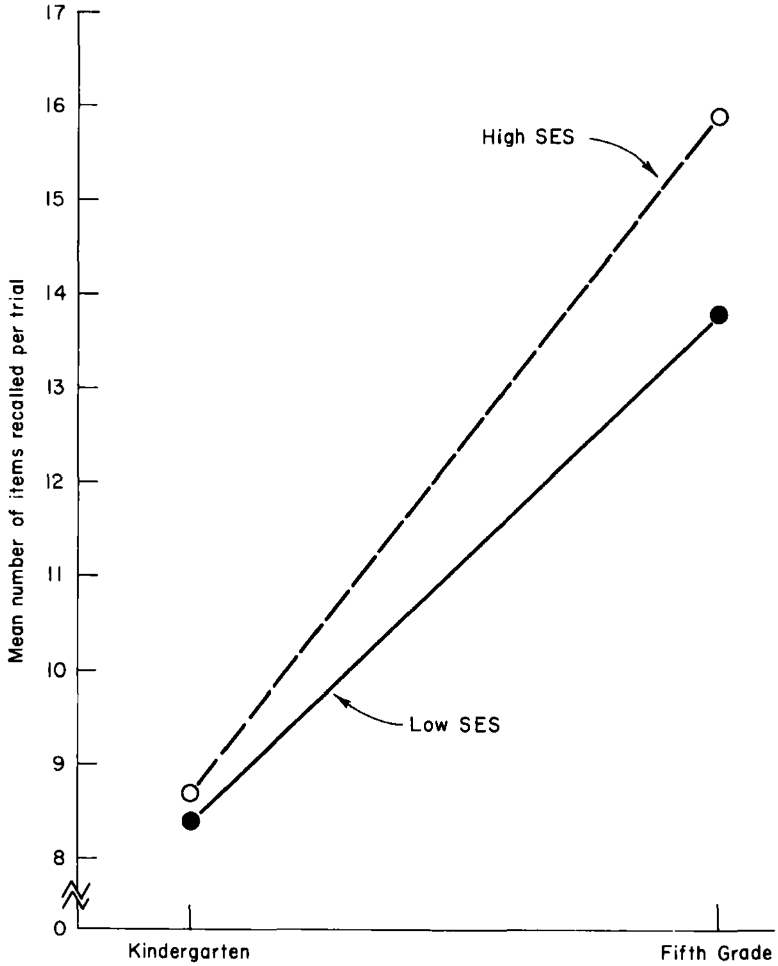

FIG. 9. Mean number of items per trial (over 5 trials) in free recall of a categorized list, as a function of Grade and Socioeconomic Status (SES) (from Glasman, 1968).

task by a Level I process. (The correlation of MA and recall is .06 at Kindergarten and .59 at Grade 5; the correlation between MA and clustering is .02 at Kindergarten and .68 at Grade 5.) These results are highly consistent with predictions based on the hypothetical growth curves for Level I and Level II abilities as a function of SES, shown in Fig. 5. FR_c performance is so strongly related to MA that when the data of Tables X and XI were subjected to an analysis of covariance, with MA as the control variable, all the main effects and the interactions were completely wiped out.

Although Glasman's study demonstrated age and social class differences in the free recall of *categorized* lists, it was not designed to study age and SES differences in performance on the free recall of categorized versus noncategorized lists. A noncategorized list is made up of unrelated or remotely associated items which cannot be readily grouped according to supraordinate categories. Subjective organization of the items in the list is likely to consist of pairs of items related on the basis of pri-

