INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.
- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.
- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17"x 23" black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6"x 9" black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.

8713806

Van Wess, Laurie E.

A TWIN STUDY OF GENERAL INTELLIGENCE AND SPECIFIC COGNITIVE ABILITIES

City University of New York PH.D. 1987

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

A Twin Study of General Intelligence and Specific Cognitive Abilities

by

Laurie E. Van Wess

A dissertation submitted to the Graduate
Faculty in Psychology in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy, The City University of
New York.

1987

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation for the degree of Doctor of Philosophy.

4/28/87
Date

David R. Owen
Chair of Examining Committee

April 29, 1987
Date

Herbert D. Saltzstein
Executive Officer

Professor David R. Owen

Professor Arthur Reber

Professor Don Scarborough
Supervisory Committee

The City University of New York

Abstract

A TWIN STUDY OF GENERAL INTELLIGENCE AND SPECIFIC COGNITIVE ABILITIES

by

Laurie E. Van Wess

Advisor: Professor David R. Owen

The performances of monozygotic and same-sex dizygotic twins were examined on an abbreviated form of the Wechsler Adult Intelligence Scale--Revised (WAIS-R) as well as on sequential pattern learning and frequency of occurrence tasks. The heritablity estimate ($\underline{H}$) for the WAIS-R of .61 showed excellent agreement with the previous literature on general intelligence. The heritability estimates for the sequential pattern learning task ($\underline{H} = .79$) and for the frequency of occurrence task ($\underline{H} = .29$) provided evidence of differential heritability for specific cognitive skills. In addition to studying genetic influence, this research examined the effects of two independent variables on sequential pattern learning performance. The type of operation required to complete the pattern, the number of items which change within a string, and their interaction were found significantly to influence performance. Finally, the study revealed a moderate correlation between general intelligence and sequential pattern learning performance and, as hypothesized, virtually no relationship between general intelligence and performance on the frequency of occurrence task.

Acknowledgements

I dedicate this effort to my family and to my friends. I should like to express my appreciation for their unending support and cooperation. Special appreciation is due Dr. David R. Owen for his unfailing encouragement and advice during the planning and execution of this study. I should also like to thank, perhaps mostly of all, the 50 pairs of twins who unselfishly gave of their time to serve as subjects.

TABLE OF CONTENTS

SECTION PAGE

SECTION PAGE

LIST OF TABLES

LIST OF FIGURES

A Twin Study of General Intelligence and Specific Cognitive Abilities

Quantitative Genetic Theory

Genotype is the genetic composition of the individual and phenotype is the apparent, visible, measurable characteristic. Quantitative genetic theory was developed to study phenotypic differences in a population and the genetic and environmental factors that create them. Phenotype (P) will not be measured as an individual's phenotypic value, but will be rescaled as an individual's absolute deviation from the population mean. Quantitative genetic theory starts with a model in which some observed phenotypic deviation from the mean for a particular characteristic in a population is a function of genetic (G) and environmental (E) deviations, which combine in a linear way. Symbolically,

$$P = G + E \, .$$

This model may also include an interaction term (GxE) to deal with possible nonadditive combinations of genetic and environmental effects. In other words, an environmental factor may have a greater effect on some genotypes than on others, and a genotype may be expressed differently in some environments than others. Thus,

$$\underline{P} = \underline{G} + \underline{E} + (\underline{G}\text{x}\underline{E}) \ .$$

The genotypic value ($\underline{G}$) may be subdivided into its component parts. One concept basic to this model is the additive genetic value. Each gene in the genotype that influences a particular character has some average effect on that character. The additive genetic value, symbolized $\underline{A}$, is the sum of these average effects, where the summation is across both genes at each locus. In addition to adding up in a linear fashion, it is also possible for alleles at a given locus to interact with one another. Thus, dominance deviation, represented as $\underline{D}$, is the difference between the expected additive genotype and the actual genotypic value. In the single-gene model, the genotypic value, $\underline{G}$, can be partitioned into the additive component, $\underline{A}$, and the nonadditive component, $\underline{D}$, so

$$\underline{G} = \underline{A} + \underline{D} \ .$$

Instead of considering the additive and

nonadditive effects of alleles at a single locus, the polygenic model considers the sum of these effects across loci. Just as additive genetic values are the summation of the average effects of the alleles at a single locus, they may also be summed across the many loci that may influence a particular genotypic character. In the same way, dominance deviations from additive genetic values may also be summed for all the loci influencing a character. At this point, however, the polygenic model requires the introduction of an additional concept. Epistasis should not be confused with dominance, which is the nonadditive interaction of alleles at a single locus. As several loci are considered, it is also necessary to consider the possibility that a particular allele may also interact with alleles at different loci. This type of interaction is called epistasis, $\underline{I}$. In the polygenic model, genetic effects may be partitioned into three components: additive, dominance, and epistatic. In symbolic terms,

$$\underline{G} = \underline{A} + \underline{D} + \underline{I} .$$

In this way, the basic model of quantitative genetic theory simply states that the phenotype of an individual is due to a genotypic value, including $\underline{A}$, $\underline{D}$, and $\underline{I}$, plus an environmental effect (all nongenetic causes).

Each of the values and deviations previously defined has a corresponding variance symbolized by $\underline{V}$ and an appropriate subscript. These variances and their symbols are indicated in Table 1. The fundamental problem of quantitative genetics is to partition the observed phenotypic variance, $\underline{V}_p$, into these theoretical components in order to estimate their proportional contributions to population variability. In the absence of genotype-environment interactions, $\underline{P} = \underline{G} + \underline{E}$; thus,

$$\underline{V}_p = \underline{V}_g + \underline{V}_e + 2\mathrm{Cov}_{(g,e)} .$$

If $\underline{G}$ and $\underline{E}$ are uncorrelated, that is, if environmental effects are distributed at random across genotypes,

$$\underline{V}_p = \underline{V}_g + \underline{V}_e .$$

If genotype-environment interactions are present,

$$\underline{V}_p = \underline{V}_g + \underline{V}_e + \underline{V}_{(gxe)} .$$

Table 1

A Summary of Variances and Their Symbols

Variance	Symbol
Phenotypic	$\underline{V}_p$
Genotypic	$\underline{V}_g$
Additive genetic	$\underline{V}_a$
Dominance	$\underline{V}_d$
Epistatic	$\underline{V}_i$
Environmental	$\underline{V}_e$
Genotype-environmental interaction	$\underline{V}_{(gxe)}$

Since $\underline{G} = \underline{A} + \underline{D} + \underline{I}$,

$$\underline{V}_g = \underline{V}_a + \underline{V}_d + \underline{V}_i \ .$$

No covariance terms are necessary here because $\underline{A}$, $\underline{D}$, and $\underline{I}$ have been derived in such a way as to be independent.

If it were possible to measure genetic and environmental effects for individuals, it would be possible to directly estimate $\underline{V}_g$ and $\underline{V}_e$ in populations. Instead, it is necessary to indirectly estimate these values. All methods that estimate these genetic and environmental components of variance utilize relationships that differ in their degree of genetic or environmental similarity.

The genetic contribution to the phenotype of an individual can be divided into two parts. One part refers to that portion of genetic variance which the individual shares or has in common with the relative. This component is typically symbolized as ($\underline{G}_c$) and refers to the differences between families. The second part refers to the genetic variance that is not shared with the relative. This component is usually designated as ($\underline{G}_w$) to indicate differences within families. Likewise, some environmental influences are shared by relatives while other aspects of the environment are not. Thus, the environmental contribution to the phenotype can also be divided into influences shared with the relative ($\underline{E}_c$) and those independent of the relative ($\underline{E}_w$). In this way, the phenotypic covariance between relatives can be expressed in terms of components of variance:

$$\mathrm{Cov}_{(\underline{P}_1,\underline{P}_2)} = \underline{V}_{gc} + \underline{V}_{ec} \; .$$

In other words, the covariance between relatives for a particular character equals the genetic variance and the environmental variance resulting from shared genetic and environmental influences.

Recall that the genetic component, $\underline{G}$, can be partitioned into three components: the additive component ($\underline{A}$), dominance ($\underline{D}$), and epistatsis ($\underline{I}$). Similarly, $\underline{V}_g$ may be partitioned into: $\underline{V}_a$, $\underline{V}_d$, and $\underline{V}_i$. These components of genetic variance contribute differently to each type of family relationship. For instance, parents and their offspring share one-half of their additive genetic variance since each parent contributes only one-half of the genotype of each offspring. Half siblings are individuals that have one and only one parent in common. The additive genetic variance for half sibs is equal to one-half the additive genetic value of the common parent, or 1/4 $\underline{V}_a$. Full siblings have both parents in common. Thus, the additive genetic variance for full siblings should be twice that of half siblings or 1/2 $\underline{V}_a$. Full siblings also share one-fourth of the dominance variance, since full siblings can be expected to receive the same alleles from both parents one-fourth of the time, and thus have the same dominance deviation. Dizygotic (DZ) twins are like other siblings and share half of the additive genetic variance and one-fourth of the variance due to dominance. In contrast, monozygotic (MZ) twins share all genetic variance, $\underline{V}_a$, $\underline{V}_d$, and $\underline{V}_i$. Thus, it may be observed that the value of the shared genetic variance is dependent upon the particular type of family relationship.

One frequently misunderstood, but nonetheless important, concept of quantitative genetics is the notion of heritability. Simply stated, broad-sense heritability, $\underline{h}_b{}^2$, is a statistic that describes the ratio of genetic to phenotypic variance in a population. Thus,

$$\underline{h}_b{}^2 = \underline{V}_g \,/\, \underline{V}_p \ .$$

It is important to emphasize that heritability is a population parameter.

Another concept is narrow-sense heritability, h^2 , which is the proportion of phenotypic variance solely due to additive genetic variance. Thus,

$$h^2 = V_a / V_p .$$

In contrast, broad-sense heritability (h_b^2) is the proportion of phenotypic differences due to all genetic variance, both additive and nonadditive.

Several methods have been used to estimate the heritabilty of a trait in human beings. To begin, studies of parents and their offspring commonly utilize correlational and regression analyses. These statistics are based on covariance. Covariance literally means shared variance and it indicates the extent to which the measures relate to one another. The covariance between two variables , $\mathrm{Cov}_{(x,y)}$, is:

$$\mathrm{Cov}_{(x,y)} = E[(x-\bar{x})(y-\bar{y})] .$$

Covariances are easier to interpret if they are standardized by an appropriate variance or a product of standard deviations. A correlation coefficient standardizes the covariance by dividing it by the product of the standard deviations of x and y . Thus,

$$r_{xy} = \mathrm{Cov}_{(x,y)} \;/\; \sqrt{(V_x)(V_y)} \;.$$

The regression coefficient divides the covariance by the variance of just one of the variables. For example, if one was predicting offspring scores (y) from parental scores (x) , the regression of y on x (b_{yx}) is the covariance divided by the variance of x . Thus,

$$b_{yx} = \mathrm{Cov}_{(x,y)} \;/\; V_x \;.$$

It has been demonstrated that squaring the correlation yields the percent of variance in one variable related to the variance of the other. Thus, one formula for heritability is

$$h^2 = r^2_{ap} \;,$$

with a representing the additive genetic component and p representing the total phenotypic value. Similarly when regression is used, heritabilty is equal to the regression of additive genetic variance on phenotypic variance,

$$h^2 = b_{ap} .$$

The regression of offspring on one parent is a measure of 1/2 h^2 , and that of offspring on mid-parent is a measure of h^2 .

At this point, the issue of assortative mating may be introduced. Sometimes the mating of parents is not random but according to their phenotypic resemblance. Thus, the assumption of random mating implicit in this model is often violated. Assortative mating increases the correlation between parents and the genetic variance in a population. However, according to Falconer (1960, p. 171), the regression of offspring on midparent is minimally affected and is still a valid measure of h^2 . In addition, several researchers have found little assortative mating for specific cognitive abilities after adjusting for educational level and/or similarity in age (Johnson et al., 1976; Zonderman, Vandenberg, Spuhler, & Fain, 1977). Therefore, the issue of assortative mating is generally not problematic provided that the regression of offspring on mid-parent is measured.

In addition to measuring heritability from parent-offspring studies, a variety of mathematical expressions has been used to analyze data from twin studies. Generally, analysis of variance is used. One well-known measure is Holzinger's (1929) heritability coefficient, H .

$$H = (r_{mz} - r_{dz}) / (1 - r_{dz})$$

$$H = (\delta^2_{dz} - \delta^2_{mz}) / \delta^2_{dz}$$

In order for these two formulas to be equivalent, the intraclass correlations, r's, must be computed by the following relationships:

$$r_{mz} = 1 - (\delta^2_{mz} / \delta^2)$$

$$r_{dz} = 1 - (\delta^2_{dz} / \delta^2)$$

where δ^2 is the total sample variance. This coefficient is not the same as that which was symbolized as h^2 previously. Holzinger's ratio gives the proportion of differences produced only by genetic differences within families.

Several criteria must be met before using

Holzinger's estimator. First, genetic and environmental variances must be additive. Second, the MZ and DZ pairs must have equivalent means and between-pair variances on the particular character being studied. Third, environmental similarity for MZ and DZ groups is assumed (Fuller and Thompson, 1978). Falconer (1960) indicated that this measure "could be taken as an estimate of half the heritability if there were no non-additive genetic variance.... But since non-additive variance cannot reasonably be assumed to be absent, the difference can only be regarded as setting an upper limit to half the heritability." Nichols (1965) developed a formula that corrected for this underestimation:

$$\underline{HR} = 2(\underline{r}_{mz} - \underline{r}_{dz}) / \underline{r}_{mz} .$$

It is usually important to evaluate the significance of the difference between MZ and DZ variances. This is commonly accomplished by the use of the $\underline{F}$:

$$\underline{F} = \sigma^2_{dz} / \sigma^2_{mz} .$$

A significant $\underline{F}$ indicates that heredity and environment produce greater differences in DZ twins than environment alone produces in MZ twins. A simple algebraic relationship exists between *H* and *F*, thus the *F* enables one to estimate heritability as well:

$$H = 1 - (1/F) .$$

Once the basic model of quantitative genetics has been introduced, the next logical step is to review the literature. After reviewing various family, adoption, and twin studies with respect to general intelligence, numerous specific cognitive abilities will be summarized as well.

Behavior Genetics of General Intelligence

Family Studies. One of the basic methods of studying the relative influences of genetic and environmental factors on intelligence is to conduct family studies. Family studies typically involve obtaining various measures of intelligence from both parents and their offspring. According to the genetic hypothesis, genetically related individuals should be similar phenotypically to the extent that genes influence intellectual behavior.

Two major reviews have been conducted in the past 25 years. First, Erlenmeyer-Kimling and Jarvik (1963) conducted a comprehensive review of family studies of intelligence. Based on 12 studies, they reported a median correlation of .50 for parent and child relationships. The median correlation for siblings reared together was .49, .53 for dizygotic twins, and .87 for monozygotic twins.

Bouchard and McGue (1981) conducted a more recent review of familial studies of intelligence. Several of their reported correlations are presented in Table 2.

Since the publication of these literature reviews, many additional family studies have been reported. For instance, Reed and Rich (1982) tested 2,029 pairs of parents with one or more offspring. The IQ values for both parents were obtained when they were teenagers and their offspring were also tested during their teens. The midparent-offspring correlation was .531. The regression coefficient for each offspring's IQ on the midparental IQ was .613 ± .022.

In addition, Vogler and DeFries (1983) recently reported on 843 families from the Americans of European ancestry sample in the Hawaii Family Study in Cognition (HFSC). A composite score derived from a principal component analysis of data from 15 tests of specific cognitive abilities was used as a measure of general cognitive ability. The regression of midoffspring on midparent was .584 ± .034.

In summary, family studies typically report correlations and regression coefficients ranging from .46 to .64. These data may be interpreted as being consistent with the genetic hypothesis but not as conclusive proof of genetic influence.

Adoption Studies. One problem with studying familial resemblance is that genetic and environmental factors covary. In other words, it is generally found that bright offspring are raised in favorable environments and dull offspring are reared in less favorable environments. Due to this relationship, studies that find resemblances among family members as a function of their increasing biological relations can be interpreted to support either a nature or a nurture theory. In order to attribute influence independently to nature or nurture it is necessary to unconfound genotype and environment. Adoption studies provide an opportunity to estimate the relative influence of nature and nurture on intelligence. In essence, the resemblances of adopted children's IQ to their biological and adopting parents' IQs are measured. The similarity between adopted children's IQs and their adopted parents' intelligence is a direct estimate of the environmental influence, while

Table 2

Familial Correlations for IQ

Relationship	Number of Pairs	Median Correlation	Weighted Average
MZ twins			
reared together	4,672	.850	.86
DZ twins			
reared together	5,546	.580	.60
Midparent-offspring			
reared together	992	.475	.50
Siblings			
reared together	26,473	.450	.47

the similarity between adopted children's IQs and their biological parents' intelligence is a direct estimate of the genetic component.

Several methodological criteria must be satisfied before a set of adoption data can be used to estimate the relative influence of nature and nurture on intelligence (Munsinger, 1975). A list of the most important problems that one must surmount in adoptive studies includes: (a) possible initial sampling biases; (b) differential attrition of the adoptive families over time; (c) possible selective placement of children in adoptive homes; (d) unreliable measurement of intelligence; (e) lack of reliable information about parents; (f) confounding of nature and nurture by late separation from the biological family, late adoption into the foster home, or both; (g) uncontrolled variation in the age of parents, and sex or race of children; (h) differential practice effects from repeated measurements; and (i) numerous statistical problems. Each methodological problem will be described and the effects of violating each criterion will be discussed.

The particular problem of initial sampling was considered by Leahy (1932). Leahy collected information on 9,973 illegitimate births in Minnesota between 1918 and 1928. More specifically, she recorded (a) the number of these illegitimate children who were actually put up for adoption, (b) the age of each child when adopted, and (c) the parental characteristics of adopted illegitimate children compared with illegitimate children who were neither put up for adoption nor adopted into a new home. Her careful analysis revealed that mothers who put their illegitimate children up for adoption earned significantly higher occupational status scores and had completed almost a year more education than mothers who retained their illegitimate children. Since the average age of these two groups of mothers was almost identical, the occupational and educational differences were interpreted as differences in ability and achievement. In addition, it was found that illegitimate children who are released for adoption have brighter parents than do illegitimate children who are retained by their mothers.

Another result of this study indicated that those children who are adopted early are brighter than children adopted later. Since most published studies contain only samples of illegitimate children who were adopted by six months of age and show no signs of physical or mental health problems, one should expect the average group IQ of adopted children to be significantly higher than the average of the population. It is almost impossible to find an adequate comparison group to test for increases in group IQ. This selection bias might also reduce the variation of scores and attenuate correlations. However, all correlations should be attenuated about the same amount so that the relative influences biological and adoptive parents exert on their children may still be compared. In addition, it has been noted that most groups of adopted children have IQ standard deviations that are very similar to the total population. Thus, if the effect exists, it is very small.

A second possible source of bias in adoption studies is differential attrition of the adoptive sample over time when longitudinal measures of intelligence are involved. More specifically, it has been reported that in longitudinal studies the higher social-status families are more likely to remain in a sample than lower social-status families. If all available scores are included in the analyses, this attrition bias will make the adopted children appear to grow brighter and brighter over time. On the other hand, if only the adopted children remaining at the end of the study are included in the analysis, then the result will be to make the group seem consistently superior in average IQ. Also, differential attrition may decrease the variation of scores and thus tend to lower correlations over time if all the data are included, or it may simply lower all correlations if only the subjects remaining at the end are used in computing the correlations. Consequently, any within group comparisons for change over time are not affected by differential attrition if all incomplete data are excluded from the analyses, but any between group comparisons will still be affected if the other group has not undergone identical differential attrition over time.

A third possible source of bias is the selective placement of adopted children into adopting homes that are similar to their biological parents' social and educational backgrounds. This has been a standard practice of adoption agencies. Selective placement can confound genetic endowment with environmental influence. This invalidates the basic logic of an adoptive study. A number of investigators have estimated the extent of selective placement. Generally, a correlation between biological mothers' education and adoptive midparent education is obtained. Correlations between adoptive and biological parents in the United States have averaged from .20 (Horn, Loehlin, and Willerman, 1979) to .27 (Skodak and Skeels, 1949). Selective placement could inflate a correlation between the IQs of adopting parents and their adopted children. Accordingly, it would appear as if the adoptive environment exerts a stronger influence on children's cognitive development than it actually does. Two methods can be used to correct this problem: (a) use data only from studies that show no selective placement, or (b) make a statistical correction.

A fourth source of possible bias is invalid measurement of IQ. This can occur because of poor testing procedures, the use of different tests without a reliable means of comparing these different scores, age changes in the content of the tests themselves, or the use of poor measures of intelligence for the parents. Random errors of measurement will decrease the size and therefore the statistical significance of all correlation coefficients. The result will be that unreliable measurements attenuate the measured effect of both heredity and environment. However, these unreliable measurements should not disturb their relative values.

A fifth source of possible bias is the lack of information about biological parents. There is especially a lack of information concerning the intelligence, education, age, and occupation of biological fathers of illegitimate children put up for adoption. It is difficult to measure the direction and strength of this bias, but some trends are clear. Generally, biological fathers on whom information is available tend to be approximately five years older than biological mothers and also had an average of one to two years more education than the biological mothers of illegitimate children put up for adoption. Thus, the loss of paternal data would tend to bias downward the expected IQ of the adopted children because the predictions are based on young mothers. These offspring are expected to be less intelligent than they become.

A sixth problem concerns the ages at which adopted children are separated from their biological parents, the type of institution that they may be placed in prior to adoption, and their ages when placed in a new adoptive family. Differential age of separation from the children's biological parents can exert different kinds of biases. Firstly, it has previously been noted that infants who are released for adoption early (before two months of age) score significantly higher on IQ tests compared with children who are released much later. In addition, if the children spend a significant amount of time with their biological parents before being released, the whole logic for doing an adoption study is invalidated. Under these circumstances, it is impossible to determine whether any correlations that are found between the biological parents and the children are due to nature, nurture, or both. The same problem applies if children are placed in an institution for a significant period of time prior to adoption. The apparent solution to this situation is to include in the final analysis of the adoptive data only those children who were released for adoption within a few days after birth and ultimately placed in the new adoptive homes at an early age.

