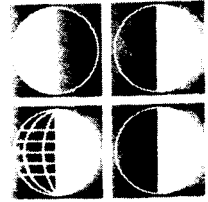# Intelligence and Family Size Reconsidered

Robert D. Retherford and William H. Sewell

Population Institute, East-West Center, Honolulu, Hawaii; and Department of Sociology, University of Wisconsin, Madison, Wisconsin

ABSTRACT: The major purpose of this study is to examine the association between the measured intelligence and fertility of over 9,000 persons who graduated high school in Wisconsin in 1957. Various measures of association are considered, including the IQ selection differential, which provides an estimate of what the generational change in mean IQ would be if, hypothetically, each child in the birth histories had the same IQ as the mean of its parents' IQ's. This is calculated not only for graduates but also, more realistically, for the complete cohort, including dropouts. The IQ selection differential for the complete cohort is estimated to be eight-tenths of an IQ point decline in a generation. The contribution of females to this decline is estimated to be almost five times greater than the contribution of males. The value of eight-tenths may be viewed as an upper bound of the generational decline in mean genotypic IQ for this cohort and its offspring. An educated guess, based partly on genetic models and findings from IQ heritability studies as well as on the above estimate of the IQ selection differential, is that the generational change in mean genotypic IQ is about one-third of an IQ point decline for this cohort and its offspring.

The major purpose of the research reported in this study is to examine the association between the measured intelligence and fertility of over 9,000 persons who graduated high school in Wisconsin in 1957. We have information on measured intelligence from tests taken when our sample members were in high school, and on fertility from birth histories collected in the most recent reinterview survey in 1975, when the respondents were approximately 36 years of age.

Our specific aims are the following:

First, we critically review the research literature on intelligence and family size, focusing especially on studies appearing in the last 25 years.

Second, we examine the relationship between the measured intelligence of our respondents and the number of children born alive to them by age 35. To achieve this aim, we calculate the mean number of children ever born, F(35), by sex and IQ, for IQ deciles. From these values of F(35) and estimates of age-specific fertility at ages below 35, we use a demographic model to extrapolate age-specific fertility after age 35 and hence to estimate total fertility rates (TFR) and gross reproduction rates (GRR) by sex and IQ decile. By making use of information on sample attrition between 1957 and 1975, we also estimate age-specific death rates by sex and IQ decile. From the age-specific birth and death rates by sex and IQ decile, we calculate net reproduction rates (NRR), intrinsic rates of natural increase (r), mean lengths of generation (T), population doubling times (t_d), and relative fitness (W) by sex and IQ decile. As a final measure of the relationship between intelligence and fertility, we calculate estimates of the IQ selection differential (S), which provides an estimate

1

of what the generational change in mean IQ would be if, hypothetically, each child in the birth histories had the same IQ as the mean of its parents' IQ's.

Third, we estimate the mean IQ, fertility, and mortality of high-school dropouts (all those who never graduated high school), which the Wisconsin sample did not include. This is difficult to do, because we lack direct information on dropouts. We are forced to use estimates of the mean IQ and mortality of dropouts based on less than adequate information from other studies. Our estimates of the fertility of dropouts are more satisfactory, since we are able to use census data to estimate the fertility of persons of similar ages and dates of birth in the Wisconsin population who completed less than 12 grades of school.

Fourth, we use the estimates of intelligence, fertility, and mortality of both graduates and dropouts to estimate the IQ selection differential for the complete cohort.

Fifth, we discuss the generational change in mean genotypic IQ for the Wisconsin cohort. In this case our results are highly tentative and imprecise, because of our inability to meet the assumptions and conditions for adequate estimation of change in mean genotypic IQ.

We now turn to the first of these objectives, a critical review of previous studies.