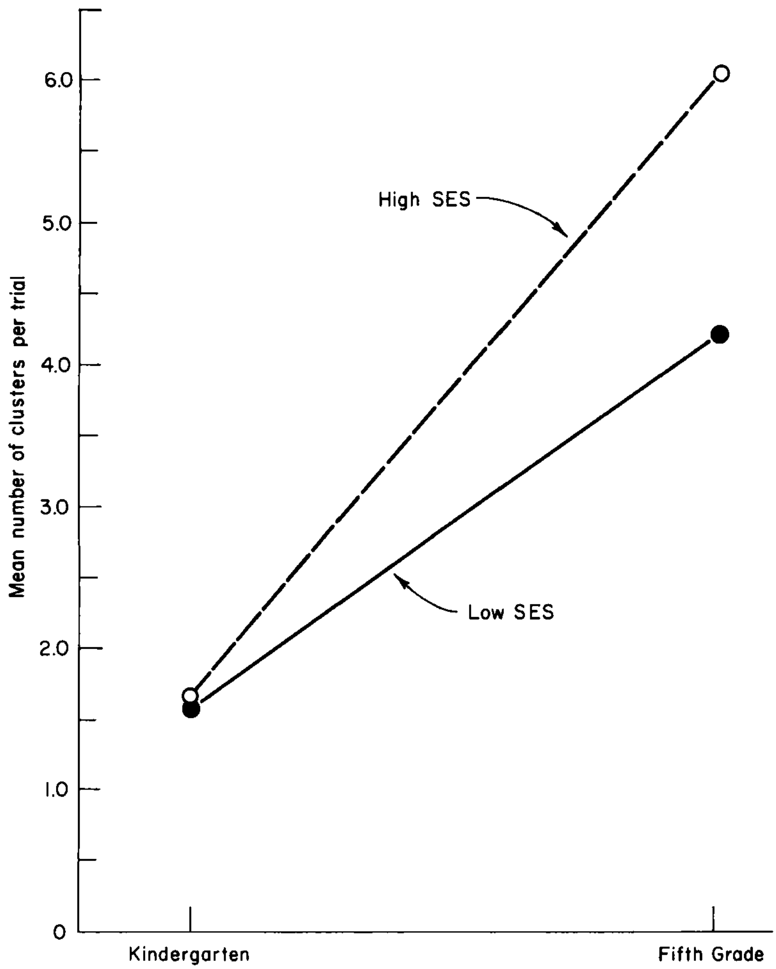

FIG. 10. Mean number of associative clusters per trial (over 5 trials) in the free recall of a categorized list, as a function of Grade and Socioeconomic Status (SES) (from Glasman, 1968).

mary generalization, clang association, or functional relationship. A noncategorized list therefore lends itself less than a categorized list to evoking Level II processes. Consequently, subjects differing in Level II ability (but not in Level I) should show less difference in FR_u than in FR_c.

Jensen and Frederiksen (in press) tested this prediction directly. The low SES and high SES groups were drawn from essentially the same populations as those in the Glasman study, i.e., lower-class Negro and middle- to upper-middle-class white children. The age factor was again investigated by comparing grades 2 and 4. Sets of 20 objects were used for the noncategorized and categorized lists; the 4 categories of the latter were: clothing, tableware, furniture, and animals. Forty *S*s received the noncategorized list, consisting of 20 common but unrelated objects, including 1 object from each of the 4 categories of the categorized lists. Forty *S*s received the categorized list with the items presented in a random order, and another 40 *S*s had the same categorized lists with the items presented in a "blocked" fashion, i.e., all items within a given category are presented in immediate sequence – a procedure which prompts clustering and facilitates recall. Five trials of presentation followed by free recall were given in all conditions. For the categorized lists, the results were essentially the same as those of the Glasman experiment: Grade 4 was superior to Grade 2 under all conditions, and the SES differences were greater at Grade 4 than at Grade 2. Whereas at Kindergarten there was no difference between SES groups, a difference in free recall clearly emerges by Grade 2, in favor of the high SES group. At Grade 4 there is a large interaction between SES level and FR_u vs. FR_c for both random and blocked lists, although the blocked condition reduces the SES difference by boosting the recall performance of the low SES group. In other words, when the input is already categorized and therefore no transformation of the input is called for, the output is facilitated in the low SES group. The high SES group, on the other hand, spontaneously transforms the random input into clustered (i.e., categorized) output and obtains approximately the same facilitation as when the input is already blocked into categories. Recall of the noncategorized list showed a relatively small difference in favor of the high SES group at both second and fourth grades. Also, for the noncategorized list there is no significant interaction between SES and grades – the SES difference is nearly the same at Grades 2 and 4. This is in marked contrast to the categorized lists, which show a large SES × Grades interaction.