A seventh possible source of bias in adoption studies is the common finding that the mean age of biological parents (20 years) is approximately 10 years less than the mean age of adoptive parents. Consequently, estimates of the biological parents' cognitive abilities based on indicators such as number of years of education or occupational status may be low, provided their education has not been completed, compared with the same estimates for the older adoptive parents. This may bias any expectations about the average intelligence of the adopted children. A researcher may inappropriately attribute any difference between the expected average IQ scores of the adopted children and their actual IQ scores to the beneficial effect of a superior adoptive environment.

Probably the best way to correct for the age-of-parents bias would simply be to accept the existence of a possible biological parents' age bias. More specifically, researchers should pay particular attention to samples of biological parents who are over 20 years of age. They should then observe if the pattern of results is different between the studies with younger and older biological parents. If there are no serious differences, then the ages of biological parents may be ignored. However, if a difference appears, then either statistical corrections should be utilized or only the data from older parents should be used in the final analysis.

An eighth source of possible bias concerns differential practice effects. Adopted children may be repeatedly tested and then these results are compared with the average IQ of a general population, which have only been tested once. A practice effect bias is most relevant to group differences and should have little if any effect on correlations.

Finally, several statistical problems may create bias in adoption studies. These include: (a) restriction in the range of scores within a group, (b) possible regression to the mean with repeated measurements, and (c) the fact that multiple regression predictions may include chance associations between particular variables and the child's IQ score. Careful consideration must be given to each of these issues.

Thus, there are many methodological issues that should be dealt with as one attempts to gather reliable data on the IQ development of adopted children. Each of these potential problems should be considered as several published adoption studies are reviewed. The first adoption study to be reviewed was a comprehensive longitudinal study from 1936 through 1949 conducted by Skeels and Skodak. The criteria for inclusion in this study were as follows: (a) The child was placed in an adoptive home before six months of age; (b) the child had been given an intelligence test prior to November 1936 but after one year residence in the adoptive home; (c) some information existed concerning the natural and the adoptive parents; and (d) the child was white and of North European background. There were 180 children of the initial sample who met these criteria. In 1937, 154 of these children were retested and in 1940, 139 were again retested. The final follow-up sample in 1946 consisted of 100 children who were tested for a fourth time.

The mental development of the adopted children was measured for the first time at an average age of 2 years 2 months. The Kuhlmann-Binet Scale was administered if they were under 3.5 years of age while the Stanford-Binet Scale was used if they were over 3.5 years of age. All test scores reported for the second and third examination were based on the 1916 Stanford-Binet Intelligence Scale. The fourth and final examination consisted of the 1916 revision and Form L of the 1937 revision. The mean age at the second examination was 4 years 3 months, 7 years and 0 months at the third examination, and 13 years 6 months at the fourth examination. Approximately one-third of the biological mothers and one-fourth of the biological fathers were administered the 1916 Stanford-Binet IQ test. In addition, educational, economic and vocational information was gathered for both the biological and adoptive parents.

The correlations between the adopted children's IQs and the adoptive mothers' education for the four tests were as follows: -.03 for Test 1, .04 for Test 2, .10 for Test 3, and .04 for Test 4. All were insignificant. On the other hand, the correlations between the adopted children's IQs and the biological mothers' IQs were as follows: .00 for Test 1, .28 for Test 2, .35 for Test 3, and .38 or .44 for Test 4 (for the 1916 and 1937 versions of the Stanford-Binet Scale, respectively). The correlations between the adopted children's IQs and their biological mothers' education were .04 for Test 1, .31 for Test 2, .37 for Test 3, and .31 for Test 4. In sum, the adopted children's IQ scores were more strongly correlated with their biological mothers' education (and IQ scores) than with their adopted mothers' education.

It is important that the adequacy of these data be evaluated with respect to the methodological criteria previously discussed. Firstly, there is no doubt that biased initial sampling occurred in selecting these children. All of the adoptive children were: illegitimate, adopted before six months of age, white and of Northern European descent, and without obvious physical health problems. In addition, the adoptive parents had a period of time in which to accept or reject an infant, and the only children included in the initial sample were those for whom the adopting parents had applied for final legal adoption.

The two adoption agencies participating in this research had a total of 319 children under the age of six months between 1933 and 1937. Selection bias of the initial sample was strong since only 180 of these children met the initial selection criteria. Under these conditions, the expected average IQ of the selected adopted children would be considerably higher than the population of all adopted children. It is difficult to estimate exactly how much the selection process will bias the average IQ of these adopted children, but the effect of selection factors on the children's average IQ is serious. It is important to note that although the average IQs of the children would be affected, the parent-offspring correlations would not be influenced by selection bias.

The second criterion dealt with differential attrition of the adoptive sample over time. There is no doubt that this sample suffered differential attrition. They began with 180 subjects and finished with only 100. Skodak and Skeels (1949) indicated that the major factor in the reduction of the size of the sample had been time and expense. The families, all originally in Iowa, had now scattered over many states and Canada. Based on comparisons between the mean IQs of the group at various re-examination periods, Skodak and Skeels concluded that differential attrition was not evident. Munsinger (1975) indicated that biased attrition was occurring and estimated that the bias was approximately four IQ points.

The third criterion dealt with the selective placement of the adopted children. Munsinger (1975) reported a correlation between biological mothers' education and adoptive midparent education of .27. Thus, there was selective placement. Skodak and Skeels proposed that this may be responsible for the low correlations that were found between the adoptive midparent education and their adopted children's IQs.

With respect to test validity, there is a serious validity problem with Test 1 and Test 2. Several longitudinal studies have concluded that infant tests have virtually no validity in predicting subsequent intelligence test performance (Bayley, 1970; Lewis, 1973; McCall, Hogarty, & Hurlburt, 1972). A major reason for these extremely low correlations between infant tests and subsequent performance is to be found in the changing nature of the intelligence test.

Infant intelligence tests measure motor skills whereas later tests evaluate verbal ability. However, Test 3, which was administered at around 7.5 years, and Test 4, at around 13.5 years, are certainly valid tests of adult intellectual potential. Recall that these two later tests also produced high correlations between the adopted children's IQs and the biological mother's IQs.

The fifth criterion dealt with parental data. These data represent one of the most complete pools of information that has been published. This is certainly not a problem with this study especially when compared with previously published data (Burks, 1928; Leahy, 1935). In addition, the methodological criterion of early separation and placement was also adequately met by Skodak and Skeels. However, as previously described, this requirement of early placement introduces a selection bias toward obtaining brighter adopted children. Once again, however, it is necessary to point out that the parent-offspring correlations would not be affected by this practice effect.

The seventh criterion dealt with parental age, and it is difficult to evaluate this criterion. More specifically, Skodak and Skeels have only indicated that the biological fathers were young. Since we cannot determine their exact ages, one would tend to rely on other studies to evaluate this issue. Other studies have shown that the biological parents tend to be much younger than are the adoptive parents. Thus their levels of education and socio-occupational status would be biased toward a low value because of their relative youthfulness.

The eighth criterion was practice effects, and there is little doubt that practice effects occurred among the adopted children in this study. Each child in the final follow-up study had been given the 1916 version of the Stanford-Binet Intelligence Scale at least four times. It has been found that with two administrations of the Stanford-Binet the practice effects have been as high as eight points. So, four or more exposures would tend to significantly raise the average performance of all the adopted children.

Finally, several potential statistical problems were previously described. Munsinger (1975) has concluded that the Skodak and Skeels study did not suffer from these statistical biases. In other words, serious differential variance, regression to the mean, and shrinkage due to multiple correlation were not found.

Munsinger (1975) has conducted a more recent adoption study. Estimates of biological and adopting parents' intelligence and their common offsprings' IQ were gathered for 20 Mexican-American and 21 white families. The children in this sample were all separated from their biological mothers at birth. The average ages of the infants at adoption were 3.90 months for Mexican-American families and 2.76 months for white families. All children were placed before six months of age. Lorge-Thorndike IQ scores were obtained at an average age of 8.5 years on all children. All parents' protocols contained information about education, occupation, and age.

The data for the two ethnic groups were identical, so they were essentially pooled into one group. The correlation between adopting midparent's socio-economic status and their adopted children's IQs was -.14, whereas the correlation between biological midparent's socioeconomic status and their adopted away children's IQs was .70. These data suggest that adopting parents' socio-economic index is not related significantly to their adopted children's intelligence, and the biological parents' socio-economic index is strongly related to their children's intelligence, even though the children were separated from their biological mothers at birth. There was no significant selective placement among the two ethnic groups.

Once again it is important that the adequacy of these data be evaluated with respect to these methodological criteria. As previously described, the issue of biased initial sampling must be addressed. In this study, the San Diego City and County Schools provided IQs for those adopted children still residing in this area. From an initial sample of about 400 family groups, there were 21 white and 20 Mexican-American family groups with complete data on the biological parents, the adoptive parents, and the adopted children. There were many reasons for incomplete data: (a) there was often no information on the biological father; (b) parents' ages were unknown; (c) the family had moved from the San Diego area, and there was no way to assess the child's IQ. Munsinger concluded that the families with complete information differed in no significant way from the total population of families listed in the San Diego County Adoption Agency records. The fact that no mental illnesses or birth difficulties were reported among this sample of adopted children suggests that they may be a selected subset. Because of this selection against defect and mental illness, these adopted children were probably phenotypically superior to the total population. In fact, the average IQ of adopted children reported in the literature is around 106, rather than 100. Thus, this was a biased sample of the total population of the possible adoptive children.

The second criterion has been concerned with the differential attrition of the adoptive sample over time. Since this study was not a longitudinal study, this concern with longitudinal measures is not relevant here.

The third issue is that of selective placement. The product-moment correlation between biological midparent and adopting midparent social-education index is -.22 for the Mexican-American families and .07 for the white families. These nonsignificant relationships led Munsinger to recognize trivial selective placement.

The fourth criterion is concerned with test validity. There is no problem with unreliable IQ measurement of the children because they were all given a Lorge-Thorndike test around eight years of age. The intelligence estimates of the parents in this study were based on socioeconomic class rather than directly estimated by IQ testing. Although this is a potentially unreliable influence on the results, it was present for both adoptive and biological parents and therefore makes a comparison of their relative association with the children's IQ scores possible.

The fifth problem concerns the lack of information on biological parents. In this study, complete information was gathered about both biological and adoptive parents.

The sixth problem concerned the ages at separation and placement. All children were separated early and placed prior to six months of age. Thus, it is readily apparent that each of these criteria were satisfied in this study.

The seventh problem dealt with the issue of parental age. The biological parents averaged around 20 years of age when they relinquished their child, while the adopting parents were typically in their early 30s when they received the baby for adoption. Consequently, estimates of the biological parents' cognitive abilities based on indicators such as number of years of education or occupational status will be low, provided their education has not been completed, compared with the same estimates for the older adoptive parents. According to Munsinger, the differential age of the biological and adoptive parents creates a problem for the comparison of group means, but not for correlations. Thus, this age difference doesn't create a significant problem with this study.

The eighth problem is concerned with practice effects. This factor is not relevant here since all subjects were only tested one time.

Finally, the ninth criterion concerned statistical problems. As previously mentioned, these include: (a) restriction in the range of scores within a group, (b) possible regression to the mean with repeated measurements, and (c) the fact that multiple regression predictions may include chance associations between particular variables and the child's IQ score. With respect to restriction in the range of scores, moderate differences in the standard deviations of the biological and adoptive parents were observed. More specifically, the variation of the biological parents' social ranks was smaller than the variation of the adoptive parents' social ranks. Recall that restriction in the range of scores within a group will probably lower the correlation of these scores with any other variable. Accordingly, this restriction in range may have lowered the correlation between biological midparents' socioeconomic status and their children's IQs rather than the correlation between adoptive midparents' socioeconomic status and their adopted children's IQs. Thus, restriction in range cannot explain the difference in association between the children's IQ and the parents' socioeconomic status for adoptive and biological families. In addition, there could be no regression to the mean because only one measure was taken, and there could be no shrinkage of multiple correlation estimates because one wasn't computed.

In the past adoption studies, few families had both adopted and biological children (Burks, 1928; Munsinger, 1975; Skodak and Skeels, 1949). Scarr and Weinberg (1977) have conducted an investigation of the similarities in IQ scores among members of families with adopted and biological children. All of the adopted children were adopted in the first year of life. Both parents and all children in the family over four years of age were given an age-appropriate IQ test. Both parents and all children 16 years of age and older completed the WAIS. The WISC was administered to children between 8 and 15, and children between 4 and 7 were administered the

Stanford-Binet Intelligence Scale.

Analysis indicated that the intraclass correlations of standardized IQ scores between the adopting parents and their biological children were higher than those between the adopting parents and their adopted children. The biological midparent-biological child intraclass correlation was .49 and the adoptive midparent-adopted child intraclass correlation was .26. These correlations between parents and children were all lowered by the restricted range of the parents' IQ scores. The standard deviations of the parents' scores were approximately two-thirds those of the standardization population. The intraclass correlations, corrected for restriction of range, were .64 and .37, respectively. Heritability estimates were also determined. The heritability estimate used was $2(\underline{r}_{RP-C} - \underline{r}_{UP-C})$, where $\underline{r}_{RP-C}$ was the intraclass correlation of genetically related parent-child pairs, and $\underline{r}_{UP-C}$ was for unrelated parent-child pairs. The first heritability estimate was .46. A second heritability estimate calculated by the same formula but with the intraclass correlations corrected for the restriction of range in the parents' IQ scores, was .54. A third heritability estimate, calculated by the same formula but with the intraclass correlations corrected for both restriction of range and selective placement, was .66.

As with the other adoption studies, this research must be evaluated with respect to the previously described criteria. To begin, the issue of biased initial sampling must be addressed. There is evidence of biased initial sampling in this study. This is reflected in the fact that both the biological and the adopted children in these families had mean IQ scores in the high average range of intellectual functioning.

The second criterion is concerned with differential attrition. This issue does not apply here since this was not a longitudinal study.

The third issue is that of selective placement. A procedure to test for the effects of selective placement on the adoptive and biological parent-child correlations was suggested by Horn, Loehlin, and Willerman (1975). This procedure basically involves correlating the education of the adoptive children's natural parents and the IQ scores of the biological children of the adoptive parents. Since these individuals are neither genetically related nor live together, any similarity which exists between them must be due to selective placement. The correlations between the education of the biological parents of the adopted children and the IQ scores of the biological children of the adoptive parents ranged from .14 to .21. As previously reported, this correction was utilized in calculating some of the intraclass correlations and heritability estimates in this study. Thus, the criterion of selective placement was satisfied through statistical correction.

The fourth criterion is concerned with test validity. Parents and all children 16 years of age and older were given the WAIS. The WISC was administered to children between 8 and 15, and children between 4 and 7 were tested with the Stanford-Binet Intelligence Scale. Thus the tests which were administered were both age-appropriate and valid tests of intelligence.

The fifth criterion concerns a lack of reliable information about biological parents. As a result of utilizing this particular design, it was not necessary to collect data from the biological parents of the adopted children.

The sixth criterion deals with the requirement of early separation and placement. The adopted children in this study were all adopted by 12 months of age. It is not clear when these children were separated from their biological parents, nor is it clear what proportion of them experienced preadoption placements. The fact that many of these children weren't adopted by six months of age indicates that this criterion has not been fully satisfied in this study.

The seventh criterion dealt with the common finding that the mean age of biological parents is approximately 10 years less than the mean age of adoptive parents. Again, due to the particular design of this study, this issue is not relevant.

The eighth source of possible bias concerns differential practice effects. Since all subjects were tested only once, the potential for any practice effects was clearly eliminated.

Finally, with respect to any statistical problems, the corrections for the restriction in the range of scores has already been mentioned. Thus, careful examination of this research indicates that most of these methodological considerations have been satisfied.

Horn, Loehlin, and Willerman (1979) have conducted the Texas Adoption Project. Estimates of biological mothers' and adopting parents' intelligence and their common offsprings' IQs were gathered for 300 families. A total of 119 families contained adopted children as well as biological children. All of the adopted children in this study were permanently separated from their biological mothers within the week following birth and placed in the adoptive homes at an early age. The WISC was administered to all 5- to 16-year-old children. Children above the age of 16 were given the WAIS while those 3 or 4 years of age were given the Stanford-Binet (1960). Children below the age of 3 were not tested. Since each unwed mother took the Revised Beta Examination, each adoptive parent was given this test as well as the WAIS.

Analyses indicated that biological parent-child IQ correlations were consistently greater than adoptive parent-child IQ correlations. The biological mother-child correlation was .31 whereas the adoptive mother-child correlation was .17.

As with the previous studies, it is important to consider whether this study satisfies each of the previously described methodological criteria. Once again, the question of biased initial sampling must be answered. For the years 1963-1971 a total of 1,848 women gave up their children for adoption through a particular agency. A search of the files indicated that 1,381 had been given the Revised Beta

Examination. Horn et al. expected to be able to obtain test information from at least 442 adoptive families. This expectation was based on the assumption that they would be able to secure the cooperation of the same proportion of adoptive parents that had responded to a previous survey conducted by this particular agency. In 1966, questionnaires were sent to all parents who adopted children through the agency between 1953 and 1965. A total of 76% of the questionnaires were returned with 70% of these respondents still residing in the state, and with 60% of these residents living in communities of 100,000 or more population. Since these investigators were interested in state residents in the large population centers, they expected to secure cooperation from about 32% of the 1,381 parents on the list. The parents who adopted the children of these tested unwed mothers were then contacted by mail and asked if a representative of the agency could visit with them and describe the study. A total of 416 families were interviewed and agreed to be tested, but funds restricted actual testing to 300 families. The major factors which determined whether or not an eligible family was tested were any scheduling difficulties and the distance from a testing site. Thus the sample of unwed mothers was ascertained indirectly through the decision of their children's adoptive parents to participate in the study.

Analyses indicated that the unwed mothers included in the sample appear to be quite representative of the population of women tested in the agency during this period. The above average intellectual level of this population (mean IQ was 108.2) is probably best explained by the fact that the agency requested significant amounts of money from the parents of the unwed mothers. This practice would reduce the number of girls from lower socio-economic backgrounds who were referred to this agency. In addition, since the data from the adoptive parents and the children were significantly above average in intelligence, there is evidence indicating that this initial sample was biased.

The second criterion concerning differential attrition often found in longitudinal studies does not apply in this study. The third criterion addresses the issue of selective placement. The correlation between biological mothers' and adoptive parents' IQs was approximately .17. Thus, one may conclude that there was modest selective placement practiced by the agency.

The fourth criterion is related to test validity. The selection of tests for administration to each member of the adoptive family was guided by two basic considerations: the adoptive parent data should match the data available on the unwed mothers and there should be as much continuity as possible between the measures given the adoptive parents and their children. The actual tests that were administered have already been discussed. It is evident that these tests are certainly valid tests of intelligence.

The fifth criterion concerns the lack of information on biological parents. Complete intellectual information was gathered on biological mothers and adoptive parents. IQ data on the biological fathers were not available.

The sixth problem concerns the age of separation and placement. All children were separated early and placed in their adoptive homes at an early age.

The seventh criterion deals with parental age. Biological mothers were significantly younger than the adoptive parents. Most of the adoptive parents were in their 30s and 40s when they were tested. However, with respect to intelligence, this is not a major problem since IQ measurement corrects for age differences automatically. This is generally accomplished by comparing an individual's performance with that of a normative sample in his or her own age group.

Finally, the eighth issue of practice effects doesn't apply in this study. In addition, there were no statistical problems. Thus, the ninth criterion has been satisfied as well.

In summary, several adoption studies have been evaluated with respect to various methodological criteria. It is apparent that some of these criteria are more easily satisfied than others. For instance: it is impossible in principle to meet the criteria of representative sampling and early separation and placement in the same sample. In essence, it has been observed that early-placement adopted infants are selected from a higher family IQ pool than are later-placement adopted children. Despite some of these problems, as seen in Table 3, a summary of the results of these four adoption studies, a general trend can be observed. At this point, it is necessary to mention that several additional adoption studies have provided evidence consistent with these previously described studies. The results of research by Leahy (1935) and Fisch, Bilek, Deinard, and Chang (1976) are presented in Table 4. Table 5 presents the results of a recent

Table 3

A Summary of the Results of Four Adoption Studies

Skodak & Skeels (1949)

Adoptive mother's IQ - Adopted child's IQ

Test 1	-.03	Test 3	.10
Test 2	.04	Test 4	.04

Biological mother's IQ - Adopted child's IQ

Test 1	.00	Test 3	.35
Test 2	.28	Test 4	.38/.44

Biological mother's education - Adopted child's IQ

Test 1	.04	Test 3	.37
Test 2	.31	Test 4	.31

Munsinger (1975)

Adoptive midparent's SES - Adopted child's IQ: -.14

Biological midparent's SES - Adopted child's IQ: .70

Scarr & Weinberg (1977)

Adoptive midparent's IQ - Adopted child's IQ: .26

Biological midparent's IQ - Adopted child's IQ: .49

Horn, Loehlin, & Willerman (1979)

Adoptive mother's IQ - Adopted child's IQ: .17

Biological mother's IQ - Adopted child's IQ: .31

Table 4

Parent-Offspring Correlations for Various Measures of Intelligence

Factor	Adopted Child	N	Biological Child	N
Leahy (1935)				
Midparent Otis IQ	.18	177	.60	173
Midparent Stanford-Binet Vocabulary	.24	174	.56	164
Fisch et al. (1976)				
Maternal IQ WISC (4 yrs.)	.07	94	.35	50
Maternal IQ WISC (7 yrs.)	.08	94	.26	50

sibling study conducted by Teasdale and Owen (1984). Based on all of these adoption studies, it is very clear that the adoptive parents' IQ and home environment have only a slight effect on their adopted children's intellectual development. In contrast, the heredity of the biological parents have a stronger effect on their own children's intellectual development.

Twin Studies. Twin studies provide another opportunity to estimate the relative influence of nature and nurture on intelligence. By comparing the average within-pair differences of monozygotic twins (MZ) with the average within-pair differences of dizygotic twins (DZ), it is possible to obtain an estimate of the importance of heredity in the determination of intelligence. The basic logic for conducting a twin study is as follows. Within-pair differences of MZ twins can only be due to nongenetic (environmental) factors because MZ twins have identical genetic makeups. On the other hand, the within-pair differences of DZ twins are due to hereditary differences as well as environmental ones. Therefore, DZ twins should be more dissimilar than MZ twins on traits which are under genetic control. In this way, the discrepancy between the two sets of

Table 5

Intraclass Correlations for Siblings' Intelligence

Sibling Relationship	BPP Correlation	*N*
Teasdale & Owen (1984)		
Full sibs apart	.47	28
Half-sibs apart	.22	64
Maternal half-sibs apart	.11	34
Paternal half-sibs apart	.30	30
Unrelated sibs reared together	.02	24
Full sibs reared together	.52	73

differences for any specific trait will constitute an indication of the degree of genetic control over that trait.