REVIEW OF PREVIOUS STUDIES

It is 32 years since Anastasi published her classic review, "Intelligence and Family Size" (Anastasi, 1956). Her review began by noting that early interest in the relationship between family size and intelligence grew mainly out of eugenic concerns. The commonly occur-

ring negative association between socioeconomic status and size of family led investigators to postulate dysgenic population effects, based on the assumption that higher status positions tend to be occupied by persons of higher intelligence. Investigations relating the measured intelligence of school children to the socioeconomic status of their parents furnished evidence in support of this assumption. More direct evidence came from studies in the United States, England, France, and Greece, all of which found a negative correlation, ranging from -0.17 to -0.32, between the measured intelligence of school children and the size of their family of origin. From these findings several writers, using various estimation procedures, predicted an inevitable decline in the intelligence of the population. These predictions, depending on the sample studied and the estimation procedures used, indicated declines of 2 to 4 IQ points per generation. Although there was not complete agreement on the magnitude of the predicted decline, there was little or no dispute about the negative association between family size and measured intelligence and, hence, the dysgenic effect of negative differential fertility.

Confidence in these predictions was seriously undermined by the findings of surveys undertaken by the Scottish Council for Research in Education (1933, 1949). These surveys, conducted in 1932 and 1947 with response rates of 88 and 89 per cent respectively, used psychometrically equivalent group tests to measure the intelligence of all Scottish eleven-year-old school children. Although the usual negative relationship between the measured intelligence of the students and the size of their sibships was found in both surveys, there was a statistically significant increase of about

2 points in mean test score over the 15-year period between the two surveys, rather than the expected decline. Even after making various psychometric adjustments to the data to take into account possible sources of bias, the most favorable adjustments produced virtually identical means for the two surveys. Subsequently, the results of the Scottish surveys were supported by several other studies that also failed to substantiate the expected decline in test scores for samples drawn from other populations (see especially Cattell, 1951); however, none of these were as adequate in coverage as the Scottish studies. The finding of an increase in the mean IQ of school children over time was surprising also because of the increase in the proportion of the school-age population attending school. Presumably school-goers are brighter on average than those who do not attend school, so that an increase in the proportion attending school should reduce the mean IQ of school children quite independently of other factors.

Needless to say, the results of the Scottish surveys led to controversy and speculation in which scholars either tried to discount the findings because of methodological inadequacies or took the findings as evidence that environmental improvements related to social and education trends had more than offset the dysgenic effects of negative differential fertility by IQ. For a discussion of this controversy, see Duncan (1952).

After examining the available evidence on intelligence and family size, Anastasi concluded that because of complex problems of sampling, measurement, and statistical analysis encountered in investigating the relationship between family size and intelligence, it was impossible with the data at hand to choose among alternative interpretations based on (1) inherited mental abilities; (2) differences in environments provided by parents of varying intellectual levels; and (3) factors inherent in family size itself, such as the fact that in large families, per capita resources available for housing, food, medical attention, parental contact with children, and education are reduced. She further argued that more definitive conclusions about the relationship between family size and intelligence must await better designed studies that would begin with the testing of large and representative samples of children or youth prior to their educational and vocational differentiation. The subjects should be followed until their families are completed, with careful recording of the date of birth of each child, the parents' ages at marriage, education, and socioeconomic status. She believed that with these data it would be possible to determine the extent and nature of the relationship between intelligence and fertility, and that "this approach would help to separate the many interrelated variables which are now intricately intertwined and should thus bring us closer to a causal interpretation of the empirically observed relationship between intelligence and family size" (p. 206).

Unfortunately, no study reported since the publication of Anastasi's 1956 review article has come close to meeting all the criteria she set forth. Nor, for that matter, could one be certain that, had these criteria been met, it would have been possible to unravel the causal relationships among intelligence, social background, and fertility. We, too, are interested in understanding these causal relationships and intend to study them in detail in a subsequent paper. In this paper, however, we are concerned with

the prior question of whether there is a significant relation between the measured intelligence of parents and the number of children they produce.

RECENT STUDIES OF INTELLIGENCE AND FAMILY SIZE

Our review of literature published since Anastasi's article is limited primarily to studies based on relatively large samples that purport to be representative of American cities. states. or the nation as a whole. As our review will show. the question of the extent and direction of the relationship between intelligence and fertility is by no means settled.