All of these findings on free recall are highly consistent with our theory that social class differences in ability involve mainly Level II processes rather than Level I.

V. IMPLICATIONS FOR EDUCATION

If the theory of primary and secondary retardation becomes fully substantiated by further research, it should raise important questions for educational practices. The first question concerns whether different approaches to instruction can yield more optimal effects if they take account of the differences between primary and secondary retardation. It would seem that this distinction should imply quite different techniques and goals of instruction.

Why has traditional schooling been so unsuccessful in teaching children with low IQs but with quite normal Level I learning ability? Many such children do not acquire the basic scholastic skills even in 12 years of schooling. How can one account for this in cases where the child has normal learning ability? One hypothesis is that basic skills are generally taught in such a manner as to make their acquisition heavily dependent upon abstract, conceptual abilities. The criterion of learning in the eyes of many teachers, and the types of pupil performance on which reinforcements from the teacher are contingent, often emphasize the signs of Level II competence—evidence of broad transfer, of broad conceptual generalization of specific learning, of the ability to perform verbal transformations and elaborations on what has been learned, such as being able to "tell it back in your own words" and the ability to say something formally different but conceptually similar. Teachers look for these signs of Level II performance in their pupils. Teachers encourage it, and reward it. The manifestation of Level I ability in its own right is not encouraged or rewarded. It is viewed only as a means to Level II performance. Consequently, the children with the better than average IQs experience a schedule of reinforcements from the teacher and from their perception of their own progress, a schedule of reinforcements which is quite ample for sustaining the behaviors that promote further learning. The low IQ child, on the other hand, even though he may be average or above in Level I learning ability, experiences, in effect, a schedule of non-reinforcement, which results in the experimental extinction of the behaviors that promote learning. One of the major tasks of future research is to determine the full extent to which Level I abilities can be capitalized upon in the teaching of scholastic skills. When Level II performance is made (a) the criterion of learning, (b) the basis for teacher dispensed reinforcement, and (c) the demonstration of having learned by passing achievement tests, the child who is deficient in Level II ability will fail to learn much that could easily be learned by means of Level I.

The writer observed one first grade class of presumably "slow-learning" children called culturally disadvantaged. The majority of these children could not say the alphabet or name the letters of the alphabet. Many apparently could not even discriminate the letters of the alphabet, despite the fact that their teacher had spent part of every school day for 6 months in trying to impart a knowledge of the letters to these children. In their ability to learn school subjects, these children appeared so extremely retarded that the writer suspected primary retardation. The writer's colleague, Dr. William Rohwer, offered to test these children individually on a picture paired-associates learning test which had already been shown to differentiate primary and secondary retardation (see Fig. 9). The children in this class learned, on the average, 16 of the 24 paired-associates in 1 presentation of the list, presented at the rate of 3 seconds per pair. Their performance was completely on a par with that of middle-class children of comparable age in another school who were making normal progress scholastically. Why, then, were the disadvantaged children not learning even letters and simple number facts, to say nothing of reading and writing? Some hours spent in systematic observation of this class and similar classes have led to some psychological speculations that might help to explain these phenomena.

First of all, it was quite apparent that the children's exceedingly poor scholastic performance could not be attributed to any lack of good will, dedication, or effort on the part of the teachers. Furthermore, the teachers had learned well the principle of reinforcement, and readily dispensed encouragement and approval. However, what seemed to be getting reinforced more than anything else was the child's *efforts* rather than his successes. Reinforcing the behaviors that are signs of effort, when the effort does not eventuate in success, indeed increases motivation—but it also leads to frustration. Probably the most potent reinforcement for learning is the child's self-perception of his own *success*, that is, of his own increasing mastery of whatever it is he is attempting to do. Much too few of these instances of success were in evidence in the classes I observed, although the children's effortful but failing attempts at teacher-determined tasks were frequently reinforced by the teacher's well-intentioned praise and approval. Why were there so few opportunities for success? Partly because some of the things being taught were too far beyond the children's present capabilities, but mostly because the teachers seemed to be operating under a preconception of what kinds of behavior constitute learning and should be shaped through reinforcement—it is mainly the child's verbal behavior which evinces Level II processes. Since at the beginning of the term the children were good at Level I associative learning, the teachers tend not to want to "waste their time" on rote activities but instead try to elicit and reinforce almost exclusively those forms of behavior, mostly verbal, which are most characteristic of children with superior IQs. Conceptual brightness, verbally expressed, is the supreme value, even to many devoted teachers who pride themselves on being specialists in teaching the culturally disadvantaged. A child's learning of $2+2=4$ is perceived as being inferior to learning to solve $2+?=4$. The school places excessive valuation and emphasis on what Sheldon White (1965) has called cognitive learning as contrasted with associative learning. Is this possibly the cause of the seemingly poor scholastic potential of many "disadvantaged" children with normal Level I abilities?