There are certain methodological assumptions that must be met before a twin study of intelligence will be considered valid. Firstly, there must be careful ascertainment of zygosity for same-sex twin pairs. One method to establish zygosity involves blood-typing procedures. It is possible to perform highly accurate blood analyses. Members of each pair of twins are compared for various antigens in the blood. Tests are conducted for the following factors : A, B, O, M, N, S, s, P_1, P_2, Rho, $rh^{'}$, $rh^{''}$, Miltenberger, Verweyst, Lewis, Lutheran, Duffy, Kidd, Sutter, Martin, Kell, Cellano, and occasionally some others. Twin pairs concordant on all antisera tests are classified as monozygotic, while pairs discordant on one or more tests are classified as dizygotic. Smith and Penrose (1955) have presented the method for calculating the probability that twins concordant for a set of blood groups are, in fact, DZ twins. If all available antisera are used the probability that concordant sets are DZ is extremely small.

Another method to determine zygosity involves several additional phenotypic criteria. More specifically, Slater (1963) has used the following heritable phenotypic characteristics to differentiate between monozygotic and dizygotic twin pairs: hair color, hair type, eye color, standing height, fingertip ridge count, blood type, and shape of ears. Slater assumes that if there is any difference between the twin pair in heritable phenotypic characteristics, the difference automatically means the twins are dizygotic. On the other hand, the assignment of pairs to the monozygotic twin category can only be made on a probabilistic basis since there is a small possibility that dizygotic twins could be concordant for several hereditary traits by chance alone. The procedure is to compute the joint probability of obtaining concordance between twin pairs, using individual probabilities for each trait, based on the relative frequency of phenotype characteristics. The criterion for including a twin pair in the monozygotic class is often a chance concordance probability of less than .0001. Recently, a number of studies have reported that physical similarity indexes have had approximately 95% accuracy when cross-validated against blood diagnosis (Nichols and Bilbro, 1966; Cohen, Dibble, Grawe, & Pollin, 1973, 1975; Kasriel & Eaves, 1976). Cohen et al. (1973) had mothers of 155 sets of twins complete a questionnaire. In addition to family background information, the mother was asked whether she believed the twins were monozygotic or dizygotic or if she was uncertain about zygosity. The remainder of the questionnaire consisted of 10 questions that were known to differentiate between monozygotic and dizygotic twins. Six of the questions concerned the degree to which the twins were similar for specific physical characteristics: height, weight, facial appearance, hair color, eye color, and complexion. These were rated on a 3-point scale: "not at all similar," 0; "somewhat similar," 1; and "exactly similar or identical," 2. Four questions involved general identity and confusion: "Do they look as alike as two peas in a pod?" "Does either mother or father ever confuse them?" "Are they sometimes confused by other people in the family?" and finally, "Is it hard for strangers to tell them apart?" These four questions were scored dichotomously: "no," 0; or "yes," 1. The results of a discriminant function analysis produced clear differentiation between the monozygotic and dizygotic groups. In addition, there was 98% agreement between zygosity assigned by the responses to the questions and zygosity determined by blood typing.

Cohen et al. (1975) demonstrated the cross-validity of this questionnaire method. The mothers of a new population of 275 sets of same-sex twins completed the previously described questionnaire. The results obtained by the use of the discriminant function method with this new population were almost identical with those found with the previous population. In addition, to study test-retest reliability of the zygosity questionnaire, parents who completed the initial set of forms were asked by letter if they would be interested in continued participation. The parents of 199 of the 275 pairs filled out the zygosity questionnaire about one year after their first completion of the form. There was a very high correlation between the initial scores and the replication scores (r = .97). Thus, test-retest reliability appeared to be well substantiated.

Kasriel and Eaves (1976) had 178 pairs of twins complete a questionnaire which included two questions relating to similarity: "In childhood, were you frequently mistaken by people who knew you?" and "Do you differ markedly in physical appearance and coloring?" The responses to these questions by all the twins who had been blood-grouped were tabulated in order to relate the diagnosis of zygosity on the basis of blood grouping to the twins' own judgment of their similarity. If all twins who agree that they were confused in childhood and are alike in appearance were to be classified as MZ on the basis of the questionnaire alone, and all others as DZ, the diagnoses of only seven pairs out of the total 178 (3.9%) would disagree with those based on blood group data.

A second methodological criterion that should be addressed is the issue of environmental bias in twin studies. More specifically, the comparisons of intraclass correlations for MZ and DZ pairs and the resulting estimates of heritability are based upon the assumption that the environments of DZ co-twins are no more dissimilar than the environments of MZ co-twins. This assumption has been seriously questioned and has been the source of much criticism of twin studies. Since excess MZ correlations are interpreted as genetic in origin, additional DZ differences created by greater environmental variance would lower DZ intraclass correlations. This would then bias the results in favor of high genetic estimates.

Although Smith (1965) and Scarr (1968) have reported data supporting the contention that the home environments of MZ co-twins are generally more similar than those of DZ co-twins, they proceeded to explain why these findings do not invalidate twin research. If parents are simply reacting to the existing differences between their DZ twins' behavior, then no bias is introduced into twin studies. However, if the parents are training differences, then these environmentally determined differences would bias the comparisons of intraclass correlations in favor of the genetic hypothesis, by reducing the possible similarities of DZ co-twins. Similarily, the parents of MZ twins who know their twins are identical may react to existing similarities or seek to train greater similarities than would ordinarily exist. When parents are correct about their twins' zygosity, it is impossible to distinguish between parental behavior of their twins and parental treatment that seeks to train greater differences and similarities.

Scarr (1968) has described a method for estimating environmental bias. Parents who are not correct about their twins' zygosity offer a critical test of environmental bias in twin studies. When parents are wrong about their twins' zygosity, it is possible to separate parental reactions to similarities and differences based on genetic relatedness from parental behaviors which arise from their belief that their twins should or should not be similar. Although the samples reported are small, a clear trend is apparent. The data generally confirm that genetic relatedness of the twins determines the similarity of parental treatment. In other words, the mothers of MZ twins, whom they wrongly believe to be DZ, treat them more like correctly identified MZ twins. Likewise the mothers of DZ twins, whom they believe to be MZ, treat them more like correctly classified DZ pairs. Despite the mother's erroneous beliefs, the twins are recognized as having similarities and differences appropriate to their degree of genetic relatedness.

In addition to studying the effect of labeling a twin pair as fraternal or identical, there is another type of study that can be conducted addressing the question of equal environments. This approach basically asks whether or not observed differences in the environment of the two types of twins make a behavioral difference. Loehlin and Nichols (1976) reported data from a large scale twin study of personality and ability. As with previous studies, their data indicated that the identical twins experienced slightly more similar environments than did the fraternal twins. They went on to try to determine whether these environmental differences affect behavior. Because MZ twins are genetically identical, differences within pairs of MZ twins can be only caused by environmental factors. Some identical twin pairs are more similar behaviorally than others, and some identical twins are exposed to more similar environments than others. Loehlin and Nichols tested the adequacy of the equal environments assumption by correlating identical twin differences for cognition and personality with the measures of similarity of environment. If similarity of environment makes a difference behaviorally, then identical twins whose environments are more similar should be more similar behaviorally. A positive correlation would mean that identical twins who are exposed to more similar environments were more similar in behavior. The analysis yielded low correlations. The median correlation was .06. The range of correlations was -.15 to .22. It is clear that the greater similarity of identical twins' environments cannot plausibly account for more than a very small fraction of their greater observed similarity on the personality and ability variables in this study. In other words, environmental variables that were unequal for identical and fraternal twins did not make a difference in cognition, personality, vocational interests, or interpersonal relationships.

Vandenberg and Wilson (1979) conducted a more recent study to determine whether the strength of the influence exerted by the twin situation is related to the magnitude of twin differences in performance on cognitive ability tests. A 28-item questionnaire was designed to measure the closeness of the relationship between the twins as seen by themselves and by their mothers. Only 7 of the 28 indices differed significantly between MZ and DZ twins. Analyses were performed to determine whether these 28 indices were correlated with twin differences in scores on the six Primary Mental Abilities (PMA) subtests. The 168 correlations ranged from .26 to -.29 and most did not differ significantly from zero. Thus, Vandenberg and Wilson found that the environment of MZ twins differs somewhat from that of DZ twins. However, once again these differences were not sufficiently strong to be related to twin differences in cognition. In summary, this evidence suggests that environmental determinants of similarities and differences between MZ and DZ co-twins are not as potent as the critics charge.

A related methodological issue concerns the assumption that prenatal conditions of MZ twins be similar to the prenatal conditions of DZ twins. A positive association has been reported between the birthweight and the adult intelligence of MZ twins. However, no such association exists in the case of DZ twins. Investigators consistently find that the heavier-birthweight MZ twin is brighter than his lighter-birthweight partner. This suggests that some factor in the prenatal environment of MZ twins is affecting both birthweight and their IQ, and that the factor is not present in the prenatal environment of DZ twins. Munsinger (1977) described a phenomenon known as the identical twin transfusion syndrome. The identical twin transfusion occurs when blood leakage between the pair of MZ twins encased in a single chorion is great enough to produce a 35% decrement in hemoglobin for one twin compared with the other. The important point to realize is that the inclusion of MZ twins with very different birth weights (>500 g) in a behavior genetic analysis will tend to decrease estimates of heritability.

A fourth methodological issue concerns possible volunteer bias in twin research. Several investigators have described this potential problem in great detail. Blewett (1954) noted that the method of sampling through the schools is subject to selective error. The twins most likely to be overlooked by the teachers are those fraternal twins who are quite dissimilar in appearance and behavior. This selective factor leads to an overestimation of the closeness of relationship between results obtained by fraternal twins. If this sampling bias is operative, its tendency is to reduce the difference between MZ and DZ intraclass correlations.

Lykken, Tellegen, and DeRubeis (1978) reported that studies of adult same-sex twins which rely upon volunteer subjects typically consist of about two-thirds female and two-thirds monozygotic (MZ) pairs. They have demonstrated that because of this recruitment bias, the male and dizygotic (DZ) twins in such studies will show smaller between-pair variance on many traits. This reduction in between-pair variance results in underestimation of the intraclass correlation of DZ twins and an overestimation of the heritablity of the trait under study. It was suggested that sufficient extrinsic incentive be provided to overcome this strong recruitment bias.

It is helpful to keep these methodological issues in mind as the twin study literature is reviewed. Several representative studies will be described. To begin, Blewett (1954) administered the Primary Mental

Abilities test (PMA) to 26 pairs of MZ twins and 26 pairs of same-sex DZ twins. All subjects ranged in age from 12 to 15 years. The intraclass correlations for PMA Total Score were .754 and .394 for the MZ and DZ twins, respectively.

Nichols (1965) collected scores on the National Merit Scholarship Qualifying Test (NMSQT) which had been taken by high school juniors. A total of 687 pairs of MZ twins and 482 pairs of DZ twins participated in this study. The MZ and DZ intraclass correlations were .87 and .63, respectively.

Block (1968) administered the Wechsler Adult Intelligence Scale (WAIS) to 60 pairs of MZ and 60 pairs of same-sex DZ twins. There was an equal number of pairs in each of six age groups ranging from 13 to 18 years. The within-pair variance of the DZ twins was compared with the within-pair variance of the MZ twins. An $\underline{F}$ yielded a value of 3.47 ($\underline{p} < .001$). In addition, Holzinger's $\underline{H}$ measure of heritability was .71.

Loehlin and Nichols (1976) updated the results of high school students who were administered the NMSQT. Now based on 1,300 MZ pairs and 864 DZ pairs, they reported intraclass correlations of .86 and .62, respectively. In addition, they reviewed 18 studies where various intelligence measures were administered to more than 25 pairs of MZ twins and 25 pairs of DZ twins. Based on 2,154 pairs of MZ twins, an intraclass correlation of .86 was reported. Similarily, based on 2,021 DZ twins, an intraclass correlation of .61 was presented.

Wilson (1977) administered the Wechsler Intelligence Scale for Children (WISC) to 7 and 8 year old twins. An intraclass correlation of .86 was indicated for 74 pairs of MZ twins. Likewise, $\underline{r}$ = .60 for 56 pairs of DZ twins. Segal (1985) also administered the WISC-R to MZ and DZ twin children. The mean age of the participants was 8.03 years. The intraclass correlation for full-scale IQ score was .85 for 69 pairs of MZ twins and .46 for 35 pairs of DZ twins.

Finally, Bouchard and McGue (1981) reviewed 111 familial studies of intelligence. Based on 4,672 MZ pairs they reported a median correlation of .85 and a weighted average of .86. For 5,546 DZ pairs, the median correlation was found to be .58 and the weighted average was .60.

In summary, the intraclass correlations for MZ twins have been consistently greater than the corresponding values for DZ twins. In addition, a careful chronological analysis of the studies indicates a trend. Plomin and DeFries (1980) have suggested that individual differences in IQ are less heritable than previously believed. In essence, the smaller difference between the identical and fraternal twin correlations in recent studies results in an estimate of broad-sense heritability closer to .50 than to the older estimate of about .70. In this way twin data suggest with respect to intelligence that about 50 percent of the phenotypic variance is due to genetic variance.

Behavior Genetics of Specific Cognitive Abilities

Once it became evident that there is a strong genetic component influencing general intelligence, many investigators turned their attention to more specific cognitive abilities. In other words, they became concerned with the question of whether the genetic influence was equal for all skills or whether there appeared to be differential heritabilities for these various cognitive skills. Researchers have examined numerous cognitive capacities including spatial, verbal, and numerical ability, perceptual speed, visual memory, divergent thinking, and abstract reasoning. Investigators have primarily conducted twin and family studies.

In a large study of mental ability, known as the Hawaii Family Study of Cognition (HFSC), data were obtained on 15 cognitive tasks. The families were Americans of European ancestry (AEA) or of Japanese ancestry (AJA). The results reported here were made possible by the efforts of G. C. Ashton, R. C. Johnson, M. P. Mi, and M. N. Rashad at the University of Hawaii and J. C. DeFries, G. E. McClearn, S. G. Vandenberg, and J. R. Wilson at the University of Colorado.

The battery of 15 cognitive tests was administered to parents and their offspring (see DeFries et al., 1976, for list of tasks and reliabilities). DeFries et al. (1974, 1976) analyzed phenotypic correlations among the 15 tasks, performing principal component analyses with varimax rotations. Four readily interpretable factors emerged: spatial visualization, verbal ability, perceptual speed, and visual memory. As noted in DeFries et al. (1976), since environmental correlations between family members are likely to be greater than zero, the regressions of offspring on midparental values can only be regarded as measures of phenotypic similarity and upper-bound estimates of heritability. With respect to the four rotated factors, the midparent-offspring regressions were moderately high for verbal and spatial ability, but somewhat lower for visual memory and perceptual speed (DeFries et al., 1976, 1979). More specifically, DeFries et al. (1976) reported midparent-offspring regressions for these four factors. The regressions for the spatial, verbal, perceptual speed, and visual memory factors were: .65 ± .04, .61 ± .04, .46 ± .04, and .44 ± .05, respectively, for 739 AEA families. The corresponding regressions for 244 AJA families were: .58 ± .07, .44 ± .08, .32 ± .09, and .31 ± .09.

Park et al. (1978) administered 14 of these 15 tasks to Korean families. Phenotypic correlations among the 14 cognitive variables were obtained and subjected to principal component analysis as previously described. Although the same four-factor solution was obtained, a more interpretable pattern of factor loadings was observed when a five-factor solution was forced. The additional factor was labeled inferential ability. The regressions of offspring on midparent value tended to be considerably higher for the Korean sample than the AEA and AJA families. More specifically, 9 of the 14 regressions were significantly higher for the Korean sample.

Two possible explanations were considered. First, tests in Korea were administered to individual nuclear family units, whereas those in Hawaii were administered to large groups. It is possible that this difference in test administration may have increased the Korean parent-offspring regressions to some extent. However, it is highly unlikely that this factor is solely responsible for these differences. A second issue to be considered is that of assortative mating. Assortative mating results in a higher heritability. Assortative mating increases the variance in the sense that the offspring differ more from the average than they would if mating were random. In other words, assortative mating adds to the genetic similarity between parents and their offspring. This increase in the additive genetic variance results in a higher heritability. Spouse correlations were considerably higher than those for AEA and AJA. Thus one would predict higher heritability estimates in the Korean sample.

The findings reported by Park et al. differ from the previously described studies in another important way. Whereas previous studies reported regressions for verbal and spatial dimensions that were somewhat greater than those for perceptual speed and visual memory, this trend was not indicated in the Korean sample.

Spuhler and Vandenberg (1980) administered a battery of cognitive tests identical to that used in Hawaii to a set of Caucasian families in Boulder, Colorado. This sample, like the Korean sample, did not show a trend of differential regressions for these specific cognitive skills. Although the regressions for parent-offspring resemblance in the four samples tended toward the midrange values, the rankings of the values for the cognitive tests was not highly consistent across the samples. This finding prevents a definitive general conclusion to be made concerning differential heritability for specific cognitive abilities in these family studies.

Spatial Ability. In addition to these researchers in Hawaii and Colorado, many other researchers have also striven to measure the genetic influence on various specific cognitive abilities. Spatial ability, in particular, has received a great deal of attention. Table 6 presents intrafamilial correlations of spatial test scores from many of these investigations as well as the HFSC studies previously mentioned.

It has been proposed that spatial ability may be transmitted by an X-linked recessive gene. A recessive X-linked trait will only be expressed in

Table 6

A Summary of Family Correlations for Tests of Spatial Ability

Studies	F-S	M-D	M-S	F-D
Stafford (1961)				
Identical Blocks Test	.02	.14	.31	.31
Hartlage (1970)				
Differential Aptitude Test	.18	.25	.39	.34
Bock and Kolakowski (1973)				
Guilford-Zimmerman Spatial Visualization Test, Form B	.15	.12	.20	.25
Guttman (1974)				
Raven's Progressive Matrices				
Subtest A-Continous Patterns	.22	.09	.04	.14
Subtest B-Analogies	.15	.33	.23	.06
Subtest C-Progressive Development of Figures	.31	.25	.02	.01
Subtest D-Permutations and Alternations of Patterns	.32	.23	.14	.03
Subtest E-Resolving Figures Into Constituent Parts	.28	.38	.08	.46
Total Score	.36	.39	.24	.23

(table continues)

Table 6 (continued)

Studies	F-S	M-D	M-S	F-D
DeFries et al. (1976) AJA/AEA[a]				
Elithorn Mazes	.08/.07	.04/.17	.14/.17	.19/.07
Shepard-Metzler Mental Rotations	.30/.15	.10/.32	.14/.16	.34/.23
Paper Form Board	.24/.27	.24/.36	.26/.30	.21/.40
ETS Card Rotations	.26/.26	.09/.36	.12/.19	.11/.22
ETS Hidden Patterns	.10/.21	.18/.32	.12/.20	.26/.27
Raven's Progressive Matrices	.10/.20	.17/.29	.22/.38	.26/.27
Spatial Composite	.33/.27	.13/.41	.21/.28	.31/.32
Bouchard and McGee (1977)				
Mental Rotation Test	.23	.16	.20	.17
Loehlin, Sharan, & Jacoby (1978)				
Card Rotations	.27	.40	.27	.32
Cube Comparisons	.16	.19	.04	.17
Hidden Patterns	.40	.22	.44	.38
Paper Folding	.27	.21	.24	.30
Spatial Composite	.28	.28	.34	.30
McGee (1978)				
Mental Rotation Test	.23	.16	.20	.17
Hidden Patterns Test	.07	.35	.23	.39

(*table continues*)

Table 6 (continued)

Studies	F-S	M-D	M-S	F-D
Park et al. (1978)				
Elithorn Mazes	.42	.57	.52	.48
Mental Rotations	.22	.46	.26	.41
Paper Form Board	.59	.57	.63	.53
Card Rotations	.12	.61	.36	.54
Hidden Patterns	.51	.65	.58	.56
Raven's Progressive Matrices	.33	.51	.25	.39
Spatial Composite	.33	.53	.37	.45
DeFries et al. (1979) AJA/AEA[a]				
Elithorn Mazes	.13/.11	.09/.17	.15/.16	.19/.07
Mental Rotations	.20/.20	.11/.30	.17/.13	.24/.20
Paper Form Board	.21/.28	.29/.33	.20/.29	.27/.35
Card Rotations	.24/.26	.17/.34	.10/.21	.11/.23
Hidden Patterns	.09/.24	.19/.32	.13/.24	.22/.27
Raven's Progressive Matrices	.26/.33	.22/.38	.20/.29	.32/.31
Guttman and Shoham (1979)				
Hidden Patterns	.24	.18	.17	.19
Hidden Blocks	.28	.12	.17	.13
Identical Blocks Test	.24	.25	.22	.25
Shepard-Metzler	.32	.21	.12	.08
Thurstone Figure Analogies Test	.36	.26	.27	.17
DAT-Abstract Reasoning Test	.34	.22	.23	.25
Ilana Rotations	.36	.18	.24	.16
Raven's Progressive Matrices	.25	.31	.30	.40

(*table continues*)

Table 6 (continued)

Studies	F-S	M-D	M-S	F-D
Corley, DeFries, Kuse, and Vandenberg (1980) AJA/AEA[a]	.25/.35	.28/.15	.10/.14	-.03/.13

Note. F-S, M-D, M-S, and F-D refer to father-son, mother-daughter, mother-son, and father-daughter correlations, respectively.

[a] AJA refers to American families of Japanese ancestry and AEA refers to American families of European ancestry.

hemizygous recessive males and homozygous recessive females. In addition, the X-linkage model predicts a characteristic pattern of family correlations. The model predicts a higher father-daughter (F-D) than father-son (F-S) correlation and a higher mother-son (M-S) than mother-daughter (M-D) correlation. More specifically, the model predicts the following pattern of family correlation: $0 = F\text{-}S < M\text{-}D < M\text{-}S = F\text{-}D$. The predicted F-S correlation is zero because the father doesn't pass an X chromosome to his son. The correlation between mother and daughter should be an intermediate value. The mother passes one of her two X chromosomes to her daughter. Finally, as indicated, M-S and F-D relationships should have the strongest correlations and they should be roughly equivalent as well. This occurs because the son inherits his only X chromosome from his mother and because the father passes his only X chromosome to his daughter.