Interest in the relationship between intelligence and family size and particularly its possible dysgenic effect seems to have declined in the years since the publication of Anastasi's review. although studies bearing on these questions have continued to appear. Some of these studies have shown that previous research based on the relationship between measured intelligence and size of sibship is flawed and cannot by itself provide an adequate basis for predicting trends in population quality.

The first of these studies. which was carried out at the Dight Institute for Human Genetics at the University of Minnesota. sought a possible resolution of the paradox that family size (number of siblings) was consistently found to be negatively associated with measured intelligence but that predictions of a substantial decline in the IQ of the population had failed to materialize (Higgens et al.. 1962). The investigation employed a sample of families for whom IQ scores and other information were available for both parents and one or more of their children from the files of the Institute. This sample consisted of 1,016 mothers. an equal number of fathers.

and their 2,039 children. The IQ scores of the parents were from tests taken before marriage. in most cases while they were still in school. The children's scores were from tests given in school. Like other investigators before them. the Minnesota group found the expected negative correlation (-0.30) between size of family of origin and the measured intelligence of the child. However, when the never-married and nonreproductive siblings of the mothers and fathers were included in the analysis. the relationship between IQ in the parental generation and number of offspring became slightly positive. In other words. the higher reproductive rate of those in the lower IQ categories who were parents was offset by their siblings who never married or failed to reproduce if married. On the basis of this evidence. the authors concluded that "the IQ level of the whole population should remain relatively static from one generation to the next, or at least not drop rapidly" (p. 199). It must be noted that the sample in this investigation was not a probability sample of the Minnesota population and may not be as representative as the authors believed. Moreover. nothing is said about the comparability of various test scores used in the study.

Several investigators during the past 25 years have reported on research in which persons of known measured intelligence have been followed up over their childbearing years to determine their fertility. The pioneering study of this type was that of Bajema (1963). who obtained life history data on 979 native-white individuals born in 1916–17 whose Terman group intelligence test scores were available from tests given in the Kalamazoo. Michigan. public school system to children in the sixth grade. He later obtained life history information in

a follow-up survey carried out when the cohort of individuals born in 1916 or 1917 reached ages 45–46. The survey included questions on date of birth, sex, number of siblings, marital status, number of offspring, date of death if deceased, and place of residence.

The mean number of children born to each of five IQ categories was computed and showed a bimodal distribution: the highest means were 2.60 for those with IQ ≥ 120 and 2.46 for those with IQ 80–94. The difference in mean fertility between these two IQ categories was not statistically significant, but the differences between each of them and the IQ categories 69–79 and 95–104 were significant. However, the means for the extreme categories were not significantly higher than the mean for the 105–119 IQ category. Further analysis using the intrinsic rate of natural increase and correlation analysis showed that the mean IQ of the population under study had increased slightly due to the small positive relation between parental IQ and the number of offspring produced. The seeming paradox that the intelligence of the individual is positively associated with number of offspring but negatively associated with size of sibship was explained essentially as it was in the study by Higgins et al. (1962). In Bajema's study, however, two different generations are involved, and there is no reason why the relationship between intelligence and fertility must be the same for each; that is, the data on IQ and sibship size reflect, however imperfectly, differential fertility by IQ in the respondents' parents' generation. The data on IQ and number of offspring reflect differential fertility by IQ in the respondent generation itself.

A minor flaw in Bajema's calculations of the intrinsic rate of natural increase is that survivorship is measured from age at IQ testing, which averaged 11.6 years, rather than from age 1. Age 1 should have been used, because Bajema used offspring surviving to age 1 in place of births in his calculations, apparently because reliable data on births were not available. It seems unlikely, however, that correction of this flaw would alter the study's general conclusions.