Is there a failure to capitalize on existing Level I abilities? To reinforce effort but not success? To make success dependent on Level II abilities when these are meager or undeveloped in some children? These are the conditions that could produce behavioral consequences reminiscent of phenomena described by Pavlov: experimental extinction, conditional inhibition, and experimental neurosis. Accordingly, when the behaviors that are necessary for learning are repeatedly unreinforced, the behaviors extinguish. In addition, the stimulus conditions under which such extinction takes place become conditioned inhibitors. Not only are conditioned inhibitors the stimuli for not responding, but conditioned inhibitors also become aversive stimuli, from which the subject turns away, either passively or actively. Unresponsiveness, drowsiness, inattentiveness, as well as aimless hyperactivity are some of the symptoms of conditioned inhibition. Nearly all the stimuli in the classroom, and especially the teacher and all those things on which the child must focus his attention—books, papers, pencils, and blackboards—all can become conditioned inhibitors for the kinds of behavior essential for learning. Pavlov found in his attempts to establish differential conditioned responses in dogs that when the discriminative stimuli were so similar as to be beyond the dog's capacity to discriminate them, the dog's behavior deteriorated, a condition that Pavlov called "experimental neurosis." It is a condition that can occur without there being any punishment. It occurs simply by withholding reinforcements when the animal fails to make impossibly difficult discriminations. The dog's behavior becomes unstable, hyperactive, and highly resistant to further training. After an experimental neurosis has developed, even the simplest discriminations, which the dog could normally have learned without difficulty, become inordinately difficult or even impossible for the dog to learn. Itard observed manifestations of this condition in Victor, the wild boy of Aveyron, while training him in color and form matching tasks. When the required discriminations were made too difficult, Victor's once normal responding turned to violent anger (Broadhurst, 1961, p. 728). The writer has observed children's behavior in some elementary school classes that closely resembles the manifestations of extinction, conditioned inhibition, and experimental neurosis as described by Pavlov.

Being importuned simply to "try harder" also could be expected to hinder the emergence of whatever Level II processes the child might otherwise evince in learning and problem solving. The well-established Yerkes-Dodson principle states that the optimum level of motivation for performance on complex tasks is lower than for performance on simple tasks. Consequently, if relatively complex learning and problem solving require Level II processes, and if the degree of motivation and arousal is beyond the optimum level for these complex processes, performance will be hindered, and the less complex Level I processes, being nearer their optimal level of motivation, will predominate over Level II. Since the relationship of the Yerkes-Dodson principle to Level I and Level II functions remains speculative, it points to an important area for future research, viz., the relationship of drive states to the potentiation of Level I and Level II functions.

Undoubtedly the most urgent research for its implications for education concerns the question of the extent to which Level II processes can be acquired through appropriate instruction by children of normal Level I ability. The fact that siblings and unrelated children reared in the same family can differ markedly on measures of Level II ability strongly suggests that individual differences in Level II are not solely a product of environmental influences but probably have a substantial genetic component. But this should not rule out the possibility that at least some aspects of Level II functioning can be learned through Level I processes, especially when these are average or above. Some of the cognitive strategies that can facilitate learning and can be acquired by all children of normal Level I ability have been described by Rohwer (1968b), who is conducting an extensive program of research on instructional methods for inculcating, stimulating, or simulating Level II processes in children who do not evince them spontaneously. It is most important that the many children of seemingly meager educational potential in terms of the traditional criteria, but who evince normal Level I abilities, should be given every opportunity to use these abilities in acquiring the basic skills and in achieving realistic educational and vocational goals. Among the important tasks for future research is the further investigation of the theory here proposed and the discovery of means for making the most of Level I abilities in the educational process.