Although the expected pattern of parent-child correlations were obtained in the earlier studies (Stafford, 1961; Hartlage, 1970; and Bock & Kolakowski, 1973), it has not been achieved in the more recent studies (Guttman, 1974; DeFries et al., 1976; Bouchard & McGee, 1977; Loehlin, Sharan, & Jacoby, 1978; Park et al., 1978; DeFries et al., 1979;

Guttman & Shoham, 1979; and Corley, DeFries, Kuse, & Vandenberg, 1980). These results do not provide evidence for the spatial-enhancing effect of the X-linked recessive gene. Thus, although there is a great deal of evidence indicating genetic influence with respect to spatial abilities, it is not clear how this is transmitted from parent to offspring.

Additional evidence for this genetic influence comes from twin studies. Osborne and Gregor (1966, 1968) administered a battery of psychological tests to MZ and DZ twins. Intraclass correlations were determined on selected spatial tests. All of the MZ correlations were greater than the corresponding correlations for the DZ twins; often these MZ-DZ differences achieved statistical significance.

Verbal Ability. Another widely studied cognitive capacity is verbal ability. Table 7 is a summary of the HFSC data on verbal ability. Regressions of midchild on midparent (± standard error) are presented for several verbal tests. The regressions are moderately high.

Twin studies also provide evidence for a genetic influence in verbal ability. Osborne, Gregor, and Miele (1968) gave MZ and DZ twins a battery of comprehensive tests which included the Heim

Table 7

A Summary of the HFSC Data on Verbal Ability: Regressions of Midchild on Midparent

Tests	DeFries et al. (1976)		Park et al. (1978)	DeFries et al. (1979)	
	AJA	AEA	Korean	AJA[a]	AEA[a]
Primary Mental Abilities (PMA) Vocabulary	.59 ± .07	.69 ± .04	.51 ± .07	.57	.67
Things (A Fluency Test)	.36 ± .09	.55 ± .04	.80 ± .06	.47	.55
ETS Word Beginnings and Endings	.55 ± .08	.56 ± .04	.80 ± .06	.59	.55
PMA Pedigree (A Reasoning Test)	.64 ± .07	.74 ± .04	.88 ± .04	.63	.72
Whiteman Test of Social Perception	.35 ± .07	.39 ± .05		.26	.38

(table continues)

Table 7 (continued)

Tests	DeFries et al. (1976)		Park et al. (1978)	DeFries et al. (1979)	
	AJA	AEA	Korean	AJA[a]	AEA[a]
Composite	.58 ± .07	.65 ± .04	.73 ± .06	.55	.61

Note. The regression coefficients are reported here ± standard errors.

AJA refers to American families of Japanese ancestry and AEA refers to American families of European ancestry.

[a] Standard errors for AEA coefficients range from .04 to .05 and from .06 to .08 for AJA coefficients.

Self-Judging Vocabulary Test, the Wide-Range Vocabulary Test, the Heim Vocabulary Test, and the Spelling Achievement Test. All the MZ intraclass correlations were greater than the corresponding DZ intraclass correlations, although most of the differences did not achieve statistical significance.

Perceptual Speed. Perceptual speed was examined by the HFSC studies as well as in a twin-family study by Rose, Miller, Dumont-Driscoll, and Evans (1979). Table 8 is a summary of the regressions of midchild on midparent for the ETS Number Comparisons Test. The MZ intraclass correlation for the Number Comparisons Test was .74. The corresponding DZ intraclass correlation was .56. The difference between these values was significant ($F = 2.07$, $p < .006$).

Memory. The HFSC administered two memory tests. The immediate visual memory test yielded regressions of midchild on midparent scores ranging from .21 to .49. The delayed visual memory test yielded slightly higher regression values ranging from .29 to .64. Composite measures for visual memory ranged from .25 to .55.

Pezzulo, Thorsen, and Madaus (1972) conducted a twin study and administerd a modified version of the Digit Span, an auditory memory test. Pezzulo et al.

Table 8

A Summary of Regressions of Midchild on Midparent for the ETS Number Comparisons Test

Studies	Regressions
DeFries et al. (1976)	
AJA	.40 ± .08
AEA	.48 ± .04
Park et al. (1978)	.68 ± .06
DeFries et al. (1979)[a]	
AJA	.36
AEA	.46
Rose et al. (1979)	.17 ± .07

Note. The regression coefficients are reported here ± standard errors.

AJA refers to American families of Japanese ancestry and AEA refers to American families of European ancestry.

[a] Standard errors for AEA coefficients range from .04 to .05 and from .06 to .08 for AJA coefficients.

used twins ranging in age from 10 to 15 years and did not partial out the effect of age on the twin correlations. Jensen and Marisi (1979) were provided with the original data of this study. The age-corrected twin intraclass correlations were $\underline{r}_{mz}$ = .716 and $\underline{r}_{dz}$ = .497. The conventional test of the significance of a genetic effect ($\underline{F} = \underline{MS}_{wdz} / \underline{MS}_{wmz}$) was significant.

Divergent Thinking. Vandenberg (1968) administered Guilford's Test of Divergent Thinking to identical and fraternal twins. Results indicated that only one of the nine subtests was significantly heritable.

Pezzulo et al. (1972) investigated the heritability of verbal and figural divergent thinking. MZ and DZ twins received the Torrence Test of Creative Thinking which was scored for fluency, flexibility, and originality. The results of the analysis of variance provided no evidence of heritability.

Olive (1972) conducted a study of siblings. Guilford's Tests were administered to 13 pairs of twins. Included were Word Fluency, Expressional Fluency, Associational Fluency, Ideational Fluency, Alternate Uses, Consequences--Obvious, and

Consequences--Remote. The results showed no evidence for a genetic factor.

Finally, Reznikoff, Domino, Bridges, and Honeyman (1973) investigated the possibility of a genetic component in creative ability. A battery of ten creativity tests, including five developed by Guilford, were administered to MZ and DZ twins. Intraclass correlations for the MZ twins were greater than for the DZ twins. However, further analysis led to only one significant F value. This research is compatible with the previously mentioned studies on divergent thinking. Although identical twins were somewhat more alike than fraternals on various measures of creativity, there is little compelling evidence to support the notion of a genetic component in creativity.

Numerical Ability and Quantitative Reasoning. Osborne and Miele (1969) investigated the heritability of numerical ability. Mukherjee's Simple Arithmetic Test was given to MZ and DZ twins, ranging in age from 13 to 18 years. This test has seven parts involving the fundamental numerical operations. The parts are arranged in order of decreasing difficulty so that subjects will not lose interest. Except for Part 1 (the most difficult level), all of the MZ correlations were greater than the corresponding results for DZ twins. Four different heritability ratios were also computed. The researchers concluded that a genetic component was strongly influencing numerical ability.

Stafford (1972) administered a test of quantitative reasoning to MZ and DZ twins, ages 12 to 18. The test consisted of arithmetic problems such as, "How many pencils can you buy for 50 cents if they are 2 for 5 cents?" Multiple-choice answers were provided. Intraclass correlations were computed for each sex and zygosity group. For the females, $\underline{r}_{mz} = .66$ and $\underline{r}_{dz} = .51$ and for the males, $\underline{r}_{mz} = .81$ and $\underline{r}_{dz} = .52$. The computed $\underline{F}$ for females was nonsignificant whereas for males this value was significant. Next, dichotomic analysis was applied to determine if there was evidence for an underlying bimodality in the continuous distribution of scores. The plotted histogram suggested bimodality. Upon further analysis, Stafford hypothesized that having a proficiency in quantitative reasoning was due to the presence of a sex-linked recessive trait.

Judgment of Number. Another quantitative ability concerned with numbers is the ability to judge visual number. Family members were asked to name the number of marbles or ping pong balls shown to them in each of five transparent bags. The data provided evidence that the absolute judgment of numerosity is composed of at least two kinds of cognitive processes: subtizing is the evaluation of 6 or fewer objects and estimation is for 7 or more objects. Subtizing was found to have high heritabilities whereas the heritabilities for estimation were low. The heritability estimates were based on parent-offspring correlations. Father-son, mother-son, father-daughter, and mother-daughter correlations were virtually equivalent. It was concluded that the ability to judge sets of less than 7 objects may be inherited while estimating sets of 7 or more objects showed low to no genetic influence.

Linguistic Skills. Munsinger and Douglass (1976) studied the syntactic abilities of identical twins, fraternal twins, and their siblings. Two different language measures were used: the Assessment of Children's Language Comprehension (ACLC) and the Northwestern Syntax Screening Test (NSST). To unconfound language ability and general intelligence, nonverbal IQ variance was partialed out before the language scores were analyzed. Since the subscales of the ACLC and NSST were highly intercorrelated, the subtests were pooled into one single score of syntactic ability. In order to compare twin-pair resemblances, intraclass correlations were calculated for all three groups. The identical twin pairs were very similar ($\underline{r}_{mz}$ = .831) whereas the fraternal twin pairs were much less similar ($\underline{r}_{dz}$ = .436). Sibling pairs closely resembled the fraternal twins ($\underline{r}_{s}$ = .492). Heritabilty of language skills was computed using Falconer's (1960) method. It was concluded that the heritabilty of language skills as measured by these tests was .79 in this sample.

Mental Ability Profiles. A profile analysis allows a test of the hypothesis that the relative strengths and weaknesses of mental ability are genetically influenced. Segal (1985) presented a profile analysis of the WISC-R for a sample of MZ and DZ twin children. MZ pairs showed significantly greater concordance for subtest profile than DZ pairs. More specifically, the intraclass correlation for 67 MZ pairs was .45 whereas the corresponding correlation for 34 DZ pairs was .24. These results are very similar to the Wechsler Preschool and Primary Scale of Intelligence pattern intraclass correlations reported by Wilson (1975). At age six, the correlation for 70 pairs of MZ twins was .43. The correlation for 46 pairs of DZ twins was .27.

These subtest correlations raise the possibility that the specific cognitive abilities have a differential underlying heritability. However, some twin studies (Foch & Plomin, 1980; Plomin & Vandenberg, 1980), using other measures of intelligence have failed to provide evidence of differential hertiability for various cognitive skills. The previously described studies of familial resemblance were also inconclusive with respect to this issue. At present, this issue is not resolved.

Frequency of Occurrence

The extent of familial resemblance and twin similarity with respect to particular cognitive abilities has been reviewed. Two additional specific cognitive skills are also examined. The first specific cognitive skill to be described involves estimating the frequency of occurrence of some particular event.

There is a great deal of empirical support for the assertion that adults are extremely sensitive to frequency of occurrence. Several different methodologies have been employed and this ability has been demonstrated under numerous conditions. After presenting this research, the issue of automaticity will be addressed. It has been proposed that the encoding of frequency information is an automatic process. Several criteria for automaticity have been introduced. These characteristics will be described and the empirical support for this hypothesis will be presented as well.

Leicht (1968) noted that the presentation of a word results in the occurrence of an implicit associational response (IAR). The IAR is similar to a response to a word in a word-association task. Accordingly, the occurrence of a word as an IAR is one way in which a word's frequency may be increased. It was hypothesized that the judged frequency of a word should increase if the associative frequency of a word is increased. In this research, list words were presented 0, 1, 3, or 5 times and were associates of 0, 1, 3, or 5 other list words. In one experiment, frequency rating was measured. Subjects were instructed to remember as many words as possible. Results confirmed the hypothesis. Judged frequency of a word increased with the number of times the word was an associate of other list words. In addition, judged frequency increased as presentation frequency increased.

Hintzman (1969) conducted three experiments. He was primarily concerned with individual's ability to judge the frequency of words as well as with investigating the spacing effect (the general superiority of spaced versus massed repetitions on item recall). In Experiment I, words were repeated 0, 1, 2, 4, 6, and 10 times. Subjects were not told that they would be required to judge the frequencies of occurrence. Two types of tests were administered: a paired-comparison test in which subjects were instructed to select the more frequent of the two alternatives and an absolute judgment test in which subjects gave numerical judgments of occurrence. In this experiment, apparent frequency in both the discrimination and the judgment tasks increased with the logarithm of the actual frequency (Fechner's Law). In addition, there appeared to be a slight primacy effect; judgments for the words that appeared early in the list were slightly higher than for words that appeared later in the list.

Experiment II focused on the spacing effect as well as on the primacy effect. The number of items intervening between two repetitions were 0, 1, 2, 4, 8, or 16 items. Subjects participated in two tasks. In order to investigate the influence of spacing, subjects had to indicate which word was more frequent when in actuality the words appeared equally often but differed with respect to spacing. To study the primacy/recency issue, subjects chose between words of equal frequency; however, which appeared in different positions in the presentation list.

The purpose of Experiment III was to study the spacing effect using the absolute judgment task. Both the discrimination task in Experiment II and the absolute judgment task in Experiment III showed effects of spacing. Apparent frequency increased as spacing increased. Both experiments also found a slight primacy effect.

In research by Jacoby (1972), subjects were presented simple sentences; each sentence contained an adjective, a noun, and a verb. Frequency of presentation was varied. In addition, modifiers (adjectives and verbs) accompanying presentations of a given noun were either identical or different for each repetition. Thus, the critical nouns were either repeated in intact sentences or by recombining words from sentences for repetitions in new sentences. Subjects were led to believe that they would participate in a memory test after the study trial but they did not know the nature of this test. After the study trial, subjects were instructed to indicate the number of times particular sentences had occurred. Analyses indicate that increased frequency of presentation resulted in an increase in frequency judgments only when nouns were repeated with identical modifiers. Repeating words in new sentences did not increase apparent frequency of the original.

Hintzman and Stern (1978) also investigated the effect of contextual variability on memory for frequency. Two experiments were conducted. The stimuli in Experiment I were nouns, and in Experiment II the stimuli were statements about famous people. The stimuli in both experiments either varied from presentation to presentation, or they remained the same. Subjects made frequency judgments and performed on recognition and recall tasks as well.

Results in both experiments were consistent. The experiments provide evidence that recall, recognition, and judged frequency are not always affected in the same ways by the same manipulations. Specifically, judged frequency was higher under low variability than under high variability conditions. The effect of variability on free recall and recognition memory was just the opposite; high variability led to better recall and recognition than did low variability.

The studies also add support to the notion that the encoding of frequency information is obligatory, rather than optional. In essence, showing that frequency judgments are not enhanced by the availability of redundant information in variable contexts argues strongly that the processing of the information underlying frequency judgments is obligatory.

Gude and Zechmeister (1975) presented a series of sentences to college students. Critical sets of sentences were repeated in identical or paraphrased form. One group of subjects was asked to judge the number of times that a sentence had occurred in the precise wording (group W) that appeared on a test sheet. Another group was asked to judge the frequency with which the basic meaning (group M) of a sentence had occurred. Also, in an effort to study the spacing effect, either 0, 4, or 8 sentences appeared between repetitions.

Analyses of the results indicated that subjects can accurately judge the frequency of repeating identical sentences as well as the frequency of repeating altered sentences having the same underlying meaning. Sentences repeated in identical form resulted in a spacing effect when subjects were asked to judge their frequency. However, this typical spacing effect was not found in the frequency judgments of paraphrased sentences.

Burnett and Stevenson (1979) conducted a similar study and obtained additional evidence that individuals are very accurate in judging the frequencies of both the exact wording and the gist of the sentences. Subjects were instructed to attend to either the gist or the exact wording of the sentences. Half of each group was then given either gist or exact-wording retrieval instructions, creating four experimental groups: (a) gist-gist, (b) gist-wording, (c) wording-gist, and (d) wording-wording. It was found that retrieval of frequency for the exact wording of a sentence and the gist of a sentence were very accurate regardless of the orienting instructions. This appears to strengthen the argument that frequency is processed automatically.

Alba, Chromiak, Hasher, and Attig (1980) have demonstrated that adults can estimate the frequency of implicitly referenced events, such as category names, when they were presented with category instances. Subjects saw a list of categorized words. Each critical category was represented by 0, 3, 6, or 9 exemplars. After they had seen the list, subjects were informed that, when cued by the category name, they were to judge the number of exemplars that they had seen. Four different pacing rates were used. Along with determining whether or not adults store occurrence rates for category names when they only see exemplars, this study was concerned with the issue of when the frequency information is encoded. It was reasoned that if this information was stored upon presentation, then the test pacing should have little effect on the judgments. On the other hand, if this information is inferred at the test then the rate of testing should influence the accuracy of those judgments.

Analyses indicated that judgments of the number of instances of a category increased as the number of exemplars increased. Additional analyses indicated that the information needed to make the judgment was stored incidentally during stimulus presentation since subjects given only two seconds to read the label and produce a response performed no differently from subjects given 10 seconds to perform the same task.

A second experiment was also conducted. This experiment compared the performance of subjects who either received intentional frequency instructions or incidental instructions. Equivalence in performance led the authors to conclude that superordinate frequency is automatically encoded.

How does one know that something is encoded automatically? Hasher and Zacks (1979, 1984) proposed a set of criteria which must be satisfied for an experience to be considered one which is automatically encoded. Firstly, individuals are sensitive to frequency information without necessarily intending to be. In addition, intention doesn't improve frequency performance. Training and practice at processing frequency information doesn't improve encoding and neither does explicit feedback. Individual differences in the ability to encode frequency information are minimal. Such variables as age and education do not significantly influence performance. Finally, it will be noted that disruptions from arousal or from simultaneous processing demands will have virtually no affect on the processing of frequency information.

There is substantial evidence on the automaticity of encoding frequency information. Howell (1973) presented instructions to subjects stressing either recall or frequency. Half of each group were then instructed to participate in either a free recall task or a frequency task. Results indicated that instructions emphasizing frequency produced consistently inferior free recall when compared to memory instructions. However, both sets of instructions produced essentially the same frequency responses. This experiment provides further evidence that people can process frequency information whether they are instructed to do so or not.

Hasher and Chromiak (1977) set out to demonstrate the absence of any developmental trend in the ability to discriminate frequency. The subjects in this experiment were students in the second, fourth, and sixth grade as well as college students. Half of the subjects at each of the grade levels were instructed prior to list presentation that their task would be to judge the frequency of occurrence of the words in the list. The other half of the subjects received nonspecific memory instructions prior to presentation of the list.

A second experiment questioned whether or not frequency discrimination improved with practice. In this study all subjects were instructed prior to the presentation of the list that they would be asked to judge the frequency with which items occurred. After presentation of the first list, subjects had two minutes to make their judgments. Half of the subjects were then given one minute to look over a feedback sheet. This sheet contained a list of items and recorded alongside of each item was its actual frequency of occurrence. They were told that this feedback might help them to make judgments on the next list. Subjects in the nonfeedback condition spent this one minute interval waiting for the experimenter to get ready. Evidence from these studies led the researchers to conclude that the ability to encode frequency information doesn't show a developmental trend across the range of ages from second grade to college. In addition, this ability is not affected by explicit instructions about the task. Finally, specific feedback on a prior list does not significantly influence the accuracy of judgments on a subsequent list.

As previously mentioned, automatic processes can occur without the intention of the information processor and they require minimal capacity. In contrast, nonautomatic processes require both attention and effort. It has been proposed that although old subjects should do worse on tasks requiring effortful processing, they should do just as well as younger subjects on tasks requiring automatic processing. In Attig and Hasher (1980), three groups of adults (mean ages 22, 43, and 68 years) listened to a list of study words. The critical words occurred from zero to seven times each. Half of the subjects at each age level were fully informed about the nature of the task prior to the presentations of the study list. The other half of the subjects were only generally informed about the task before presentation of the study words. All subjects were then specifically informed about the frequency task just before presentation of the test list. Subjects were tested on a two-alternative forced choice procedure. They were instructed to circle the word in each pair that had appeared more frequently.

Results indicate that the instructional manipulation had no significant impact on performance. In addition, analysis of variance confirmed that age did not influence performance. In summary, these data support the argument that some memory processes are unaffected by age as well as the intention of the information processor.

Kausler and Puckett (1980) replicated the results of Attig and Hasher (1980). Both studies found that age had an insignificant effect on subjects' performance on a relative frequency judgment task. In addition, this study extends the generalizability of these findings by introducing several procedural variations. The first procedural change involved visual, rather than auditory, presentation of the study list. Secondly, in Attig and Hasher's study, subjects either received the list under incidental learning conditions or under intentional learning conditions. In Kausler and Puckett's study, a within-subjects design was used with subjects performing under both incidental and intentional instructions. The third dimension involved the number of test trials. Attig and Hasher administered only one test trial whereas in this study a test trial was administered again, but with different pairings.

Two kinds of relationships were investigated correlationally in this study as well. The first was between frequency judgment proficiency and proficiency on an effortful learning memory task. After subjects performed on the incidental and intentional frequency judgment tasks, they participated in a paired-associate learning task.

The second kind of relationship was between frequency judgment scores and scores on both crystallized and fluid intelligence tests. The psychometric tasks consisted of the WAIS Vocabulary Test as a measure of crystallized intelligence and the

Figural Relations Test of Horn's Gf-Gc Sampler as a measure of fluid intelligence. The order of these two tests were counterbalanced.

Analysis of variance revealed that neither the main effect for age nor the main effect for instructions approached significance. The age by instructions interaction did not approach significance either. Thus, Attig and Hasher's results were replicated despite the procedural changes introduced in this study. The fact that intentionality did not improve performance lends itself as support for the automaticity of these processes. Since elderly adults did not differ from the younger adults, there is evidence for the absence of age changes in these automatic processes. The main effect for trials was significant, but neither the age by trials nor the instructions by trials interaction approached significance. As expected, judgment scores were lower on Trial 2 than on Trial 1. This adverse effect of prior testing was seen for both age groups and both instructional conditions.

Although no age effect was found in these frequency tasks, performance on the paired-associate task revealed a significant age difference favoring the young adults. Of even greater interest was the covariation between frequency judgment and paired-associate performance. The correlations between these two scores for both age groups were not significantly different. It seemed as though effortful processes involved in the processing of frequency information are no more involved in young than in elderly adults.

A significant age difference, biased toward the elderly group, was found on the vocabulary test, whereas, on the figural relations test, the young subjects were significantly favored. In general, measures of intelligence were positively correlated with frequency scores for young but not for elderly adults.