Guided by the findings of Higgins et al. (1962) and Bajema (1963), Waller (1971) sought a sample that would include the nonreproducing as well as the reproducing members of the parental generation. Using the large data base developed over a period of sixty years by the Dight Institute for Human Genetics, Waller selected a sample of families for a follow-up study. The sample consisted of families for whom IQ test scores were available for both parents and in which the mother was born in or before December 1927. These criteria yielded a study population of 529 families. To these parents were added their reproducing and nonreproducing siblings, if IQ scores were available for them. If an IQ score was available for the spouse of a sibling, that individual was also included. An attempt was made to locate the qualified families and to obtain information from them using a mailed questionnaire. After three follow-up letters were sent, information was obtained on 80 per cent of the sample families. The total sample numbered 1,603 persons, consisting of 846 parents and 757 siblings and spouses of siblings. Information on date of birth, schooling, and IQ of the subject was obtained from the files of the Institute. Occupation of father, subject's occupation, information on the spouse of the subject, number of children in the completed family,

and names and dates of birth of the children were obtained from the questionnaire. The investigator claimed that his sample was representative of the Minnesota population. but given the unusual sampling procedures employed. the degree of representativeness must be viewed as uncertain.

Using the data from these sources. Waller computed a number of correlation coefficients. Those most relevant here are the coefficients for subject's IQ with size of family of origin and size of completed family. The results of the analysis for males revealed the expected significant negative correlation between IQ and size of family of origin. but virtually no correlation between IQ and size of family of procreation. For females. there was no correlation between IQ and size of family of origin but a significant positive correlation between IQ and size of family of procreation. Again it must be noted that two different generations are involved. and that there is no reason why the relationship between intelligence and fertility must be the same for each.

Following Bajema's lead. Waller also computed the intrinsic rate of natural increase for the 699 persons who had completed their fertility. The highest IQ group in the study had the largest rate of increase and the lowest IQ category had the smallest rate of increase. After determining relative fitness in relation to IQ for those persons, he concluded. "There is no evidence from these considerations in this sample that natural selection is acting to decrease the average IQ score, and it may in fact be effecting a very slow increase." Thus, the results of the Waller study support those of Higgins and associates and of Bajema.

It must be noted that although these three studies used more sophisticate analytic techniques and made less naiv assumptions than the earlier studies their findings may be faulted on a num ber of grounds, particularly selectio bias and sample representativeness Also, the fact that IQ scores probabl' came from a number of different test given at varying ages invites doubt abou the comparability of test scores. Bu even in these respects. their method were superior to those used in most o the earlier studies that came to opposit conclusions.

In a more recent study. Olneck an Wolfe (1980) reported on research i which they examined the relationshir between intelligence and family size us ing longitudinal data from two samples the Kalamazoo Sample of Brothers and the National Bureau of Economic Research Thorndike-Hagan (NBER-T-H sample of white men. The Kalamazoo Brothers sample includes 1.200 mer traced and interviewed in 1973 and 197⁴ by Olneck (1976). The respondents were drawn from an original sample o¹ 3.000 males. identified as siblings, for whom sixth-grade test scores from 1928 to 1950 were available in Kalamazoo public school records. Olneck and Wolfe's Kalamazoo findings are based on 352 weighted independent pairs of brothers for whom test scores. age, and self-reported education, earnings, and marital status were available. The intelligence test scores for the period 1928–1943 are based on the Terman group test; those for siblings tested after 1943 are from the Otis test. Both tests are reported to measure similar skills and to stress verbal rather than quantitative items. The authors assert that in their data the correlations involving the two tests are quite similar and that in the literature there is no evidence that the

variances or reliabilities of the two tests differ significantly (Flemming. 1925; Cattell, 1930; Ratcliff, 1934; Buros, 1965). However, the Otis test is scaled to a lower mean than the Terman test. Therefore, Olneck (1976) adjusted the mean scores of those who had taken the Otis test and pooled the subsamples. The Kalamazoo respondents were virtually all white, Protestant, and of non-farm origins. Their educational and occupational levels are higher than for men of their age when compared with the national sample used in the Occupational Changes in a Generation replication study (Featherman and Hauser, 1977, 1978). The investigators assert that while caution must be used in viewing results from the Kalamazoo sample, there are no obvious biases which would greatly distort the findings in the present analysis (Olneck and Wolfe, 1980, p. 244).