REFERENCES

Barnett, C. D., Ellis, N. R., & Pryer, M. W. Serial position effects in superior and retarded subjects. *Psychological Reports*, 1960, 7, 111-113.

Barrett, B. H. Behavioral individuality in four cultural-familially retarded brothers. Mimeo, undated.

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. (a)

Bayley, N. Research in child development: A longitudinal perspective. *Merrill-Palmer Quarterly*, 1965, **11**, 183–208. (b)

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

Bilodeau, E. A. (Ed.) *Acquisition of skill*. New York: Academic Press, 1966.

Bousfield, A. K., & Bousfield, W. A. Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports*, 1966, **19**, 935–942.

Broadhurst, P. L. Abnormal animal behavior. In H. J. Eysenck (Ed.), *Handbook of abnormal psychology*. New York: Basic Books, 1961. Pp. 726–763.

Burt, C. The evidence for the concept of intelligence. *British Journal of Educational Psychology*, 1955, **25**, 158–177.

Burt, C., The inheritance of mental ability. *American Psychologist*, 1958, **13**, 1–15.

Burt, C. Class differences in general intelligence: III. *British Journal of Statistical Psychology*, 1959, **12**, 15–33.

Burt, C. Intelligence and social mobility. *British Journal of Statistical Psychology*, 1961, **14**, 3–24. (a)

Burt, C. The gifted child. *British Journal of Statistical Psychology*, 1961, **14**, 123–139. (b)

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

Cassell, R. H. Serial verbal learning and retroactive inhibition in aments and children. *Journal of Clinical Psychology*, 1957, **13**, 369–372.

Cattell, R. B. Occupational norms of intelligence and the standardization of an adult intelligence test. *British Journal of Psychology*, 1934, **25**, 1–28.

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

Cooper, G. D., York, M. W., Daston, P. G., & Adams, H. B. The Porteus Test and various measures of intelligence with Southern Negro adolescents. *American Journal of Mental Deficiency*, 1967, **71**, 787–792.

Curti, M., Marshall, F. B., Steggerda, M., & Henderson, E. M. The Gesell schedules applied to one-, two-, and three-year old Negro children of Jamaica, B. W. I. *Journal of Comparative and Physiological Psychology*, 1935, **20**, 125–156.

Durham Education Improvement Program, Duke University, 1966–1967. (a)

Durham Education Improvement Program, Research. Durham, North Carolina, 1966–1967. (b)

Eckland, B. K. Genetics and sociology: A reconsideration. *American Sociological Review*, 1967, **32**, 173–194.

Edgerton, R. B. Anthropology and mental retardation: A plea for the comparative study of incompetence. In H. J. Prehm, L. A. Hamerlynck, & J. E. Crosson (Eds.), *Behavioral research in mental retardation*. Eugene, Oregon: University of Oregon Press, 1968. Pp. 75–87.

Eells, K., Davis, A., Havighurst, R. J., Herrick, V. E., & Tyler, R. *Intelligence and cultural differences*. Chicago: University of Chicago Press, 1951.

Ellis, N. R. The stimulus trace and behavioral inadequacy. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 134–158.

Erlenmeyer-Kimling, L., and Jarvik, L. F. Genetics and intelligence: A review. *Science*, 1963, **142**, 1477-1479.

Fuller, J. L., and Thompson, W. R. *Behavior genetics.* New York: Wiley, 1960.

Gagné, R. M. The acquisition of knowledge. *Psychological Review*, 1962, **69**, 355-365.

Gagné, R. M. Contributions of learning to human development. *Psychological Review*, 1968, **75**, 177-191.

Gates, A. I., & Taylor, G. A. An experimental study of the nature of improvement resulting from practice in mental function. *Journal of Educational Psychology*, 1925, **16**, 583-593.

Geber, M., & Dean, R. F. A. Precocious development in newborn African infants. In Y. Brackbill & G. Thompson (Eds.), *Readings in infancy and childhood.* New York: Free Press, 1966.