As well as comparing the performance of elderly and younger subjects, some researchers have been interested in comparing depressed subjects and nondepressed subjects. Hasher and Zacks (1979) reported that the performance of the depressed and nondepressed adults were virtually identical. Estimates of the frequency of occurrence increased with actual frequency of the stimuli. The two groups did not differ significantly.

Numerous empirical standards have been presented to establish that a particular encoding operation is automatic. Recently, Zacks, Hasher, and Sanft (1982) conducted a study in order to evaluate several of these criteria. One experiment compared practice effects on a single-trial free-recall task with those on a forced-choice frequency task. As predicted, the results for the free-recall task contrasted with those for the frequency task. Recall performance showed a positive practice effect, a benefit from appropriate practice as well as reliable individual differences. In contrast, frequency performance did not show a positive practice effect, a differential effect of appropriate versus inappropriate training, or consistent individual differences. Additional experiments demonstrated that free-recall was affected by the accuracy of test expectations and was hindered by competing demands. In contrast, frequency performance was not influenced by accuracy of test expectations and was not hindered by competing demands. In addition, individual differences were examined by comparing the performances of students from two very different universities: University A was a small, highly select, private institution, and University B was a state university with more liberal admission standards. Differences in SAT scores favored students at University A. Such differences were expected to influence recall performance but not frequency performance. The evidence supported these predictions. In total, this evidence provided further support for the hypothesis that frequency information is processed automatically.

Sequential Pattern Learning

The second specific cognitive ability which will be investigated is known as sequential pattern learning. Several theories have been proposed to account for an individual's ability to learn, to predict, and to generate patterned strings of events. Early theories of sequential pattern learning were based on simple conditioning and associations. These descriptions of the coding process were found to be too simplistic. Theorists proceeded to develop more complex models involving hierarchies, hypothesis-testing and higher-order cognitive processes. Each of these theories as well as the corresponding empirical findings will be described.

One of the earliest theories was Estes' statistical learning theory (Myers, 1970, 1976). According to Estes, event prediction is based on an underlying representation of the probability of the event occurring. The probability develops through a simple conditioning process. In the linear model, some proportion of the stimuli are randomly sampled on a given trial. The probability of a particular response "A" is equal to the proportion of sampled stimuli conditioned to this response. When the reinforcement event occurs, the sampled stimuli become conditioned to this response. In the pattern model, a pattern or small set of stimuli are sampled on a given trial. The response is based on the response conditioned to this pattern. Upon reinforcement, the entire pattern becomes conditioned to this response.

Both models predict: (a) positive recency and (b) probability matching. Positive recency proposes that, as a certain reinforced response continues to occur, the probability of that response should approach 1.00. Probability matching proposes that the probability of event prediction should match or equal the actual probability or frequency of that event. Empirical evidence does not support these notions. The data tend to demonstrate negative recency as well as overshooting. Negative recency implies that as an event continues to occur two or three times in succession, the probability of predicting the alternative event increases. Overshooting implies that, with extended practice, the probability of predicting a certain event exceeds its actual probability of occurring.

Gambino and Myers (1967) proposed a generalization gradient. The basic principle implies that the continuation (or breaking off) of a run of a given length is influenced by the continuation (or breaking off) of a run of similar length through a generalization mechanism. Subjects have an expectancy concerning continuation (or breaking off) for each run length. If a subject observes a run of length $\underline{m}$ continue (or break off) his expectancy will increase (or decrease) for a run of similar length. The effect is strongest close to $\underline{m}$ and decreases with further distance from $\underline{m}$.

Restle (1966) proposed that subjects should learn perfectly. In other words, if a run always continues, subjects would always predict it would continue. If a run always broke off, subjects should always predict it would break off. However, after several hundred trials, Restle's subjects still made errors. Gambino and Myers' generalization model provides an explanation for this imperfect learning. The following situation illustrates this point. If runs of length $\underline{m}$ + 2 always broke off, Restle would say that subjects should always break off these runs. The generalization model allows for errors. Subjects may predict continuation with runs of length $\underline{m} + 2$ since runs of length $\underline{m} + 1$ always continued. Thus, performance on $\underline{m} + 2$ runs is influenced by $\underline{m} + 1$ runs.

According to Vitz and Todd (1967), a pattern is any subset of events which, when repeated, can generate the entire string or sequence. A pattern of binary events indicates that only two possible events can occur. A simple pattern of binary events indicates that only one possible event can occur following a particular event or run. Vitz and Todd explained the learning of simple repeating binary sequences through a series of S-R connections. To illustrate, in order to learn the sequence "aaabb" the following associations would be made:

(1) a ---> a,
(2) aa ---> a,
(3) aaa ---> b,
(4) b ---> b, and
(5) bb ---> a.

This is a simple pattern since only one possible response can follow any particular stimulus. This model predicts that all positions within a sequence will be learned equally well and that learning will be independent of sequence length. In addition, the model also predicts that learning will be an all-or-none process. Empirical evidence shows no increase in performance during the presolution phase of learning.

Restle has studied sequential pattern learning. In his first model (Restle, 1961), the _K_-span model, he proposed that the subject has in memory the last _K_ events. Response on a given trial is based on sampling from a set of stimuli elements of a corresponding set of _K_ events. Restle became dissatisfied with this model (1966). It appeared that learning seemed dependent on certain runs of variable length rather than on events of a fixed length. In addition, it did not allow for any encoding processes.

He then proposed the schema (run) model. The run became the unit of encoding. He proposed that subjects have in memory the last event that occurred as well as the number of times it has occurred in succession. Based on this, he predicted that the probabliity of predicting that an event will occur (continue) should match the actual frequency with which it occurs (continues). Since data indicated overshooting, Restle then proposed that memory for longer lengths was more salient than memory for shorter length runs. Thus, event prediction was now based on a combination of the actual probability (frequency) of the event occurring (continuing) as well as consideration for the run length.

As previously mentioned, Vitz and Todd developed a theory of simple repeating binary sequence patterns. Restle's two-stage model explained the learning of more complex repeating binary sequence patterns. He proposed that subjects could learn both a set of mandatory rules as well as a set of optional rules. A mandatory rule indicates that each rule determines a single continuation whereas an optional rule permits several different possible continuations. The following situations will help illustrate the differences between mandatory and optional rules. First, in order to learn a repeating sequence "1000" the subject need only learn the following set of mandatory rules:

(1) 1 ---> 0,
(2) 0 ---> 0,
(3) 00 ---> 0, and
(4) 000 ---> 1.

However, in order to learn the sequence "1001", the subject need learn some mandatory as well as optional rules:

(1) 1 ---> 0,
(2) 0 ---> 0,

(3) 00 ---> 1, and

(4) 1 ---> 1.

The fact that 1 ---> 0 or 1 ---> 1 indicates the necessity for optional rules. Restle proposed that subjects will learn the required set of mandatory rules faster than the set of optional rules. His experiments support this notion.

Additional empirical evidence comes from a study conducted by Garner and Gottwald (1967). Subjects learned sequences of binary events. Simple patterns such as "RRRLL" and complex patterns such as "RRLRL" were presented. In essence, the simple patterns consisted of only mandatory rules whereas the complex patterns contained both mandatory and optional rules.

Results indicated that the simple patterns were much easier to learn than the complex patterns. In addition, pattern descriptions clearly showed that subjects perceived the simple patterns as organized into two runs; descriptions of the complex patterns were more variable. Thus, Restle's two-stage model is one of the first times we see any mention that some positions in a sequence will be more difficult to learn than other positions.

Restle then began to propose that subjects divide the sequence into subunits of runs and trills (Restle, 1970; Restle and Brown, 1970a, 1970b). An example of a run is "AAAA" and an example of a trill is "ABABAB". He developed his E-I theory and proposed that subjects represent the information in sequences in this fashion. E refers to the set of elements or events that occur in the sequence and I refers to the set of intervals between the elements. Thus for the sequence "12323", E = (1,2,3) and I = (+1,-1). The subject would only need to remember E and I in order to represent any sequence.

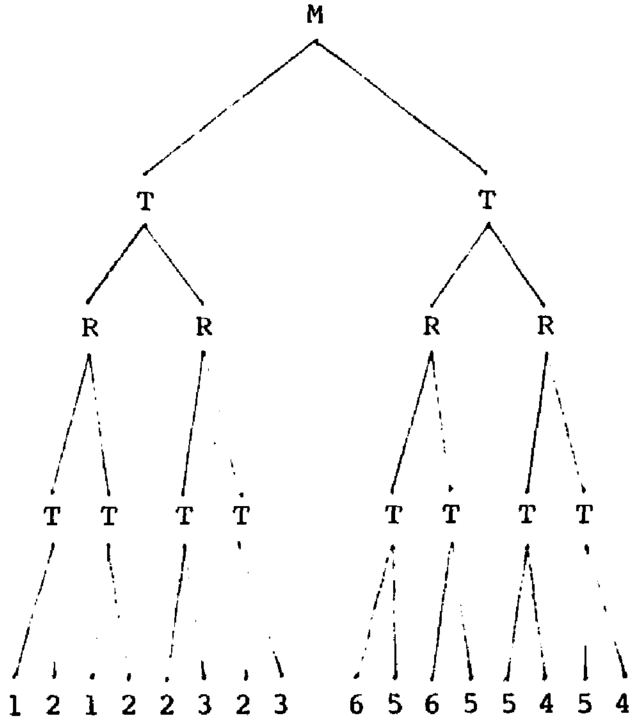
There is a great deal of empirical evidence indicating that subjects divide the sequence into runs and trills. Although these are simple ideas, they are abstract ideas about the encoding of sequences. In one study (Restle and Brown, 1970c), one group of subjects was pretrained on "run" sequences. A second group was pretrained on "trill" sequences. During a test trial, both groups were provided with an ambiguous test sequence. This sequence can be divided according to runs or trills. The hypothesis was that performance on the test trial would be consistent with performance during pretraining. The data supported this notion. Subjects pretrained on "runs" tended to divide the test sequence into "runs" while subjects pretrained on "trills" tended to divide the test trial into "trills." In another study (Restle and Brown, 1970a), pauses or rest periods occurred during the presentation of a sequence. For one group the pauses divided the sequence into runs and for another group the pauses divided the sequence into trills. An ambiguous test trial was presented. Data indicated that the positioning of the pauses had a significant effect on the encoding of the pattern. Subjects in the first group tended to divide the sequence into runs while subjects in the second group tended to divide the sequence into trills.

Restle's most elaborate model was an hierarchical rule learning model known as structural trees (Restle, 1970; Restle and Brown, 1970a). An important component of the hierarchical model is that each element can generate many other elements. This hierarchical network can allow for much greater flexibility and complexity of sequences. It is also necessary that operations be able to be performed on the product of other operations. Restle proposed a number of different operations: transposition, (T), repetition, (R), and mirror-image, (M). Given sequence X = (1,2,3) and the possible elements of (1,2,3,4,5,6), T (X) = 123234,

R (X) = 123123, and

M (X) = 123654.

As previously indicated, operations can be performed on products of previous operations. Thus, M (T(X)) = 123234654543. According to this theory each sequence of operations or sequence of events can be represented in a structural tree. M(T(R(T(X)))) with X = 1 would generate the sequence 1212232365655454. This sequence may be represented by a structural tree.



The notion of structural trees allows for elaborate representation as well as structural analysis. Based on this notion of representation of sequences, several predictions can be made:

(a) subjects will be able to learn lower-order operations faster than higher-order operations;

(b) processing will tend to be bottom-up rather than top-down;

(c) transfer for top-down will be less (worse) than transfer for bottom-up processing;

(d) subjects should tend to make their errors between branches rather than within branches (predictions about which positions in sequence will be more difficult to learn); and

(e) when there is a deviation in the sequence, subjects will tend to fill-in this position with the "appropriate" sequence (overgeneralize). There will be more fill-ins if the deviation occurs late in the sequence than if the deviation occurs early in the sequence.

Restle provided empirical evidence for each of these predictions.

Restle was not alone in investigating transfer processes. Keeney (1969) devised a paired-associate learning task. The stimuli were strings of five consonants followed by a numeral; the responses were reorderings of the same five consonants. The stimulus and response strings were related to one another by one of four permutation rules. The four rules together divided the consonant string into a hierarchical system of units. Subjects participated in an initial learning phase and then a transfer task. It was hypothesized that if learning in the initial phase only consisted of the formation of associations between the stimuli and the responses, then the presentation of five entirely new consonants should completely disrupt performance. However, if during the initial learning phase, the permutation rules were learned, then changing the letters should have a minimal effect on performance. It was found that, when transferred to a list with five entirely new consonants, but with the same rules, subjects performed virtually errorlessly. As well as providing data on transfer processes, this research also supported the notion of a hierarchical organization. When subjects responded, they often tended to pause in a pattern appropriate to the hierarchical units imposed by the permutation rules.

As previously mentioned, Restle made several predictions based on his notion of structural trees. Other researchers have also obtained empirical evidence supporting these expectations. Marmurek and Johnson (1978) had subjects learn a set of permutations of a base sequence of letters. A set of permutations either defined an hierarchical organization for the base sequence or did not. It was found that sets that defined hierarchical organizations led to more correct responses. It was also found that the transitional error probability (TEP) was higher between subgroupings than within subgroupings.

Finally, Hersh (1974) systematically investigated the effects of irrelevant relations on letter series completion problems. These irrelevant relations were positioned at the beginning, at the end, or they did not appear at all. Analysis revealed that irrelevant relations at the beginning of the series produced the longest latencies and the most errors. Perhaps these irrelevant relations increased the number of hypotheses that subjects had to test. This situation may resemble Restle's issue of "filling-in" when a deviation occurred. According to Restle, a deviation was most likely to be filled-in when it occurred late in the sequence. Apparently the subject has learned a pattern and will overgeneralize it when a deviation occurs here. Similarly, when an irrelevant relation was introduced towards the end of a sequence, Hersh found that it did not generate many hypotheses. It appears that exposure to a major portion of the sequence allows subjects not only to learn the pattern but to remain consistent in their performance as well.

Although Jones has not devloped a major theoretical orientation, she has found evidence that cannot be accounted for by previous theories. In one experiment, Jones (1973) manipulated the type of higher order operation as well as the positioning of this change in operator. The two operators were mirror-image (M) and transposition (T) and the positions were first (F) or halfway (H). Generally, it was hypothesized that (M) should be more difficult than (T) and (H) should be more difficult than (F). However, Jones found an interaction. Whereas (TF) was easier than (TH), (MH) was easier than (MF).

In another experiment (Jones and Zamostny, 1975), one group of subjects was trained according to a linear strategy and another group was trained according to an hierarchical strategy. Subjects trained linearly performed better on the first half of a sequence whereas subjects trained with the hierarchical model performed better on the second half. Thus, type of training interacted with performance on each half of the sequence. These are two illustrations of interactions that cannot be accounted for by other theories. For instance, Restle cannot explain why (M) would be more difficult than (T).

Simon and Kotovsky (1963; Kotovsky & Simon, 1973) developed a theory that predicts the relative difficulty of letter series. They began by considering what the subject brings to the task. Simon and Kotovsky (1963) assumed that people know three things about letters in the alphabet. They know when two letters are identical, they know which letter follows any letter in the alphabet, and they know which letter precedes any letter. These are known as the identical, next, and backward operations, respectively. According to Simon and Kotovsky, this information is used during series representation as well as when people continue a letter series.

During representation, people use the information to induce which letters go on a string. Simon and Kotovsky proposed that once people have assigned letters to a string, they then use the information on that string to calculate the problem's period. People then use the knowledge of the period length along with the represented string to guide their induction of which letters go on other strings. Once people have assigned all letters to a string, people supposedly form a rule describing the entire series. For Simon and Kotovsky, this rule consists of memory lists. A memory list is a psychological unit referring to a string that moves up or down the alphabet from one period to the next. For instance, for the problem ATUBSUCRU, Simon and Kotovsky would say that a person forms two memory lists, one for the next operation on the first item and the other for the backward operation on the second item. According to Simon and Kotovsky, people don't form a memory list to represent an identity string.

For Simon and Kotovsky, series continuation involves applying the same operations they use to induce its representation. This is guided by the rule formed during representation. They proposed that the primary factor of difficulty during continuation would be the number of memory lists that must be employed. Their research supports this notion that the number of memory lists predicts the difficulty of letter series problems. In addition, Holzman, Pellegrino, and Glaser (1982, 1983) recently reached the same conclusion in their studies of the cognitive determinants of number analogy performance and of number series completion performance.

Simon and Kotovsky, apparently influenced by a computer analogy, proceeded to develop their theory consisting of a formal notational language and a computer program. The basic components of Simon and Kotovsky's model include a pattern generator and a sequence generator. A pattern generator is programmed to abstract a pattern description from a sequence that has been presented. It searches for symbols recurring at regular intervals. If this isn't detected, it searches for regularity that is interrupted at periodic intervals. Once a pattern description has been obtained, a sequence generator uses this description to elicit and generate a sequence. The sequence generator is programmed to write symbols, copy symbols, etc., and can thus produce a sequence. Simon and Kotovsky compared the performance of their program with the performance of human subjects. For human subjects, there was a high correlation between the number of words in the pattern description and the difficulty in generating the sequence. Simon and Kotovsky also tended to find that memory load, the number of memory lists, was related to difficulty in generating the sequence.

Greeno and Simon then proceeded to propose three possible interpreters. Each of these interpreters uses the pattern generator to generate sequences but with different emphases on, memory. The first interpreter to be described was the push-down interpreter. Here, each element in the sequence can be generated with one single operation. The principle of "last one in, first one out" applies here. In order to reach some goal, several subgoals may need to be achieved. To illustrate, given the sequence "523," the push-down interpreter would represent this as seen in Figure 1a. In order to get to the number "3" we must first meet the subgoals of getting to number "2." If we start with $\underline{t} = 5$ and perform the operation $(\underline{t} - 3)$, then number "2" will be obtained. Now, if the operation $(\underline{t} + 1)$ is performed, the number "3" will be obtained. Thus number "3" was obtained by performing a single operation on a previous element in the sequence. This interpreter has a moderate load on both short term memory (numbers) and long term memory (operations).

The second interpreter was the recompute interpreter. According to this interpreter, each element in the sequence can be generated from the original element. For instance, the sequence "5234" would be represented as seen in Figure 1b. The

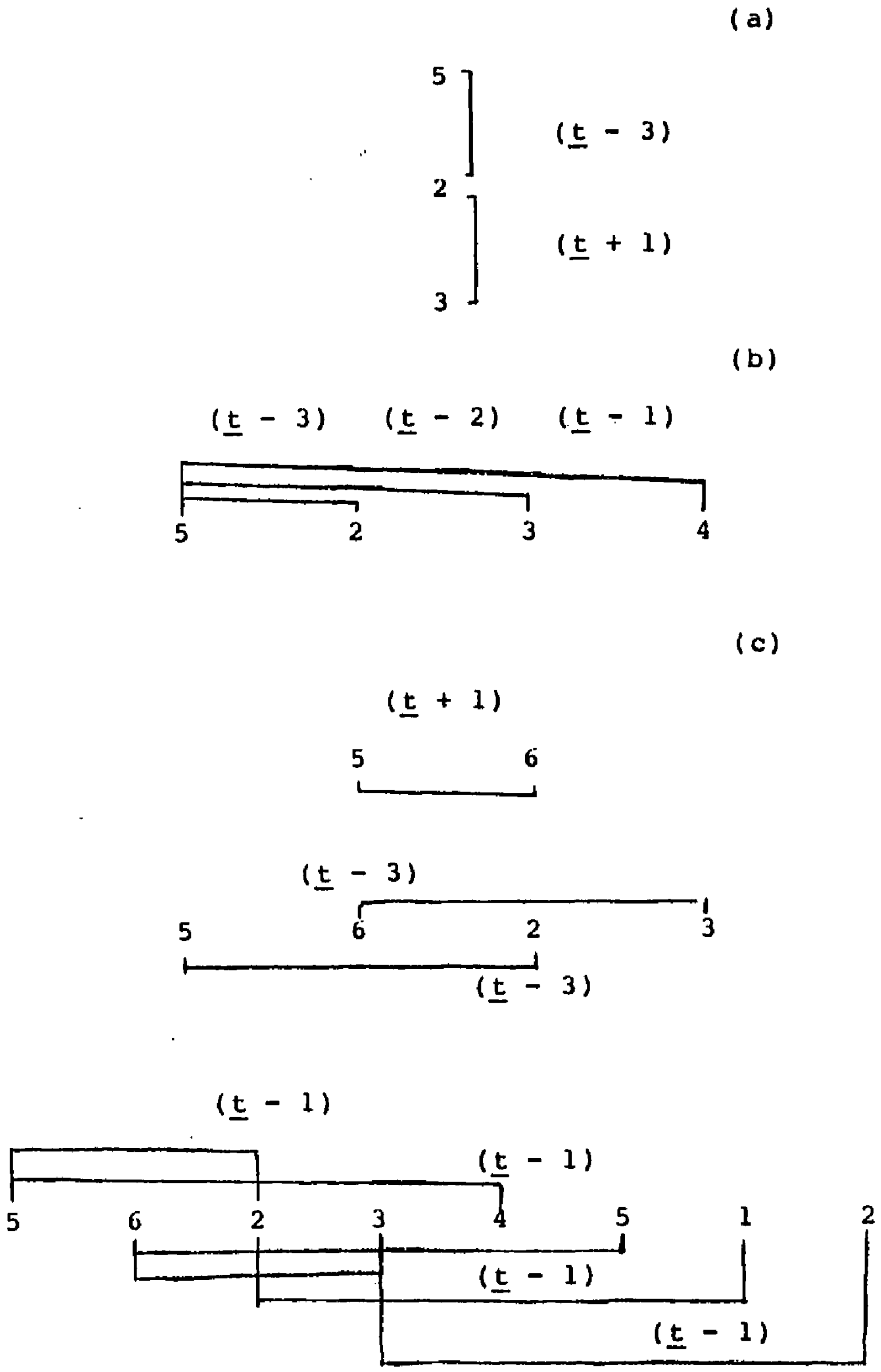


Figure 1. The Three Possible Interpreters Proposed by Greeno and Simon.

recompute interpreter places a heavy burden on long term memory since there are many operations to recall but a low burden on short term memory since one only needs to recall the original element in the sequence.