The NBER sample comprises 5,000 white men, born between 1917 and 1925, who took the U.S. Air Corps Aviation Cadet Qualifying (ACQ) Examination in 1943 and were followed up in a 1969 NBER survey. The sample is relatively homogeneous in measured intelligence and education: all respondents had at least graduated from high school or had high school equivalency and scored at or above the median on the ACQ test. The test score used is based on a pooled composite of a battery of tests and is said to represent general intellectual ability or scholastic aptitude (Thorndike and Hagan, 1962). The measure of fertility derives from the NBER follow-up survey in 1969. It is based on the question, "How many children do you have?" Because of the phrasing of this question and other related questions concerning offspring, it is thought that the respondents reported surviving children currently at home rather than children ever born. Adopted children may also be included in the responses. Never-married persons and individuals who failed to respond to the questions on children, education, or income were excluded. The analysis is based on the 4,826 remaining cases.

The results of the analysis of the Kalamazoo and NBER samples indicate that the correlation between father's intelligence test score and number of children is positive but very small (0.06) in both samples. The coefficients of IQ in regressions of number of children on IQ are statistically significant but also very small, 0.027 for the NBER-T-H sample and 0.014 for individuals in the Kalamazoo sample. The corresponding within-pairs coefficient for the pairs of brothers in the Kalamazoo sample is 0.021. Olneck and Wolfe conclude that "our evidence, therefore, suggests that intelligence, insofar as it is measured by standardized tests, is not an important determinant of fertility." The authors further suggest that the characteristics of the samples and comparisons of their findings with those of other studies based on completed families give them no reason to believe that their results are due to sample biases.

It must be pointed out, however, that results based on a sample of brothers from Kalamazoo may have limited generalizability regardless of how carefully the analysis was done. The results from the NBER-T-H sample are more impressive but are based on a nonrepresentative sample of ever-married male high-school graduates whose measured intelligence was relatively high. Moreover, as has already been pointed out, the measurement of fertility was less than ideal for a study of this kind. Another problem is that both the Kalama-

zoo and the NBER-T-H samples suffer from being restricted to males. As we shall see later, the relationship between IQ and fertility for females may be quite different from that for males.

A more recent published study on the relation of measured parental IQ and fertility by Vining (1982) was based on data collected as part of the National Longitudinal Study of Labor Market Experience (NLS). A national probability sample of 5,172 males aged 14–24 was taken in 1966 and a second sample of 5,097 females aged 14–24 was taken in 1968. Both samples intentionally over-sampled blacks, in order to equalize to some extent sampling error in the larger white and smaller black populations; proportional weighting was used to obtain unbiased estimates for the population of both races. Resurveys have been undertaken periodically. The resurveys on which Vining's fertility estimates were based were taken in 1976 for males and 1978 for females, when the respondents were aged 25–34. The fertility measure for the males was derived from information on the relationship of each member of the household to the respondent in 1976. As Vining states, the number of own children living with their father in 1976 is not an adequate measure of biological fertility of males, although he argues that it is a good proxy for fertility of white but not black men. For women, information was available on the number of children the respondent had before 1978; the woman's statement about the number of children she expects to have in her lifetime; and the number of children she considers to be ideal for a family. The intelligence variable was constructed by members of the NLS staff from "test scores" obtained from the last high school attended. The scores came from

approximately 30 different measures that were subsequently pooled by transforming the scores of each test to a mean of 100 and a standard deviation of 16. Because test scores were sought only for students who had completed the ninth grade, those aged 14 at the time of the first survey round were excluded from the analysis. Thus, the study was confined to those who had completed ninth grade and who were aged 15–24 at the time of the initial survey and 25–34 at the time of the last survey. The original sample of males contained 3,275 whites and 1,233 blacks; information was available on both test scores and fertility for 61.2 per cent of whites but only 26.8 per cent of blacks. For females, this information was available for 61.1 per cent of whites and 35.4 per cent of blacks.