Glasman, L. D. A social-class comparison of conceptual processes in children's free recall. Unpublished doctoral dissertation, University of California, 1968.

Gottesman, I. I. Genetic aspects of intelligent behavior. In N. R. Ellis (Ed.), *Handbook of mental deficiency.* New York: McGraw-Hill, 1963. Pp. 253-296.

Gottesman, I. I. Biogenetics of race and class. In M. Deutsch, I. Katz, & A. R. Jensen (Eds.), *Social class, race, and psychological development.* New York: Holt, Rinehart & Winston, 1968. Pp. 11-51.

Goulet, L. R. Verbal learning and memory research with retardates: An attempt to assess developmental trends. In N. R. Ellis (Ed.), *International review of research in mental retardation.* Vol. 3. New York: Academic Press, 1968.

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

Hardy, J. B. Perinatal factors and intelligence. In S. F. Osler & R. E. Cooke (Eds.), *The biosocial basis of mental retardation.* Baltimore: Johns Hopkins Press, 1965. Pp. 35-60.

Havighurst, R. J., & Neugarten, B. L. *Society and education.* (3rd ed.) Boston: Allyn & Bacon, 1967.

Heber, R., Dever, R., & Conry, J. The influence of environmental and genetic variables on intellectual development. In H. J. Prehm, L. A. Hamerlynck, & J. E. Crosson (Eds.), *Behavioral research in mental retardation.* Eugene, Oregon: University of Oregon Press, 1968. Pp. 1-23.

Higgins, C., & Sivers, C. A comparison of Stanford-Binet and Colored Raven Progressive Matrices IQ's for children with low socioeconomic status. *Journal of Consulting Psychology*, 1958, **22**, 465-468.

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

Huntley, R. M. C. Heritability of intelligence. In J. E. Meade & A. S. Parker (Eds.), *Genetic and environmental factors in human ability.* New York: Plenum Press, 1966. Pp. 201-218.

Iscoe, I., & Semler, I. J. Paired-associate learning in normal and mentally retarded children as a function of four experimental conditions. *Journal of Comparative and Physiological Psychology*, 1964, **57**, 387-392.

Jensen, A. R. A statistical note on racial differences in the Progressive Matrices. *Journal of Consulting Psychology*, 1959, **23**, 272.

Jensen, A. R. Learning abilities in Mexican-American and Anglo-American children. *California Journal of Educational Research*, 1961, **12**, 147-159.

Jensen, A. R. An empirical theory of the serial-position effect. *Journal of Psychology*, 1962, **53**, 127-142.

Jensen, A. R. Learning abilities in retarded, average, and gifted children. *Merrill-Palmer Quarterly*, 1963, **9**, 123-140.

Jensen, A. R. Rote learning in retarded adults and normal children. *American Journal of Mental Deficiency*, 1965, **69**, 828–834. (a)

Jensen, A. R. Individual differences in learning: Interference factor. Cooperative Research Project No. 1867, U.S. Office of Education, 1965. (b)

Jensen, A. R. The measurement of reactive inhibition in humans. *Journal of General Psychology*, 1966, **75**, 85–93.

Jensen, A. R. Estimation of the limits of heritability of traits by comparison of monozygotic and dizygotic twins. *Proceedings of the National Academy of Sciences, U.S.*, 1967, **58**, 149–157.

Jensen, A. R. Social class, race, and genetics: Implications for education. *American Educational Research Journal*, 1968, **5**, 1–42. (a)

Jensen, A. R. Patterns of mental ability and socioeconomic status. *Proceedings of the National Academy of Sciences, U.S.*, 1968, **60**, 1330–1337. (b)

Jensen, A. R. Another look at culture-fair testing. In *Western regional conference on testing problems, proceedings for 1968. Measurement for educational planning*. Berkeley, Calif.: Educational Testing Service, Western Office, 1968. Pp. 50–104. (c)

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

Jensen, A. R., & Frederiksen, J. Social-class differences in free recall learning. *Journal of Educational Psychology*, in press.

Jensen, A. R., & Roden, A. Memory span and the skewness of the serial-position curve. *British Journal of Psychology*, 1963, **54**, 337–349.