The third interpreter was the doubling interpreter. Here, as with the push-down interpreter, each element in the sequence can be generated by using one operation on a previous element, but each operation can be applied several times in succession. For instance, Figure 1c is a representation of the sequence "56234512." Starting with the number "5," the operation of (t + 1) could be applied. This would yield number "6." The operation (t - 3) can now be applied to each element. This extends the string by adding the elements "2" and "3" to it. Finally, if the operation (t - 1) is applied to each existing element, the numbers "4," "5," "1," and "2" are obtained. Thus, it can be seen that each time a new operation is performed on the existing elements, the length of the sequence doubles. The interpreting places a heavy load on short term memory (must remember elements) but a light load on long term memory (few operations to recall since they can be applied several times in a row). It is important to note that each of these three interpreters are possible and no one interpreter seems to dominate in these tasks.

Butterfield, Nielson, Tangen, and Richardson (1985) have recently extended the work of Simon and Kotovsky and by Holzman, Pellegrino, and Glaser. To begin, they developed a rule system to notate and to generate letter series. *N* means next one letter up the alphabet, *I* means identical, and *B* means back one letter down the alphabet. To illustrate, the rule *N*1 *I*1 *B*2 *I*2 describes the series DDYYEEXXFFWW. The numbers, which can be viewed as subscripts, indicate that the rule describes a series composed of two strings, the first of which has an *N* and an *I* relation and the second of which has a *B* and an *I* relation. The *N*1 *I*1 string was created by selecting a C and applying first an *N* operation, giving the letter D, and then an *I* operation, giving another D, then an *N*, giving the letter E, then an *I*, and so on recursively. The *B*2 *I*2 string was created by starting with a Z and recursively applying back and identity operations. Finally, the two strings were intermixed in the order specified by the rule. The problem has a period of 4 which is the sum of the periods of the two strings from which it is constructed.

As previously mentioned, Simon and Kotovsky would say a person forms two memory strings, whereas I1 and I2 are nonmoving or identity strings. Butterfield et al. have also attempted to incorporate letter series with spurious relations into their theory of letter series continuation. This will become evident shortly but for now it is necessary to define what is meant by spurious relations. A letter series contains spurious relations if it contains duplicate letters that are not related to one another by identity relations that lie on the same string.

Butterfield et al. have developed a knowledge hierarchy which hypothesizes how people represent letter series problems. The easiest level consists of series composed entirely of nonmoving strings. Strings composed entirely of identical letters are easy to represent. The next four levels consist of series with moving strings, but without spurious relations. They proposed that the representation of moving strings is influenced by two orthogonal features, the adjacency or nonadjacency of groups of identical letters on a string and the equality or inequality in number of identical letters among the groups. The easiest moving string to represent would be composed of units that have the same number of adjacent identical letters. Likewise, the hardest moving string to represent would be composed of units with different numbers of nonadjacent identical letters. In addition, they hypothesized that nonadjacency of letters within units would be more difficult to represent than inequality among units in number of letters. This established the following predicted rank order of representational ease: adjacent and equal; adjacent and unequal; nonadjacent and equal; and nonadjacent and unequal. Finally, because it was proposed that more abstract knowledge is required to represent series with spurious relations, this knowledge was placed at the sixth level of the knowledge hierarchy.

Since this theory is presented as a hierarchy it assumes that people who understand knowledge at any particular level also understand the knowledge at lower levels. As well as predicting the relative difficulty of letter series with different features, the hierarchy predicts the order in which representational knowledge is acquired and the order in which it is used by an individual presented with any series problem. Essentially, an individual will use the knowledge from the bottom of the hierarchy up whenever they are presented with a series problem.

Having hypothesized about how people represent letter series, Butterfield et al. turned their attention to the processes required to continue or extrapolate the series. It seemed to them that Simon and Kotovsky were correct in proposing that memory load would determine the difficulty of extrapolating a series. However, they hoped to achieve a more process-oriented explanation of memory load. Butterfield et al. described two processes that a person must participate in after representing a series. First, the person must locate the appropriate letters from which to generate each letter needed to continue the series. Second, having found the required letter by working backwards, a person must create the letter required for continuation by performing the relational transformation specified in his or her representation of the series' string structure.

The number of memory lists in the series rule was the index of continuation difficulty for Simon and Kotovsky. Similarly, one index for Butterfield et al. is the number of letters a person must work backward in order to find all of the letters to continue a problem for an entire period. Consider the series AHP BGP CFP. Every letter moved back is called a Count. In order to generate the C in the last period, the person must move back to the B in the preceding period. Likewise, in order to generate the F, the person must move back to the G. In this way, both C and F require counts of 3, but P requires no counts because it is located on a nonmoving string, which requires no memory operations for its continuation. Butterfield et al. indicated that it is not necessary to examine each series to determine its count value. This can be done from its rule using the following formula:

$$\text{Counts} = \#\text{MS}\ (P - \#\text{AIR})$$

where #MS equals the number of moving strings in the rule, P refers to the period length, and #AIR equals the number of adjacent identity relations. Although the formula for Counts shows that P might correlate with series difficulty, Butterfield et al. argue that it is not a determinant of series solution.

A second index of continuation difficulty proposed by Butterfield et al. has to do with the difficulty of transforming the found letters into desired letters. They weight N and B relations according to their difficulty whereas Simon and Kotovsky made no distinctions among the relations. They assumed that performing a back relation is harder than performing a next relation, and that performing an identity relation places no burden on memory. Accordingly, each I relation on each moving string in a rule is given a value of 0, each N relation on each moving string a value of 1, and each B relation a value of 2. The sums of the values of the relations on all moving strings is the relational value of the series, called Relations.

Butterfield et al. conducted a series of experiments to determine to what extent the predictors Number of Memory Lists, Knowledge, Counts, and Relations are each related to series continuation difficulty. Analyses of the data led them to draw several conclusions. To begin, all four predictors were significantly correlated with the percentage of children who responded correctly to problems of each rule (%R). All intercorrelations of the four predictors were significant, but Counts, Relations, and Number of Memory Lists correlated more strongly with one another than with Knowledge. The multiple correlations tended to be approximately .95. Number of Memory Lists failed to significantly contribute to this multiple prediction of series difficulty. In other words, the data indicated that the variance in difficulty associated with Number of Memory Lists was accounted for completely by Counts and Relations.

Based on these results, Butterfield et al. concluded that these memory processes are better thought of as locating needed letters and transforming them into desired letters than as the number of memory lists to be used.

In one experiment, they investigated the importance of period. Number of Memory Lists, Knowledge, Counts, and Relations were varied orthogonally with Period (_P_). In addition to varying _P_ of the problems, the order of the problems was also varied. Some subjects received randomly ordered problems while others received problems blocked according to _P_. It was reasoned that if calculating _P_ contributes to the difficulty of solving letter series, then the randomized order of problems should be more difficult. Under this condition, there should be uncertainty about _P_ for every problem. On the other hand, for blocked problems, _P_ should become apparent after a few problems. Five orders of items were formed; one random order and four blocked orders were arranged. Analysis of the data indicated that none of the five means differed from any other, and there were no significant order effects. Further analyses indicated that when Number of Memory Lists, Knowledge, Counts, and Relations are controlled, variation in P from three to six has no influence on difficulty. In addition, each of the other predictors taken alone has a strong influence on series difficulty.

Although their research generally supported their initial theory, additional experimentation suggested that Butterfield et al. were wrong in two aspects of their initial hypothesis about representational knowledge. First, instead of a single hierarchy, knowledge required to identify series' string structures should be represented as two dimensions. One dimension concerns the knowledge that a string's units can be formed of either adjacent or nonadjacent identical letters. The other dimension concerns the knowledge that units falling on the same string can have the same or different numbers of letters.

The second aspect of their theory that required modification concerned spurious relations. Butterfield et al. hypothesized that spurious relations are resolved by application of a more abstract level of knowledge. Based on their data analysis, they dropped the highest level of the knowledge hierarchy and added a separate predictor to account for the difficulty of series with spurious relations.

To conclude, Butterfield et al. proposed that almost perfect prediction of letter series difficulty would be provided by five variables: Counts, Relations, Adjacency, Equality, and Spurious Identities. They demonstrated that if Adjacency and Equality were substituted for Knowledge in multiple regression analysis, a greater proportion of variance in series difficulty was accounted for. When Knowledge was used in combination with Counts and Relations, very high multiple correlations resulted. Further experimentation indicated that if Spurious Identities was used instead of the sixth level of Knowledge, even a greater proportion of variance in series difficulty was predicted. Based on these findings, Butterfield et al. concluded that very near perfect prediction of series difficulty would result if these five predictors are used.

In summary, the relatively recent literature in behavior genetics of human intelligence and several areas in cognitive psychology has been reviewed. The behavior genetic studies of general intelligence indicate that the greater the amount of shared genetic material, the more similar the family members are intellectually. This same relationship was usually evidenced with respect to specific cognitive skills as well, although some skills appear to be more heritable than others.

As previously mentioned, two areas in cognitive psychology were also described. First, an abundance of data indicated that individuals can estimate the frequency of occurrence of various stimuli. The evidence supports the view that this process is virtually automatic. In addition, various theories and studies were presented about sequential pattern learning. Although individuals can often explicitly encode a series of stimuli and identify "the rule" which generates a particular string of events, no one theory can fully account for subjects' performance.

This research made several contributions to our present knowledge of these issues. First, as far as the author is aware, the methodologies of behavior genetics have been applied for the first time to these two specific cognitive skills. One goal of this research was to examine whether there is any evidence of genetic influence. If there is any such evidence, it was hypothesized that it is greater in the sequential pattern learning task than in the frequency of occurrence task.

Second, the effects of two different independent variables on sequential pattern learning performance were investigated. The two factors were: (a) the type of operation required in order to continue the sequential pattern and (b) the number of items which change within a string of the sequential pattern. It was proposed that each of these factors significantly influence sequential pattern learning performance.

Third, an abbreviated version of a general intelligence test was administered to a new sample of MZ and DZ twins. As the previous literature indicates, MZ intraclass correlations were expected to be greater than the DZ intraclass correlations.

Finally, the relationship between general intelligence and these two specific cognitive abilities was examined. Hasher and Zacks have demonstrated that frequency of occurrence performance is relatively independent of intellectual capacity. Based on this evidence, it was proposed that the correlation between performance on this frequency task and the four subtests of the WAIS-R is nonsignificant. In sharp contrast, it was hypothesized that the correlation between performance on the sequential pattern learning task and the WAIS-R is very significant.

Method

Subjects

The data for this study were obtained from 25 pairs of MZ twins and 25 pairs of same-sex DZ twins. Mothers of twins were asked to complete a questionnaire which consisted of 10 questions that are known to differentiate between monozygotic and dizygotic twins. As previously described, Cohen et al. (1973, 1975) reported that this questionnaire had 98% accuracy when cross-validated against diagnoses based on blood types. In addition, the mothers were asked to indicate the twins' birthweights and whether she believed the twins to be identical or fraternal. This questionnaire may be found in Appendix A.

Cohen et al. (1973, 1975) performed a discriminant function analysis on their questionnaire data. This analysis generated a set of discriminant function raw score coefficients. Each coefficient is a weighting that reflects the degree to which a response on a question uniquely contributes to the likelihood that a twinship is MZ. Table 9 presents the discriminant function coefficients for the 10 questions.

Each twin pair was assigned to the MZ or DZ group on the basis of a zygosity score derived from the raw score coefficients and the mother's replies to the 10

Table 9

Discriminant Function Coefficients

Question	Raw Score Coefficient
1. Similar height?	-0.14314
2. Similar weight?	1.81723
3. Similar facial appearance?	1.84217
4. Similar hair color?	4.67957
5. Similar eye color?	5.53449
6. Similar complexion?	2.13925
7. Like two peas in a pod?	1.63211
8. Confused by father or mother?	1.07532
9. Confused by others in family?	0.43248
10. Confused by strangers?	9.30071

questions. To obtain this zygosity score, each coefficient was multiplied by the value assigned to the mother's response for that particular question and the products for the 10 questions were summed. The discriminant analysis generated a discriminant cutoff point separating MZ and DZ groups. Pairs whose zygosity score was greater than 26.70860 were assigned to the MZ group; pairs with scores equal to or below this point were diagnosed DZ. The mean zygosity score for the MZ twins was 42.44 (SD = 1.8) and it was 14.26 (SD = 9.2) for the DZ twins.

Of the completed sample, 22% had been recruited from a twin club, with the remainder solicited through schools, personal referrals, and referrals from participating twins. Of the 11 pairs of twins from the twin club, six of these pairs were MZ twins and five of the pairs were DZ twins. Analyses of variance were conducted on each of the three tasks in order to statistically evaluate whether within-pair differences of twins from the twin club were significantly different from within-pair differences of twins that were not members of the twin club. The results of these analyses indicated that there were no significant differences between club and non-club members. The total sample was composed of twins from the five boroughs of New York City, Long Island, Westchester County, and New Jersey.

Of the 50 pairs of twins, 27 male pairs and 23 female pairs were included. The MZ group consisted of 13 male pairs and 12 female pairs. Likewise, the DZ group was composed of 14 male pairs and 11 female pairs. Three female pairs were ineligible to participate because they had very different birthweights (> 256 grams).

Participants in this study ranged in age from 16 years to 46 years. The mean age of the males was 23.89 (SD = 8.3) and the mean age of the females was 23.00 (SD = 8.1). A t test was conducted to determine whether these mean ages were statistically different from one another and it was found not to be significant, $t(48) = 0.40$, two-tailed. The mean age of the MZ twins was 26.64 (SD = 8.9) and it was 20.32 (SD = 6.0) for the DZ pairs. A t test was conducted to test whether there was a significant difference between the mean ages of these two groups as well. This test indicated that these means were statistically different, $t(48) = 2.95$, $p < .01$, two-tailed. The consequences of this age difference are discussed later.

Procedure

This study was conducted in two major parts, a psychometric test and two psychological tasks. All subjects first participated in the sequential pattern learning task. They were then given a 3-minute break. The frequency of occurrence task was next. After a 2-minute break, an abbreviated version of the WAIS-R was administered.

General Intelligence

Materials. An abbreviated form of the WAIS-R was administered to all twin pairs. The following subtests were included: picture arrangement, vocabulary, block design, and arithmetic. Scarr and Weinberg (1978) selected these same subtests in their family study of intelligence. Extensive research has demonstrated that various subtest combinations correlate over .90 with Full Scale IQ's (Doppelt, 1956; Levy, 1968). Also, as previously mentioned, mothers of twins were asked to complete a questionnaire on their twins.

Sequential Pattern Learning

Materials. The stimulus items, as shown in Table 10, were 10 Thurstone-like Letter Series Completion Problems. The first problem was used as a practice problem. The following nine problems were the actual

Table 10

Thurstone-like Letter Series Completion Problems

Pattern	Direction Factor	Number Factor
Practice		
cadaeafagahaiaja		
Test		
lhylhxlhwlhvlhul	B	1
cqmdrnesoftpguqh	F	3
ufpueougqudnuhru	F/B	2
kbskcskdskeskfsk	F	1
wlsvkrxmtujqynut	F/B	3
pliokhnjgmitlhek	B	3
nfaogaphaqiarjas	F	2
rzfqzepzdozcnzbm	B	2
aduacuaeuabuafua	F/B	1

Note. F refers to a forward operation, B refers to a backward operation, and F/B refers to a combination of both forward and backward operations.

The numbers 1, 2, and 3 refer to the number of items which change within a string of the pattern.

test problems. Two independent variables were manipulated. The first variable referred to the type of operation that must be perfomed in order to correctly continue the pattern. Three levels of this factor were presented: (a) a forward operation (F); (b) a backward operation (B); and (c) a combination of both the forward and backward operations (F/B). The second variable referred to the number of items which change within a string of the sequential pattern. Three levels of this factor were presented as well: (a) one item changing; (b) two items changing; and (c) all three items changing. The order of presentation of these test patterns was randomly selected.

Procedure. Subjects were seated at a Tandy Model 600 portable computer and the following instructions appeared on the screen:

> This is a sequential pattern learning experiment. You will first receive one (1) practice trial before the testing phase actually begins. You will see items made up of strings of letters. A sequence of three letters will be presented and you are to search for and recognize a pattern in the string of letters. Your task is to carefully

> complete this string by adding 13 additional letters; there will be a total of 16 letters in this pattern. If you wish to correct any item, please use the BKSP key prior to pressing the enter (return) key.
>
> If you are correct the next new string of letters will be presented. However, if you are incorrect, the previous letters will be presented with one additional letter added to the string. As before, you are to complete this pattern so that there is a total of 16 letters. Here is the practice trial.

Subjects started with the practice sequence and continued until they had successfully completed this item. Subjects then proceeded to the testing phase of the experiment. They were instructed to work at their own pace.

<u>Frequency of Occurrence</u>

<u>Materials</u>. A total of 16 nouns ranging from four to eight letters in length were selected as experimental items. These items are presented in Table 11. Each item in the presentation list was a noun of high (30-56 per million) Thorndike-Lorge

Table 11

Experimental Items in the Frequency of Occurrence Task

Test Items	Frequency	Buffer Items	Filler Items
Weapon	2	Skill	Vision
Partner	4	Dignity	Curve
Rice	6	Crew	Lane
Funeral	8	Emotion	Mixture
Lawyer	2	Joint	Bench
Wisdom	4	Stomach	Museum
Bride	6		Mankind
Proof	8		Luck
Mystery	2		
Bureau	4		
Error	6		
Phrase	8		
Weakness	2		
Magic	4		
Host	6		
Mayor	8		

(1944) and Kucera & Francis (1967) count. These stimulus items were selected after independent frequency estimates indicated that they were equivalent with respect to subjective frequency as well.

Stimulus presentation proceeded in a manner adopted from Hasher and Zacks (1979). The entire series consisted of a set of 94 presentations. The first three and the last three presentations were included as buffers to counteract any potential primacy or recency effects. Thus, the main body of the list consisted of 88 presentations. Of these 88 stimuli, 80 were composed of items occurring 2, 4, 6, or 8 times during the list. These critical items were distributed throughout the list so that each quarter of the main body of the list contained one item at each of the four possible frequency levels. Filler items, which were presented only once, were distributed throughout the list, with two fillers present in each quarter of the list. The list was constructed in such a way that any particular item never appeared twice in a row. In addition, there were four critical items that did not occur on the study trial but did occur on the test trial. These represent the zero-occurring items.

Procedure. This phase of the research was run in two parts, a study phase and a testing phase. Subjects were seated at a Tandy Model 600 portable computer and the following instructions appeared on the screen:

> This is a general memory experiment. You will see a list of words. Your task is to attend to each item as it is presented. Please tell the examiner when you are ready to begin.

Each word appeared individually in the center of the screen. Presentation of the entire list of 94 words took 4 minutes and 20 seconds. The following set of instructions were then presented on the screen:

> Stimulus items will be presented one at a time. Your task is to decide how frequently each item occurred during the learning phase. If you are uncertain, make a guess.

A total list of 20 test items were presented. No feedback was given about the correctness of their responses.

Results

General Intelligence

Since all twins were tested by this investigator, the subjects' protocols were independently rescored by another graduate student. The rescorer was blind with respect to the initial scores and the zygosity of the individuals. In addition, she did not have knowledge of which data sets were from members of the same twin pair. This method of rescoring is a common safeguard against some sources of bias. The inter-scorer reliability coefficients for the subscales were: .994 for picture arrangement, .936 for vocabulary, .997 for block design, and .995 for arithmetic. The discrepancies for the picture arrangement, block design, and arithmetic subtests were due to occasional arithmetic errors which were later corrected. The discrepancies for the vocabulary subtest were resolved in conference.

The four subtest scaled scores were transformed to a full scale IQ score using the regression procedure developed by Doppelt (1956), in which the sum of the scaled scores of the omitted subtests is prorated from the four available scores and adjusted by a constant, which is a function of the subject's age. Doppelt reported a correlation between the short-form IQ and the complete WAIS-R IQ ranging from .95 to .96 depending on the subject's age.

The mean full scale IQ score for this sample was 101.2 (SD = 11.9). This mean was converted into a standard score ($z = 0.08$) and it was found not to be significantly different from the population mean of 100. The mean full scale IQ score for the MZ twins was 99.6 (SD = 12.1), and it was 102.8 (SD = 11.5) for DZ twins. These means were converted into standard scores ($z_{mz} = -0.03$ and $z_{dz} = 0.19$) and they were also found not to be significantly different from the population mean.

Similarly, the standard deviation of the prorated sum of scaled scores for this sample was compared with the population sum of scaled scores standard deviation. Essentially, the standard deviation for this sample was compared with the standard deviation for the standardization sample. The standard deviation of this sample (SD = 16.4) was found to be significantly different from the standard deviation of the standardization sample (SD = 22.8), Hartley's $F_{max}(2, 99) = 1.96$, $p < .05$. The standard deviation of the prorated sum of the scaled scores was 16.3 for the MZ twins and it was 16.5 for the DZ twins. These values were also found to be significantly different from the standard deviation of the standardized sample, Hartley's $\underline{F}_{max}(2, 49) = 1.96$, $\underline{p} < .05$, and Hartley's $\underline{F}_{max}(2, 49) = 1.91$, $\underline{p} < .05$, respectively.

In addition to evaluating whether the means and standard deviations of the MZ and DZ samples differed from the population, analyses were conducted to determine whether these samples differed from one another. As presented in Table 12, Fisher's $\underline{F}$ test was conducted to determine whether the MZ and DZ means for full scale IQ statistically differed from one another. Likewise, Hartley's $\underline{F}_{max}$ test was conducted to determine whether the standard deviations for the MZ twins and DZ twins were statistically different. These means and standard deviations did not differ significantly, $\underline{F}(1, 48) = 1.07$, and Hartley's $\underline{F}_{max}(2, 49) = 1.12$.

The means, standard deviations, and standard scores for the sum of the scaled scores as well as the four individual subtests for MZ and DZ pairs are also presented in Table 12. Fisher's $\underline{F}$ and Hartley's $\underline{F}_{max}$ values are also presented. These values indicate that the means and standard deviations did not differ significantly between MZ and DZ pairs.