Mean cumulative fertility rates were computed for six IQ categories separately for white women, black women, and white men. The results showed that the mean cumulative fertility rate was highest for both white females and males in the middle IQ categories, and lowest in the bottom and top IQ categories. The sharpest drop-off was in the highest IQ category. For black women, there was a monotonic decrease in mean number of children from the lowest to the highest IQ category. For both white and black women, the overall relationship between IQ and number of children was negative. From these data Vining calculated the IQ selection differential (Falconer, 1966) to be "around 2 points" decline in IQ per generation for each of the three race and gender subgroups. After adjusting for the regression of child's IQ on mid-parent's IQ (taken to be 0.5), he concluded that the generational change in mean genotypic IQ due to differential fertility by IQ was

about one IQ point decline per generation.

Vining then regressed the number of children ever born on the parent's IQ score for the 5-year age cohorts 25–29 and 30–34 and for the 10-year age cohort. He found very small negative slope coefficients that increased slightly (i.e., became more negative) across cohorts for white women aged 25–29, from -0.014 to -0.017. He also found that the coefficients of white women in the older cohort showed a decline over the previous five years from -0.014 to -0.012. Despite the fact that the differences in coefficients were negligible and not statistically significant, he concluded that for white women, "there is evidence of a decline in the severity of the dysgenic trend in fertility differentials with respect to intelligence, both across cohorts with respect to age and within cohorts as they move through the childbearing ages" (p. 248). (The conclusion that severity decreases across cohorts is erroneous and seems to reflect a misinterpretation of the above-mentioned change in regression coefficients from -0.014 to -0.017.) He argues that the within-cohort negative relationship between IQ and fertility will not be reduced to zero by the time the sample of white women has completed its fertility. On the other hand, he found no evidence of a decline in the negative relationship between IQ and fertility for black women, either with respect to age or across time. The data for white men are not complete enough to allow these kinds of comparisons.

He also regressed mean expected fertility, as stated by the women, on IQ and found for white women a "substantially weaker" negative relationship with IQ for whites than had been found for completed fertility to date and a "moderately weaker" relationship in the case of blacks. This was taken as evidence that higher IQ white women, in particular, intend to have substantially more children in their later years of childbearing than lower IQ women in this same period. For black women, this catch-up effect was weaker.

Finally, combining the white and black women's samples, he predicts a decline in mean genotypic IQ of one-and-a-third IQ points per generation if current fertility patterns persist, and a decline of about three-fifths of an IQ point if current fertility expectations are realized (calculated by multiplying the IQ selection differentials for all women aged 25–34 in Vining's Table 4 by 0.5).

Vining suggests that previous findings of a positive association between parental intelligence and number of children reported by Higgins, et al. (1962), Bajema (1963), Waller (1971), and Olneck and Wolfe (1980) may be faulted because they are not based on nationally representative samples but rather on the white population of the Great Lakes states. Furthermore, he suggests that the cohorts studied were reproductively active during the period of rising birth rates of the post-war baby boom, which was possibly the only period in modern times in which there may have been a eugenic trend in the population. On the other hand, he claims that because his data are from a national sample and because his cohorts were reproductively active after the baby boom, they reflect the dysgenic trend that he asserts has probably characterized the normal situation over a long period of time and is likely to represent the future situation.

Vining's study suffers from at least five major weaknesses. First, there was considerable sample attrition, due ei-

ther to failure to reinterview the original respondents or more importantly to obtain measures of intelligence for a large number of those who were reinterviewed. The losses from these and other sources were approximately 39 per cent for both white women and white men, 65 per cent for black women, and 73 per cent for black men. From what we know about the effects of sampling losses of these magnitudes, one can have little confidence that parameter estimates based on the remaining cases truly represent the population originally sampled. Vining notes that only in the case of black males do the cumulative age-specific birth rates deviate significantly from those of the comparable subpopulations in the nation as a whole. It is possible, however, that the pattern of fertility by IQ is biased more than the estimates of overall fertility for all IQ's combined. In this regard, evidence is needed to throw light on the question of whether the measured intelligence of those for whom test scores were obtained is representative of the intelligence of the original sample. No attempt was made to answer this question, nor was any systematic attempt made to identify the selective factors that account for the losses. Quite plausibly, one of these selective factors is lower average intelligence among the losses.