Jensen, A. R., & Rohwer, W. D., Jr. Verbal mediation in paired-associate and serial learning. *Journal of Verbal Learning and Verbal Behavior*, 1963, **1**, 346–352. (a)

Jensen, A. R., & Rohwer, W. D., Jr. The effect of verbal mediation on the learning and retention of paired-associates by retarded adults. *American Journal of Mental Deficiency*, 1963, **68**, 80–84. (b)

Jensen, A. R., & Rohwer, W. D., Jr. Syntactical mediation of serial and paired-associate learning as a function of age. *Child Development*, 1965, **36**, 601–608.

Jensen, A. R., & Rohwer, W. D., Jr. Mental retardation, mental age, and learning rate. *Journal of Educational Psychology*, 1968, **59**, 402–403.

Jensen, A. R., & Rohwer, W. D., Jr. Experimental analysis of learning abilities in culturally disadvantaged children. Final Report on OEO Project No. 2404, U.S. Office of Economic Opportunity, 1969.

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

Kagan, J. A developmental approach to conceptual growth. In H. J. Klausmeier & C. W. Harris (Eds.), *Analyses of concept learning*. New York: Academic Press, 1966. Pp. 97–115.

Knobloch, H., & Pasamanick, B. The relationship of race and socioeconomic status to the development of motor behavior patterns in infancy. *Psychiatric Research Reports*, 1958, **10**, 123–133.

Kushlick, A. Assessing the size of the problem of subnormality. In J. E. Meade & A. S. Parkes (Eds.), *Genetic and environmental factors in human ability*. New York: Plenum Press, 1966. Pp. 121–147.

Lawrence, E. M. An investigation into the relation between intelligence and inheritance. *British Journal of Psychology, Monograph Supplement*, 1931, **16**, No. 5.

Leahy, A. M. Nature-nurture and intelligence. *Genetic Psychology Monographs*, 1935, **17**, 241–305.

Masland, R. L., Sarason, S. B., and Gladwin, T. *Mental subnormality.* New York: Basic Books, 1958.

Maxwell, J. *Social implications of the 1947 Scottish Mental Survey.* London: University of London Press, 1953.

McGurk, F. C. J. The culture hypothesis and psychological tests. In R. E. Kuttner (Ed.), *Race and modern science.* New York: Social Science Press, 1967. Pp. 367-381.

McManis, D. L. The von Restorff effect in serial learning by normal and retarded subjects. *American Journal of Mental Deficiency,* 1966, **70**, 569-575.

McNemar, Q. Twin resemblances in motor skills, and the effect of practice thereon. *Journal of Genetic Psychology,* 1933, **42**, 70-97.

Neff, W. S. Socioeconomic status and intelligence: A critical survey. *Psychological Bulletin,* 1938, **35**, 727-757.

Noble, C. E. The learning of psychomobar skills. *Annual Review of Psychology,* 1968, **19**, 203-250.

O'Connor, N., & Hermelin, B. Like and cross modality recognition in subnormal children. *Quarterly Journal of Experimental Psychology,* 1961, **13**, 48-52.

Osborne, R. T. Stability of factor structure for the WISC for normal Negro children from pre-school level to first grade. *Psychological Reports,* 1966, **18**, 655-664.

Prehm, H. J. Rote verbal learning and memory in the retarded. In H. J. Prehm, L. A. Hamerlynck, & J. E. Crosson (Eds.), *Behavioral research in mental retardation.* Eugene, Oregon: University of Oregon, 1968. Pp. 31-43.

Rapier, J. L. Measured intelligence and the ability to learn. *Acta Psychologica,* 1962, **20**, 1-17.

Rapier, J. L. The learning abilities of normal and retarded children as a function of social class. *Journal of Educational Psychology,* 1968, **59**, 102-119.

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

Research Profile No. 11. Summary of progress in childhood disorders of the brain and nervous system. U.S. Public Health Service, Washington, D.C., 1965.

Rimland, B. *Infantile autism.* New York: Appleton-Century-Crofts, 1964.

Roberts, J. A. F. The genetics of mental deficiency. *Eugenics Review,* 1952, **44**, 71-83.

Robinson, H. B., & Robinson, N. M. *The mentally retarded child. A psychological approach.* New York: McGraw-Hill, 1965.