The intraclass correlations for the full scale IQ, the sum of the scaled scores and the four subtests are

Table 12

A Summary of the Results For General Intelligence

	Mean		Standard Score		F^a	Standard Deviation		Hartley's^b F_{max}
	MZ	DZ	MZ	DZ		MZ	DZ	
Picture Arrangement	9.86	10.76	-0.0467	0.2533	2.4141	2.4496	2.5035	1.0446
Vocabulary	10.60	10.56	0.2000	0.1867	0.0045	2.3387	2.1680	1.1636
Block Design	10.36	10.74	0.1200	0.2467	0.3575	2.5052	2.4228	1.0692
Arithmetic	9.54	10.58	-0.1533	0.1933	2.7817	2.4678	2.5642	1.0796
Sum of Scaled Scores	40.36	42.64			1.5729	7.1621	6.7575	1.1234
Full Scale IQ Score	99.59	102.80	-0.0273	0.1867	1.0749	12.1444	11.4513	1.1247

Note. [a]The critical value at $\alpha = .05$ for $F(1, 48)$ is 4.048. [b]The critical value at $\alpha = .05$ for $F_{max}(2, 49)$ is 1.830.

presented for MZ and DZ pairs in Table 13. The MZ twins consistently showed a higher correlation than the DZ twins on full scale IQ, the sum of the scaled scores as well as on each of the four subtests.

The null hypothesis of no genetic variance was evaluated statistically using an *F* obtained by dividing the within-pair variance of the DZ twins by the within-pair variance of the MZ twins:

$$F = 6^2_{wdz} / 6^2_{wmz}.$$

A significant *F* indicates that heredity and environment produce greater differences in DZ twins than environment alone produces in MZ twins. These results for the WAIS-R also appear in Table 12. Group differences were statistically significant for the full scale IQ score, WAIS-R scaled score sum as well as for the vocabulary and arithmetic subtests.

Heritability is a statistic that describes the ratio of genetic to phenotypic variance in a population. A simple algebraic relationship exists between heritability, *H*, and *F*; thus the *F* enables one to estimate heritability as well:

$$H = 1 - (1/F).$$

The heritability estimates for full scale IQ, the sum of the scaled scores as well as for each of the subtests appear in Table 13.

Sequential Pattern Learning

Two independent variables were manipulated: direction and number. The first variable is the type of operation required to correctly continue the pattern. This factor contained three levels: (a) a forward operation (F); (b) a backward operation (B); and (c) a combination of both forward and backward operations (F/B). This is the direction factor (D).

The second factor is the number of items which change within a string of the sequential pattern. This factor consists of three levels as well: (a) one item changing; (b) two items changing; and (c) all three items changing. This is the number factor (N). Raw scores represent the number of letters a subject had to observe before successfully completing a pattern. The grand means and standard deviations (N = 100) for each of the nine conditions are presented in Table 14.

Since each subject served in each of the treatment combinations, a two-way analysis of variance with repeated measures was conducted. As seen in Table 15, this analysis indicated that the direction factor,

Table 13

A Summary of the Results For the Behavior Genetics of General Intelligence

	Intraclass[a] Correlation			
	MZ	DZ	F^b	H^c
Picture Arrangement	.4150	.3023	1.2429	.1955
Vocabulary	.8284	.5398	2.3043*	.5660
Block Design	.7363	.5383	1.6310	.3869
Arithmetic	.6750	.3451	2.1753*	.5403
Sum of Scaled Scores	.8314	.5102	2.5682*	.6106
Full Scale IQ Score	.8515	.5197	2.8565*	.6499

Note. [a] $r = (MS_b - MS_w) / (MS_b + MS_w)$. [b] $F = 6^2_{wdz}/6^2_{wmz}$.
The critical value at $\alpha = .05$ for $F(25, 25)$ is 1.957. [c] $H = 1 - (1/F)$.
*$p < .05$.

Table 14

The Grand Mean and Standard Deviation For Each Thurstone-like Letter Series Completion Problem

Pattern[a]	Mean	Standard Deviation
D_1N_1	5.9600	0.8519
D_1N_2	6.3400	1.4718
D_1N_3	5.9200	2.1305
D_2N_1	6.3700	1.1340
D_2N_2	6.1800	2.2580
D_2N_3	6.3600	2.9423
D_3N_1	9.8300	3.1303
D_3N_2	11.4300	2.9414
D_3N_3	12.0600	3.6731

Note. [a]The subscripts designate the particular levels of the direction (D) and number (N) factors.

(D-1), and the number factor, (N-1), as well as the direction by number interaction, (D-1)(N-1), were statistically significant.

Following the discovery of a significant direction by number interaction, a simple main effects analysis was conducted. This analytical tool is helpful in determining the locus of the interaction. The results, which are presented in Table 16, revealed a significant effect of factor D at each of the levels of factor N and a significant effect of factor N at level three of factor D.

Orthogonal contrasts were conducted to provide uncorrelated pieces of information about the results of this experiment. To illustrate, consider the simple main effect of the direction factor at level one of the number factor. This simple main effect was significant. A set of mutually orthogonal comparisons can extract independent pieces of information from these data. Two comparisons were possible since two degrees of freedom were available. One comparison of interest to this researcher concerned whether or not the forward and backward levels of the direction factor were significantly different from the level involving the combination of forward and backward directions. Hence, one comparison tested the average

Table 15

Two-Way Analysis of Variance With Repeated Measures For Sequential Pattern Learning

Source	SS	df	MS	*F*	*p*
(D-1)	4845.9356	2	2422.9678	342.8227	.0000
(R-1)	2178.9720	99	22.0098		
(D-1)(R-1)	1399.3978	198	7.0677		
(N-1)	90.0956	2	45.0478	12.6354	.0000
(N-1)(R-1)	705.9044	198	3.5652		
(D-1)(N-1)	187.2644	4	46.8161	16.8119	.0000
(D-1)(N-1)(R-1)	1102.7356	396	2.7847		
Total	10510.3050	899			

Table 16

Simple Main Effects For Sequential Pattern Learning

Source	SS	df	MS	*F*	*p*
D at N_1	903.8867	2	451.9433	125.31	.0000
D at N_2	1783.2067	2	891.6033	265.02	.0000
D at N_3	2346.1067	2	1173.0533	207.03	.0000
N at D_1	10.7467	2	5.3733	3.04	.0501
N at D_2	2.2867	2	1.1433	.31	.7307
N at D_3	264.3267	2	132.1633	35.44	.0000

of levels one and two versus level three of the direction factor at level one of the number factor.

Another issue of interest was whether or not the forward direction was statistically different from the backward direction. Accordingly, the second contrast compared level one versus level two of the direction factor at level one of the number factor. Thus, when orthogonal contrasts are constructed in this fashion they will provide very useful information.

Two sets of eight orthogonal contrasts were computed. The first set of comparisons had an emphasis on factor D whereas factor N was emphasized in the second set of comparisons. These orthogonal contrasts are presented in Table 17.

Butterfield et al. (1985) proposed Counts and Relations as ways to account for letter series difficulty. The correlation between overall letter series performance and Counts was .12, and the corresponding correlation for Relations was .71. It is important to note that the value of Counts is primarily determined by factor N whereas both factors D and N contribute to the value of Relations. Thus, these correlations are consistent with the results reported earlier in this study.

The mean total score for sequential pattern

Table 17

Two Sets of Orthogonal Contrasts For Sequential Pattern Learning

Source	SS	df	MS	*F*	*p*
N_1 & N_2 vs. N_3					
Marginal	36.6939	1	36.6939	8.75	.0039
at D_1	3.5267	1	3.5267	1.53	.2190
at D_2	0.4817	1	0.4817	0.11	.7374
at D_3	136.3267	1	136.3267	33.38	.0000
N_1 vs. N_2					
Marginal	53.4017	1	53.4017	18.18	.0000
at D_1	7.2200	1	7.2200	5.87	.0172
at D_2	1.8050	1	1.8050	0.60	.4411
at D_3	128.0000	1	128.0000	37.94	.0000
D_1 & D_2 vs. D_3					
Marginal	4838.0000	1	4838.0000	398.86	.0000
at N_1	895.4817	1	895.4817	139.61	.0000
at N_2	1781.9267	1	1781.9267	334.49	.0000
at N_3	2336.4267	1	2336.4267	287.49	.0000
D_1 vs. D_2					
Marginal	7.9350	1	7.9350	3.96	.0495
at N_1	8.4050	1	8.4050	10.52	.0016
at N_2	1.2800	1	1.2800	0.91	.3415
at N_3	9.6800	1	9.6800	3.02	.0854

learning was 70.5 (SD = 14.1). The mean total score for MZ twins was 70.9 (SD = 13.2), and the mean was 70.0 (SD = 15.0) for DZ twins. These means and standard deviations did not differ significantly, $F(1, 48) = 0.06$, and Hartley's $F_{max}(2, 49) = 1.30$.

The MZ and DZ intraclass correlations were then calculated based on these total scores for sequential pattern learning. The MZ and DZ intraclass correlations were .80 and .23, respectively. An F test was conducted to evaluate the null hypothesis of no genetic variance, $F(25, 25) = 4.83$, $p < .001$. This F generated a heritability estimate of .79 for the sequential pattern learning task.

The discovery of a significant age difference between the MZ and DZ samples made it necessary to evaluate whether or not age significantly influenced performance on the sequential pattern learning task. Accordingly, the correlations between age and within-pair differences were calculated for the MZ and DZ groups. The correlation between age and within-pair differences was .069 for the MZ twins and it was .139 for the DZ twins. Analyses were conducted to determine if these correlations were significantly different from zero. These t tests were found not to be significant. Thus, although the MZ twins were older than the DZ twins, it is evident that this age difference did not significantly influence sequential pattern learning performance.

Frequency of Occurrence

Subjects viewed a list of words which contained stimulus items that were presented either 2, 4, 6, or 8 times. Subjects were then asked to judge the number of times particular test items had been presented. Mean judged (subjective) frequency is plotted against presented (objective) frequency in Figure 2.

The correlation between presented (objective) frequency and judged (subjective) frequency was calculated for each subject. The mean frequency correlation for all subjects was .71 (SD = 0.13). The mean correlation for MZ twins was .70 (SD = 0.12) and it was .72 (SD = 0.14) for DZ twins. These means and standard deviations did not differ significantly, $F(1, 48) = 0.35$, and Hartley's $F_{max}(2, 49) = 1.30$.

MZ and DZ intraclass correlations were then calculated. The MZ intraclass correlation for this frequency task was .20, and it was .13 for DZ twins. It is important to note that the correlation between age and within-pair differences was found to be significantly less than zero for the MZ twins. Thus, the MZ intraclass correlation may be slightly inflated

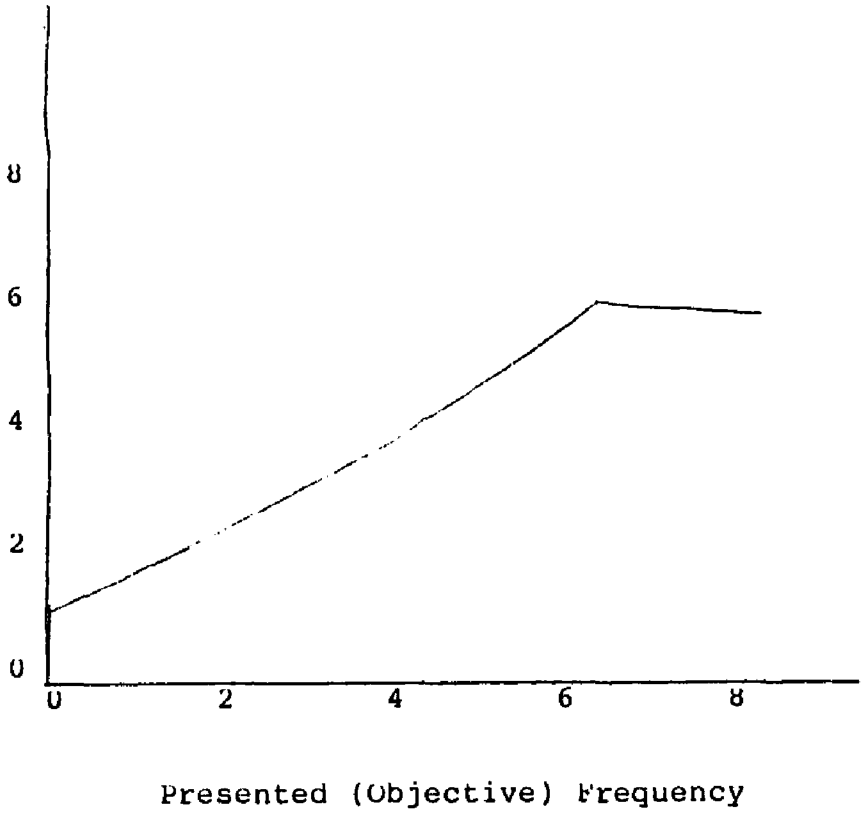


Figure 2. Mean Judged (Subjective) Frequency as a Function of Presented (Objective) Frequency.

due to this age effect. Accordingly, if certain statistical procedures were utilized in order to partial out this age effect, a slight reduction in heritability would be expected. An F test was conducted to evaluate the null hypothesis of no genetic variance, $F(25, 25) = 1.40$, and it was determined not to be significant. This F generated a heritability estimate of .29 for this frequency of occurrence task.

The discovery of a significant age difference between the MZ and DZ samples made it necessary to evaluate whether or not age significantly influenced performance on the frequency of occurrence task. Accordingly, the correlations between age and within-pair differences were calculated for the MZ and DZ groups. The correlations between age and within-pair differences was -.412 for the MZ twins and it was .026 for the DZ twins. As with sequential pattern learning, t tests were conducted to determine if these correlations were significantly different from zero. These tests indicated that the correlation between age and within-pair differences was significantly different from zero for the MZ twins, $t(23) = -2.17$, $p < .05$, two-tailed. Thus, the MZ intraclass correlation may be slightly inflated due to this age effect. If certain statistical procedures were utliized in order to partial out this age effect, a further reduction in heritability would be expected. In this way, the reported estimate of heritability would then be even closer to zero.

The Relationship Between General Intelligence and These Two Specific Cognitive Tasks

In addition to examining the behavior genetics of general intelligence, this study also addresses the relationship between performance on the WAIS-R and performance on the two specific cognitive tasks. The correlation between WAIS-R and sequential pattern learning total scores was -.41. Although a high total score for the WAIS-R indicates better performance, a high total score for sequential pattern learning implies poorer performance. Thus, since these tasks were scored in this manner it is not surprising that the correlation between them was negative. A t test was conducted to test the null hypothesis that the correlation coefficient was less than zero. This test indicated that this correlation was significantly less than zero, $t(98) = -4.40$, $p < .01$, one-tailed.

Whereas the correlation between the WAIS-R and sequential pattern learning task was moderate in strength, the correlation between the WAIS-R and the frequency task was considerably lower (_r_ = .14). As before, a _t_ test was conducted, and it revealed that this correlation was not significantly greater than zero, _t_(98) = 1.54, one-tailed.

Discussion

The primary purpose for administering an abbreviated form of a general intelligence test was for it to function as a marker variable in this study. The rationale is that if it is known how a sample should perform on a particular task and these expectations are fulfilled, then one may be more certain of the results on the other, less predictable tasks. The MZ and DZ intraclass correlations for the sum of the WAIS-R scaled scores were .83 and .51, respectively. These correlations show excellent agreement with the summary statistics of .86 and .60 reported by Bouchard and McGue (1981). They provide further support for a genetic influence on general intelligence. The heritability estimate for the sum of the WAIS-R scaled scores of .61 is also in line with typically reported values which range from .50 to .70.

Examination of the MZ and DZ intraclass correlations for each of the four subtests reveals greater resemblance within MZ than DZ pairs. However, the differences between these intraclass correlations reached statistical significance only for the vocabulary and arithmetic subtests. These MZ and DZ intraclass correlations conform very well with those reported by Segal (1985).

The variability in the magnitude of these subtest correlations raises the possibility that the specific cognitive tasks have differential underlying heritability. Evidence from twin studies using the Wechsler Intelligence Scale for Children (Vandenberg, 1968) and the Wechsler Preschool and Primary Scale of Intelligence (Wilson, 1975) support this interpretation. Performance on the WAIS-R, the frequency of occurrence task, and the sequential pattern learning task in this twin study also provide evidence for differential heritability. In addition, this study supports the finding that heritability estimates for verbal subtests are generally higher than estimates for performance subtests (Williams, 1975).

In contrast, several recent studies (Foch & Plomin, 1980; Loehlin & Nichols, 1976; Plomin & Vandenberg, 1980) have generally failed to provide evidence of differential heritability for the various cognitive skills.

Family studies have also addressed this issue but these studies are generally inconclusive. In the Hawaii Family Study of Cognition, regressions of midchild on midparent were higher for verbal and spatial abilities, and relatively lower for perceptual speed and visual memory (DeFries et al., 1979).

Loehlin and Nichols (1976) have argued against differential heritabilities of specific cognitive skills. They proposed that differences between similar types of measures are responsible for what appear to be differences in heritability for various cognitive skills. At present, the issue is not resolved.

Another aspect of this study concerns the relationship between general intelligence and these two specific cognitive tasks. Investigators have argued that performance on a frequency of occurrence task is relatively independent of intellectual capacity (Hasher & Chromiak, 1977; Hasher & Zacks, 1979, 1984; Zacks, Hasher, & Sanft, 1982). Thus, it was proposed that the correlation between performance on the four subtests of the WAIS-R and the frequency task would not be significantly greater than zero. In contrast, it was hypothesized that the correlation between performance on the WAIS-R and on the sequential pattern learning task would be significantly less than zero (negative because the tasks were scored in opposite directions). The data support these hypotheses.

The effects of two different independent variables on sequential pattern learning performance were also investigated. The two factors were: (a) the type of operation required to continue the sequential pattern, and (b) the number of items which change within a string of the sequential pattern. It was proposed that each of these factors would significantly influence sequential pattern learning performance.

The two-way analysis of variance with repeated measures revealed that each of these factors significantly influenced performance and thus supported this hypothesis. This analysis also revealed a significant interaction between these factors. The presence of an interaction indicates that conclusions based on the two main effects will not fully describe the data. Each of the factors must be interpreted with the levels of the other factor in mind.

A simple main effects analysis was conducted to determine the locus of this interaction. This analysis revealed that factor D exerted a significant effect on sequential pattern learning performance at each level of factor N. Similarly, but not as dramatically, factor N was apparently only effective at level three of factor D.

Orthogonal contrasts were calculated to extract independent pieces of information about these simple main effects. These analyses indicated that the following comparisons were statistically significant ($\underline{p} < .001$): (a) the contrasts comparing the average of levels one and two versus level three of factor D were significant at each level of factor N; (b) the contrasts comparing the average of levels one and two versus level three of factor N were only significant at level three of factor D; and (c) the contrasts comparing level one versus level two of factor N were only significant at level three of factor D.

Two key points emerge from these analyses. First, the type of operation required to continue the sequential pattern (factor D) significantly influenced sequential pattern learning performance regardless of whether one, two, or three items changed within a string of the pattern. This was primarily due to subjects' poorer performance on the patterns which required a combination of forward and backward operations.

Second, the number of items which change within a string of the pattern (factor N) significantly influenced performance only when a combination of forward and backward operations was required to correctly continue the pattern. Apparently, whether one, two or three items were changing within a cluster did not influence performance on the forward or backward patterns. However, when the combination of forward and backward operations were required, the number of items which change did significantly influence performance. It seems reasonable to expect that if the number factor were going to have an impact on performance, it would significantly influence the most difficult patterns (F/B) rather than the easier ones (F and B).

Another goal of this research was to examine whether there is any evidence of genetic influence on sequential pattern learning performance. The MZ and DZ intraclass correlations ($r_{mz} = .80$ and $r_{dz} = .23$) for sequential pattern learning provide support for a genetic influence. The data strongly suggest that identical twins are more alike than fraternal twins on this sequential pattern learning task.

Performance on a frequency of occurrence task was also evaluated in this study. The data on overall frequency performance replicate the various studies conducted by Hasher and Zacks (1979, 1984). As with sequential pattern learning, this research also examined whether there is any evidence of genetic influence on frequency performance. It was proposed that if there is any such evidence, it would be greater in the sequential pattern learning task than in the frequency of occurrence task. There was minimal evidence of heritability for the frequency of occurrence task ($\underline{H}$ = .29). Thus, the data reported here fully support this hypothesis.

In conclusion, the present study provides evidence that heredity contributes substantially to measured general intelligence. Studies that demonstrate genetic factors underlying mental ability are slowly accumulating and their consistency is impressive. Results from this present twin study are also confirmatory with respect to a genetic influence on sequential pattern learning performance. Although, as far as the author is aware, this is the first time sequential pattern learning has been studied with respect to behavior genetics, there appears to be a strong relationship between increasing genetic relatedness and increasing resemblance in performance. In contrast, this study provided virtually no evidence of heritability for performance on a frequency of occurrence task.

Appendix A

The first six questions concern the degree to which the twins were similar for specific physical characteristics. Please rate the twins on these characteristics using the following scale:

0 = not at all similar

1 = somewhat similar

2 = exactly similar or identical

1 - height ________

2 - weight ________

3 - facial appearance ________

4 - hair color ________

5 - eye color ________

6 - complexion ________

The next four questions involve general identity and confusion. Please use the following scoring system:

0 = no

1 = yes

7 - Do they look as alike as two peas in a pod? ____

8 - Does either mother or father ever confuse them? ____

9 - Are they sometimes confused by other people in the family? ____

10 - Is it hard for strangers to tell them apart? ____

Do you believe your twins to be identical or fraternal? ________________

How much did each of them weigh at birth?

______________________ ________lbs. ______oz.

______________________ ________lbs. ______oz.

We appreciate your help very much.
Thank you.