Second, the IQ scores used in this study suffer from serious psychometric weaknesses. They were obtained by pooling information from 30 different tests and measures, including a number of standardized intelligence tests with varying item content and structure and with differing validity and reliability coefficients. Also included are scores on college entrance tests, such as the ACT, SAT, and PSAT, none of which was de-signed to measure intelligence. When test scores were lacking, grade point averages in high school were substituted, despite the fact that the correlation between grade point average and intelligence test score is approximately 0.5, indicating that grade point averages can account for only about 25 per cent of the variance in intelligence test scores. No attempt was made to determine the equivalence of scores from the various intelligence tests before combining the scores, even though the techniques for doing so and the dangers in not doing so have been well known for many years (Gulliksen, 1950; Angoff, 1971). Merely norming the various scores to a mean of 100 and a standard deviation of 16 does not address the basic measurement questions; it only converts a melange of scores of unknown equivalence into a common metric. Thus, we conclude that the measurement of IQ in this study is very questionable.

Third, there are serious questions about the measurement of fertility in this study. As Vining has pointed out, the questions on which the data on the fertility of males and females are based are quite different and provide less adequate information on men than on women. Another problem is that the younger women in the sample could not have had more than 10 years and the older women more than 20 years in which to have had children, out of a potential period of some 35 years. Thus, estimates of life-time differences in fertility have varying degrees of reliability, especially when they are made for various IQ categories. This is particularly true because both men and women in higher intelligence categories tend to delay marriage for higher education and thus to have their children later than

those in the lower intelligence categories. Consequently, the projected figures based on fertility over shorter periods may seriously understate the eventual fertility of those in the higher IQ categories.

Fourth, Vining's sample excludes persons who did not complete ninth grade. Unpublished tabulations from one-in-a-thousand public use sample tapes from the U.S. censuses of 1960 and 1970 indicate that, of persons aged 15–24 at the time of the 1960 and 1970 censuses, 17 per cent in 1960 and 12 per cent in 1970 had not completed the ninth grade. Linear interpolation of these figures yields an estimate of 15 per cent for 1966. (By race, the estimates are 14 per cent for whites and 26 per cent for blacks in 1966.) Thus, Vining's estimates exclude a substantial portion of the population that probably has an average IQ well below that of the population as a whole. Depending on the fertility of this excluded group, their omission from the sample may seriously bias Vining's estimates of the generational change in intelligence.

Fifth, Vining did not take differential mortality by IQ into account in his calculation of the IQ selection differential.

Because of these deficiencies, one must raise serious questions about Vining's conclusions, particularly those having to do with estimated generational changes in intelligence.

Another recent study by Van Court and Bean (1985), based on retrospective rather than prospective data, examines the trend in the relation of measured intelligence and fertility by correlating IQ test scores and the number of children born in adult cohorts who were respondents to a series of National Opinion Research Center (NORC) general social surveys conducted from 1974 through 1982. The IQ test scores were based on a vocabulary test consisting of ten questions selected from a longer test devised by Thorndike (1942). Low, but statistically significant, negative correlations (ranging from −0.05 to −0.23) were found between the test scores and the number of children ever born for six of the nine five-year age cohorts whose fertility had been completed by the time of the survey. (The other three correlations were also negative but not statistically significant. Correlations for later cohorts were more strongly negative, but the fertility of these cohorts was incomplete when measured, so that the correlations may be biased in the negative direction, for reasons discussed earlier.) The authors did not publish correlations between fertility and IQ for each sex separately. On the basis of their findings, Van Court and Bean concluded that the relationship between intelligence and fertility was predominantly negative from 1912 to 1982. They believe that previous reports of a neutral or slightly positive relationship may be due to the fact that they involved unrepresentative samples and did not include nonwhites.

More confidence could be placed in Van Court and Bean's results had they reported response rates to the five separate surveys and particularly to the IQ test and the fertility questions. Also, one would like to know the validity and reliability correlations of the specific test used in the NORC survey, rather than the evidence the authors provide for other short tests of intelligence. Finally, although the surveys may have been representative of the U.S. population at the time they were taken, the survivors in the older age cohorts are not likely to

be representative of their original age cohort; for example, those of higher IQ may be overrepresented because of their lower mortality. There is no discussion of these points in their paper.