Rohwer, W. D., Jr., Social class differences in the role of linguistic structures in paired associate learning. Cooperative Research Project No. 5-0605, U.S. Office of Education, 1967.

Rohwer, W. D., Jr., Socioeconomic status, intelligence and learning proficiency in children. Paper read at the annual meeting of the American Psychological Association, San Francisco, August, 1968. (a)

Rohwer, W. D., Jr. Mental mnemonics in early learning. *Teachers College Record,* 1968, **70**, 213-226. (b)

Rohwer, W. D., Jr., & Lynch, S. Retardation, school strata, and learning proficiency. *American Journal of Mental Deficiency,* 1968, **73**, 91-96.

Rohwer, W. D., Jr., Lynch, S., Levin, J. R., & Suzuki, N. Grade level, school strata, and learning efficiency. *Journal of Educational Psychology,* 1968, **59**, 26-31.

Rossi, E. L. Associative clustering in normal and retarded children. *American Journal of Mental Deficiency,* 1963, **67**, 691-699.

Semler, I. J., & Iscoe, I. Concept interference and paired-associates in retarded children. *Journal of Comparative and Physiological Psychology,* 1965, **60**, 465-466.

Shields, J. *Monozygotic twins brought up apart and brought up together.* London: Oxford University Press, 1962.

Shields, J., & Slater, E. Heredity and psychological abnormality. In H. J. Eysenck (Ed.), *Handbook of abnormal psychology.* New York: Basic Books, 1961. Pp. 298–343.

Shuey, A. M. *The testing of Negro intelligence.* (2nd ed.) New York: Social Science Press, 1966.

Sperrazzo, G., & Wilkins, W. L. Further normative data on the Progressive Matrices. *Journal of Consulting Psychology,* 1958, **22**, 35–37.

Sperrazzo, G., & Wilkins, W. L. Racial differences on Progressive Matrices. *Journal of Consulting Psychology,* 1959, **23**, 273–274.

Talland, G. A. *Deranged memory: A psychonomic study of the amnesic syndrome.* New York: Academic Press, 1965.

Terman, L. M., & Merrill, M. A. *Stanford-Binet Intelligence Scale: Manual for the third revision, Form L-M.* Boston: Houghton Mifflin, 1960.

Terman, L. M., & Oden, M. *The gifted group at mid-life.* Stanford: Stanford University Press, 1959.

Tyler, L. E. *The psychology of human differences.* (3rd ed.) New York: Appleton-Century-Crofts, 1965.

Vernon, P. E. *The structure of human abilities.* New York: Wiley, 1950.

Wallace, W. P., & Underwood, B. J. Implicit responses and the role of intralist similarity in verbal learning by normal and retarded subjects. *Journal of Educational Psychology,* 1964, **55**, 362–370.

Walters, C. E. Comparative development of Negro and white infants. *Journal of Genetic Psychology,* 1967, **110**, 243–251.

Wechsler, D. *Manual of the Wechsler Intelligence Scale for Children.* New York: Psychological Corporation, 1949.

Wechsler, D. *The measurement and appraisal of adult intelligence.* (4th ed.) Baltimore: Williams & Wilkins, 1958.

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

Willerman, L., & Churchill, J. A. Intelligence and birth weight in identical twins. *Child Development,* 1967, **38**, 623–629.

Wright, L., & Hearn, C. B., Jr. Reactive inhibition in normals and defectives as measured from a common performance criterion. *Journal of General Psychology,* 1964, **71**, 57–64.

Young, M., and Gibson, J. B. Social mobility and fertility. In J. E. Meade & A. S. Parkes (Eds.), *Biological aspects of social problems.* Edinburgh: Oliver & Boyd, 1965.

Zeaman, D. Learning processes of the mentally retarded. In S. F. Osler & R. E. Cooke (Eds.), *The biosocial basis of mental retardation.* Baltimore: Johns Hopkins Press, 1965. Pp. 107–127.

Zeaman, D., & House, B. J. The relation of IQ and learning. In R. M. Gagné (Ed.), *Learning and individual differences.* Columbus, Ohio: Merrill, 1967.