Appendix B

WAIS-R P.A.	WAIS-R Vocab.	WAIS-R Block Design	WAIS-R Arith.	WAIS-R Full Scale Score	Total[a] Score For SPL	Correlation[b] For FO Task	Zygosity Score
Monozygotic Pairs							
11	11	12	12	106.5	72	.7371	44.17976
8	11	11	11	99.5	74	.6470	
7	11	13	9	97.0	70	.5775	42.33759
11	11	11	7	99.5	74	.7531	
11	10	8	8	93.5	76	.5873	43.10444
15	8	10	10	102.0	64	.6589	
8	9	10	8	90.5	68	.7952	44.17976
11	11	10	10	103.0	56	.6833	
10	9	9	6	90.0	59	.7280	44.17976
6	8	10	8	88.5	74	.7777	
12	12	8	6	94.0	77	.6130	42.50567
12	13	9	8	101.0	70	.7916	
8	10	9	11	96.0	72	.7649	42.33759
9	11	9	11	98.0	66	.8582	
8	7	10	13	95.0	74	.4891	44.17976
7	7	11	13	95.0	53	.7048	

WAIS-R P.A.	WAIS-R Vocab.	WAIS-R Block Design	WAIS-R Arith.	WAIS-R Full Scale Score	Total[a] Score For SPL	Correlation[b] For FO Task	Zygosity Score
8	9	13	8	94.5	77	.7575	44.17976
8	8	12	7	90.0	89	.7531	
10	9	8	6	86.0	66	.7113	44.17976
11	9	5	7	86.0	70	.8403	
8	9	10	9	90.5	72	.7486	40.87356
9	8	10	8	90.5	67	.6506	
9	14	15	15	122.5	68	.7624	39.63016
14	13	15	15	129.5	57	.8420	
12	11	11	10	103.0	54	.9027	44.17976
13	11	12	10	108.5	53	.8083	
8	8	5	8	83.0	84	.5302	43.10444
6	8	6	9	83.0	83	.6334	
12	13	12	7	104.5	83	.5646	42.54765
12	13	14	9	112.5	76	.6977	
13	16	10	17	126.0	61	.5448	44.17976
11	15	13	10	116.0	51	.8423	
10	11	10	10	103.5	50	.7167	40.70548
7	9	11	11	97.0	52	.8401	
7	9	8	8	87.0	99	.5178	39.63016
11	12	9	9	101.0	98	.3979	

WAIS-R P.A.	WAIS-R Vocab.	WAIS-R Block Design	WAIS-R Arith.	WAIS-R Full Scale Score	Total[a] Score For SPL	Correlation[b] For FO Task	Zygosity Score
10	14	12	10	110.5	58	.4357	37.95607
14	14	11	11	114.5	63	.7354	
10	11	12	6	96.5	70	.7387	44.17976
11	11	15	8	105.5	81	.4639	
17	13	13	13	128.0	56	.7245	40.70548
12	16	15	13	128.0	53	.8324	
7	9	11	12	97.0	79	.6815	43.36253
7	10	6	9	86.0	73	.8271	
10	11	7	8	93.5	83	.7718	40.70548
9	11	6	7	87.0	85	.7167	
8	6	11	8	88.0	62	.6792	41.28721
7	8	10	9	88.0	70	.9236	
10	10	11	10	98.0	103	.6502	42.50567
8	12	9	9	94.5	98	.5424	
Dizygotic Pairs							
9	12	11	9	101.0	85	.6784	18.40989
8	10	14	8	99.0	59	.8275	
12	9	6	9	91.5	99	.7809	22.54525
12	9	11	10	99.5	56	.9057	

WAIS-R P.A.	WAIS-R Vocab.	WAIS-R Block Design	WAIS-R Arith.	WAIS-R Full Scale Score	Total[a] Score For SPL	Correlation[b] For FO Task	Zygosity Score
11	12	13	10	109.0	54	.7899	22.54525
11	9	13	9	102.0	58	.7760	
9	10	11	13	102.0	76	.8232	25.17028
5	6	7	10	80.5	90	.6511	
12	18	13	14	124.5	53	.8397	10.92584
15	15	14	14	131.0	62	.8301	
11	10	10	14	107.0	65	.5330	1.67409
13	12	12	12	113.5	53	.8865	
12	12	15	19	132.5	63	.7088	20.02705
13	11	13	15	114.5	70	.4691	
12	11	13	11	113.0	67	.5759	5.39135
12	13	15	16	126.0	58	.9078	
12	11	8	11	102.0	74	.6802	1.81723
3	7	6	10	78.0	56	.6553	
13	11	9	11	103.0	61	.7869	25.94049
9	10	10	10	95.0	62	.9436	
12	12	14	8	111.0	54	.7493	3.83828
12	11	14	13	107.0	67	.7308	
7	11	10	10	94.0	59	.6779	2.13925
11	10	10	8	96.0	57	.5219	

WAIS-R P.A.	WAIS-R Vocab.	WAIS-R Block Design	WAIS-R Arith.	WAIS-R Full Scale Score	Total[a] Score For SPL	Correlation[b] For FO Task	Zygosity Score
11	8	10	14	102.0	54	.9021	-0.14314
12	7	12	7	91.5	92	.5305	
13	9	14	10	108.0	86	.8766	24.56348
9	13	12	13	113.0	64	.7640	
11	13	7	8	96.0	66	.6613	26.08363
15	13	9	11	112.5	73	.4969	
10	9	10	12	102.0	67	.8423	25.38548
10	11	11	10	102.0	77	.5010	
14	15	8	8	104.5	55	.7525	7.53060
17	12	9	12	119.0	70	.8945	
8	11	11	9	94.5	77	.7102	5.51237
8	10	8	10	90.0	97	.5709	
13	9	15	8	106.5	80	.3833	11.19000
8	8	11	6	88.5	88	.5061	
11	9	8	6	90.0	84	.8309	9.51591
11	10	12	11	108.0	71	.7674	
9	8	10	15	99.5	70	.6565	19.56189
12	10	10	10	99.5	81	.5522	
8	10	10	9	94.0	58	.5384	26.08363
9	9	10	9	94.0	61	.7061	

WAIS-R P.A.	WAIS-R Vocab.	WAIS-R Block Design	WAIS-R Arith.	WAIS-R Full Scale Score	Total[a] Score For SPL	Correlation[b] For FO Task	Zygosity Score
10	10	7	11	94.5	127	.7598	11.19000
7	9	8	10	89.5	85	.6037	
11	9	9	10	97.0	82	.8548	13.56224
12	10	11	10	102.0	65	.7067	
10	12	13	9	107.5	55	.8259	16.01271
11	12	10	10	102.0	59	.8648	

<u>Note</u>.[a] SPL refers to the sequential pattern learning task.

[b] FO refers to the frequency of occurrence task.

References

Alba, J. W., Chromiak, W., Hasher, L., & Attig, M. S. (1980). Automatic encoding of category size information. Journal of Experimental Psychology: Human Learning and Memory, 6, 370-378.

Allen, R., & Reber, A. S. (1980). Very long term memory for tacit knowledge. Cognition, 6, 189-221.

Attig, M., & Hasher, L. (1980). The processing of frequency of occurrence information by adults. Journal of Gerontology, 35, 66-69.

Bayley, N. (1970). Development of mental abilities. In P. H. Mussen (Ed.), Carmichael's manual of child psychology. New York: Wiley.

Blewett, D. B. (1954). An experimental study of the inheritance of intelligence. Journal of Mental Science, 100, 922-933.

Block, J. B. (1968). Hereditary components in the performance of twins on the WAIS. In S. G. Vandenberg (Ed.), Progress in human behavior genetics (pp.221-228). Baltimore, MD: Johns Hopkins University Press.

Bock, R. D., & Kolakowski, D. (1973). Further evidence of sex-linked major-gene influence on human spatial visualizing ability. American Journal of Human Genetics, 25, 1-14.

Bouchard, T. J., Jr., & McGee, M. G. (1977). Sex differences in human spatial ability: Not an X-linked recessive gene effect. Social Biology, 24, 332-335.

Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. Science, 212, 1055-1059.

Burks, B. S. (1928). The relative influence of nature and nurture upon mental development: A comparative study of foster parent-foster child resemblance and true parent-true child resemblance. Twenty-Seventh Yearbook of the National Society for the Study of Education, 27, 219-316.

Burnett, S. A., & Stevenson, J. M. (1979). Instructional effects on frequency judgments for sentence meanings. American Journal of Psychology, 92, 711-721.

Butterfield, E. C., Nielson, D., Tangen, K. L., & Richardson, M. B. Theoretically based psychometric measures of inductive reasoning. In S. E. Embretson (Ed.), Test design: Developments in psychology and psychometrics. (pp.77-147). NY: Academic Press.

Cohen, D. J., Dibble, E., Grawe, J. M., & Pollin, W. (1973). Separating identical from fraternal twins. Archives of General Psychiatry, 29, 465-469.

Cohen, D. J., Dibble, E., Grawe, J. M., & Pollin, W. (1975). Reliably separating identical from fraternal twins. Archives of General Psychiatry, 32, 1371-1375.

Corley, R. P., DeFries, J. C., Kuse, A. R., & Vandenberg, S. G. (1980). Familial resemblance for the identical blocks test of spatial ability: No evidence for X-linkage. Behavior Genetics, 10, 211-215.

DeFries, J. C., Johnson, R. C., Kuse, A. R., McClearn, G. E., Mi, M. P., Rashad, M. N., Vandenberg, S. G., & Wilson, J. R. (1976). Parent-offspring resemblance for specific cognitive abilities in two ethnic groups. Nature, 261, 131-133.

DeFries, J. C., Johnson, R. C., Kuse, A. R., McClearn, G. E., Polovina, J., Vandenberg, S. G., & Wilson, J. R. (1979). Familial resemblance for specific cognitive abilities. Behavior Genetics, 9, 23-43.

DeFries, J. C., Vandenberg, S. G., McClearn, G. E., Kuse, A. R., Wilson, J. R., Ashton, G. C., & Johnson, R. C. (1974). Near identity of cognitive structures in two ethnic groups. Science, 183, 338-339.

Doppelt, J. E. (1956). Estimating the full scale score of the Wechsler Adult Intelligence Scale from scores on four subtests. Journal of Consulting Psychology, 20, 63-66.

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

Falconer, D. S. (1960). Introduction to quantitative genetics. New York: Ronald Press.

Fisch, R. O., Bilek, M. K., Deinard, A. S., and Chang, P. N. (1976). Growth, behavioral, and psychologic measurements of adopted children: The influences of genetic and socioeconomic factors in a prospective study. The Journal of Pediatrics, 89, 494-500.

Foch, T. T., & Plomin, R. (1980). Specific cognitive abilities in 5- to 12-year-old twins. Behavior Genetics, 10, 507-520.

Fuller, J. L., & Thompson, W. R. (1978). Foundations of behavior genetics. St. Louis: Mosby.

Gambino, B., & Myers, J. L. (1967). Role of event runs in probability learning. Psychological Review, 74, 410-419.

Garner, W. R., & Gottwald, R. L. (1967). Some perceptual factors in the learning of sequential patterns of binary events. Journal of Verbal Learning and Verbal Behavior, 6, 582-589.

Greeno, J. G., & Simon, H. A. (1974). Processes for sequence production. Psychological Review, 81, 187-198.

Gude, C., & Zechmeister, E. B. (1975). Frequency judgements for the gist of sentences. American Journal of Psychology, 88, 385-396.

Guttman, R. (1974). Genetic analysis of analytical spatial ability: Raven's Progressive Matrices. Behavior Genetics, 4, 273-284.

Guttman, R., & Shoham, I. (1979). Intrafamilial invariance and parent-offspring resemblance in spatial abilities. Behavior Genetics, 9, 367-378.

Hartlage, L. C. (1970). Sex-linked inheritance of spatial ability. Perceptual and Motor Skills, 31, 610.

Hasher, L., & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? Journal of Verbal Learning and Verbal Behavior, 16, 173-184.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. Journal of Experimental Psychology: General, 108, 356-388.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information. American Psychologist, 39, 1372-1388.

Hersh, H. M. (1974). The effect of irrelevant relations on the processing of sequential patterns. Memory and Cognition, 2, 771-774.

Hintzman, D. L. (1969). Apparent frequency as a function of frequency and the spacing of repetitions. Journal of Experimental Psychology, 80, 139-145.

Hintzman, D. L., & Stern, L. D. (1978). Contextual variability and memory for frequency. Journal of Experimental Psychology: Human Learning and Memory, 4, 539-549.

Holzinger, K. J. (1929). The relative effect of nature and nurture on twin differences. Journal of Educational Psychology, 20, 241-248.

Holzman, T. G., Pellegrino, J. W., & Glaser, R. (1982). Cognitive dimensions of numerical rule induction. Journal of Educational Psychology, 74, 360-373.

Holzman, T. G., Pellegrino, J. W., & Glaser, R. (1983). Cognitive variables in series completion. Journal of Educational Psychology, 75, 603-618.

Horn, J. M., Loehlin, J. C., & Willerman, L. (1975). The Texas adoption project. In L. Carter-Saltzman (Chair), Adoption studies: Similarities of children to adoptive and natural parents. Symposium presented at the Behavior Genetics Association, Texas.

Horn, J. M., Loehlin, J. C., & Willerman, L. (1979). Intellectual resemblance among adoptive and biological relatives: The Texas adoption project. *Behavior Genetics*, *9*, 177-207.

Howell W. G. (1973). Storage of events and event frequencies: A comparison of two paradigms in memory. *Journal of Experimental Psychology*, *98*, 260-263.

Jacoby, L. L. (1972). Context effects on frequency judgments of words and sentences. *Journal of Experimental Psychology*, *94*, 255-260.

Jensen, A. R., & Marisi, D. Q. (1979). A note on the heritability of memory span. *Behavior Genetics*, *9*, 379-387.

Johnson, R. C., DeFries, J. C., Wilson, J. R., McClearn, G. E., Vandenberg, S. G., Ashton, G. C., Mi, M. P., & Rashad, M. N. (1976). Assortative marriage for specific cognitive abilities in two ethnic groups. *Human Biology*, *48*, 343-352.

Jones, M. R. (1973). Higher order organization in serial recall of digits. *Journal of Experimental Psychology*, *99*, 106-119.

Jones, M. R., & Zamostny, K. P. (1975). Memory and rule structure in the prediction of serial patterns. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 295-306.

Kasriel, J., & Eaves, L. (1976). The zygosity of twins: Further evidence on the agreement between diagnosis by blood groups and written questionnaires. *Journal of Biosocial Science*, *8*, 263-266.

Kausler, D. H., & Puckett, J. M. (1980). Frequency judgments and correlated cognitive abilities in young and elderly adults. *Journal of Gerontology*, *35*, 376-382.

Keeney, T. J. (1969). Permutation transformations on phrase structures in letter sequences. *Journal of Experimental Psychology*, *82*, 28-33.

Kotovsky, K., & Simon, H. A. (1973). Empirical tests of a theory of human acquisition of concepts for sequential patterns. Cognitive Psychology, 4, 399-424.

Leahy, A. (1932). A study of certain selective factors influencing prediction of the mental status of adopted children, or adopted children in nature-nurture research. Journal of Genetic Psychology, 41, 294-329.

Leahy, A. (1935). Nature-nurture and intelligence. Genetic Psychology Monographs, 17, 241-306.

Leicht, K. L. (1968). Recall and judged frequency of implicitly occurring words. Journal of Verbal Learning and Verbal Behavior, 7, 918-923.

Levy, P. (1968). Short-form tests: A methodological review. Psychological Bulletin, 69, 410-416.

Lewis, M. (1973). Infant intelligence tests: Their use and misuse. Human Development, 16, 108-118.

Loehlin, J. C., & Nichols, R. C. (1976). Heredity, environment, and personality. Austin: University of Texas Press.

Loehlin, J. C., Sharan, S., & Jacoby, R. (1978). In pursuit of the "spatial gene": A family study. Behavior Genetics, 8, 27-41.

Lykken, D. T., Tellegen, A. & DeRubeis, R. (1978). Volunteer bias in twin research: The rule of two-thirds. Social Biology, 25, 1-9.

Marmurek, H. H. C., & Johnson, N. F. (1978). Hierarchical organization as a determinant of sequential learning. Memory and Cognition, 6, 240-245.

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

McGee, M. G. (1978). Intrafamilial correlations and heritability estimates for spatial ability in a Minnesota sample. Behavior Genetics, 8, 77-80.

Munsinger, H. (1975). Children's resemblance to their biological and adopting parents in two ethnic groups. Behavior Genetics, 5, 239-254.

Munsinger, H. (1975). The adopted child's IQ: A critical review. Psychological Bulletin, 82, 623-659.

Munsinger, H. (1977). The identical-twin transfusion syndrome: A source of error in estimating IQ resemblance and heritability. Annals of Human Genetics, 40, 307-321.

Munsinger, H., & Douglass, A. (1976). The syntactic abilities of identical twins, fraternal twins, and their siblings. Child Development, 47, 40-50.

Myers, J. L. (1970). Sequential choice behavior. In G. H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (pp. 109-170). New York: Academic Press.

Myers, J. L. (1976). Probability learning and sequence learning. In W. K. Estes (Ed.), Handbook of learning and cognitive processes (pp. 171-205). Hillsdale, NJ: Erlbaum.

Nichols, R. C. (1965). The inheritance of general and specific ability. National Merit Scholarship Research Reports, 1, 1-10.

Olive, H. (1972). Sibling resemblances in divergent thinking. Journal of Genetic Psychology, 120, 155-162.

Osborne, R. T., & Gregor, A. J. (1966). The heritability of visualization, perceptual speed, and spatial orientation. Perceptual and Motor Skills, 23, 379-390.

Osborne, R. T., & Gregor, A. J. (1968). Racial differences in heritability estimates for tests of spatial ability. Perceptual and Motor Skills, 27, 735-739.

Osborne, R. T., & Gregor, A. J., & Miele, F. (1968). Heritability of factor V: Verbal comprehension. Perceptual and Motor Skills, 26, 191-202.

Osborne, R. T., & Miele, F. (1969). Racial differences in environmental influences on numerical abilities as determined by heritability estimates. Perceptual and Motor Skills, 28, 535-538.

Park, J., Johnson, R. C., DeFries, J. C., McClearn, G. E., Mi, M. P., Rashad, M. N., Vandenberg, S. G., & Wilson, J. R. (1978). Parent-offspring resemblance for specific cognitive abilities in Korea. Behavior Genetics, 8, 43-52.

Pezzulo, T. R., Thorsen, E. E., & Madaus, G. F. (1972). The heritability of Jensen's Level I and II and divergent thinking. American Educational Research Journal, 9, 539-546.

Plomin, R., & DeFries, J. C. (1980). Genetics and intelligence: Recent data. Intelligence, 4, 15-24.

Plomin, R., & Vandenberg, S. G. (1980). An analysis of Koch's (1966) Primary Abilities test data for 5- to 7-year-old twins. Behavior Genetics, 10, 409-412.

Reed, S. C., & Rich, S. S. (1982). Parent-offspring correlations and regressions for IQ. Behavior Genetics, 12, 535-542.

Restle, F. (1961). Psychology of judgement and choice. New York: Wiley.

Restle, F. (1966). Run structure and proability learning: Disproof of Restle's model. Journal of Experimental Psychology, 72, 382-389.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. Psychological Review, 77, 481-495.

Restle, F., & Brown, E. R. (1970a). Organization of serial pattern learning. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 249-331). New York: Academic Press.

Restle, F., & Brown, E. R. (1970b). Serial pattern learning. Journal of Experimental Psychology, 83, 120-125.

Restle, F., & Brown, E. R. (1970c). Serial pattern learning: Pretraining of runs and trills. Psychonomic Science, 19, 321-322.

Reznikoff, M., Domino, G., Bridges, C., & Honeyman, M. (1973). Creative abilities in identical and fraternal twins. Behavior Genetics, 3, 365-377.

Rose, R. J., Miller, J. Z., Dumont-Driscoll, M., & Evans, M. M. (1979). Twin-family studies of perceptual speed ability. Behavior Genetics, 9, 71-86.

Scarr, S. (1968). Environmental bias in twin studies. In S. G. Vandenberg (Ed.), Progress in human behavior genetics (pp. 205-213). Baltimore, MD: Johns Hopkins University Press.

Scarr, S., & Weinberg, R. A. (1977). Intellectual similarities within families of both adopted and biological children. Intelligence, 1, 170-191.

Scarr, S., & Weinberg, R. A. (1978). The influence of "family background" on intellectual attainment. American Sociological Review, 43, 674-692.

Segal, N. L. (1985). Monozygotic and dizygotic twins: A comparative analysis of mental ability profiles. Child Development, 56, 1051-1058.

Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. Psychological Review, 79, 369-382.

Skodak, M., & Skeels, H. (1949). A final follow-up study of one hundred adopted children. Journal of Genetic Psychology, 75, 21-58.

Smith, R. T. (1965). A comparison of socioenvironmental factors in monozygotic and dizygotic twins, testing an assuption. In S. G. Vandenberg (Ed.), Methods and goals in human behavior genetics (pp. 45-62). New York: Academic Press.

Spuhler, K. P., & Vandenberg, S. G. (1980). Comparison of parent-offspring resemblance for specific cognitive abilities. Behavior Genetics, 10, 413-418.

Stafford, R. E. (1961). Sex differences in spatial visualization as evidence of sex-linked inheritance. Perceptual and Motor Skills, 13, 428.

Stafford, R. E. (1972). Hereditary and environmental components of quantitative reasoning. Review of Educational Research, 42, 183-201.

Teasdale, T. W., & Owen, D. R. (1984). Heredity and familial environment in intelligence and educational level--a sibling study. Nature, 309, 620-622.

Thorndike, E. L., & Lorge, I. (1944). The teacher's word book of 30,000 words. New York: Teachers College, Columbia University, Bureau of Publications.

Vandenberg, S. G. (1968). Progress in human behavior genetics. Baltimore, MD: Johns Hopkins University Press.

Vandenberg, S. G. (1968). The nature and nurture of intelligence. In D. Glass (Ed.), Genetics (pp.3-58). New York: Rockefeller University Press.

Vandenberg, S. G., & Wilson, K. (1979). Failure of the twin situation to influence twin differences in cognition. Behavior Genetics, 9, 55-60.

Vitz, P. C., & Todd, T. C. (1967). A model of learning for simple repeating binary patterns. Journal of Experimental Psychology, 75, 108-117.

Vogler, G. P., & DeFries, J. C. (1983). Linearity of offspring-parent regression for general cognitive ability. Behavior Genetics, 13, 355-360.

Williams, (1975). Family resemblance in abilities: The Wechsler scales. Behavior Genetics, 5, 405-409.

Wilson, R. S. (1975). Twins: Patterns of cognitive development as measured on the WPPSI. Developmental Psychology, 11, 126-134.

Wilson, R. S. (1977). Twins and siblings: Concordance for school-age mental development. Child Development, 48, 211-216.

Zacks, R. T., Hasher, L., & Sanft, H. (1982). Automatic encoding of event frequency: Further findings. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 106-116.

Zonderman, A. B., Vandenberg, S. G., Spuhler, K. P., & Fain, P. R. (1977). Assortative marriage for cognitive abilities. _Behavior Genetics_, _7_, 261-271.