In a subsequent article, Vining (1986) reviews further evidence on trends in differential fertility by intelligence, mainly for the United States. The evidence consists of comparisons between the fertility trends of groups known or presumed to have high intelligence and the fertility trends of the nation as a whole. The groups with high intelligence include samples from *Who's Who*, the Forbes 400, Terman's high IQ sample, and the American Mensa. Much of this evidence is weakened by the fact that the fertility of men in the select groups is usually compared with the fertility of women in the nation as a whole. The difficulty here, as we shall see later in this paper, is that the pattern of differential fertility by IQ may be quite different between women and men. Therefore, the fertility of women ideally should be compared with the fertility of women, and the fertility of men should be compared with the fertility of men. Another problem is the possibility that small elite groups of very high intelligence may be atypical of the broader group of intelligent persons. Despite these difficulties and ambiguities, it must be noted that these data consistently confirm the pattern noted by Van Court and Bean and by Vining in his earlier article, that negative differential fertility by IQ tended to converge during the baby boom and diverge again in the negative direction after the baby boom.

Two further studies are worthy of note. The first of these shows large gains in the mean IQ of Americans between 1932 and 1978, based on results from IQ standardization samples used for norm-ing a series of major IQ tests in the United States (Flynn, 1984). The second study confirms similar gains in 13 other countries as well (Flynn, 1987). The studies show that the increasing trend of IQ first noted in the Scottish study discussed earlier (Scottish Council for Research in Education, 1933, 1949) and by Cattell (1951) is widespread and has continued. Together, these studies suggest that measured IQ is influenced considerably by environmental factors, most notably improving trends in education.

From this brief review of studies reported after the appearance of Anastasi's review paper in 1956, one must conclude that research to date on the relationship between intelligence and fertility can be faulted on a number of methodological grounds, including sampling, measurement of variables, and in some cases less than adequate statistical techniques. None of these studies meets the requirements for an adequate study of the relationship between intelligence and fertility. Because of the shortcomings of the studies reviewed, we are convinced that any generalizations about the trend in intelligence of the U.S. population as a whole, derived from these studies, must be viewed as highly tentative at best.

One may question the need for still another paper on the topic of intelligence and fertility that does not fulfill all of the requirements for a definitive study. But our research, although not ideal in every regard, does have better data on the critical variables than those available to the authors of the earlier studies covered in this review. Our sample, although not designed to be nationally representative, is a large random sample drawn from the state of Wisconsin, which ranks near the middle of the

states in many socioeconomic characteristics and has great diversity in its population. Our research, too, has limitations, which we shall clarify in the course of presenting our analysis.

MATERIALS AND METHODS

THE WISCONSIN LONGITUDINAL DATA

The data on which our investigation is based come from the Wisconsin Longitudinal Study of Social and Psychological Factors in Educational and Occupational Aspirations and Achievements, referred to throughout this paper as the WLS. The WLS is based on a random sample of 10,317 high-school seniors in public, private, and parochial schools in 1957. The questionnaires used in this study covered the students' socioeconomic backgrounds, high school experiences, educational and occupational aspirations and plans, and perceptions of the influence of parents, teachers, and friends on their plans and aspirations. From the schools and public sources, additional information was obtained on parents' economic status, the students' measured intelligence and high-school grades, and the characteristics of their high schools and communities. Race was not asked in the WLS, but unpublished tabulations from a 1/1000 Public Use Sample tape for Wisconsin from the 1960 census indicates that, of persons aged 15–24 with 12 or more completed years of education, 98.2 per cent were white, 0.9 per cent black, and 0.9 per cent other races. (The age group 15–24 is centered on age 20, which was approximately the average age of the WLS cohort in 1960.) WLS follow-up surveys of the students were conducted in 1964 and 1975 with response rates of approximately 90 per cent. For a detailed discussion of the 1957, 1964, and 1975 samples, nonresponse bias, and data quality, see Sewell and Hauser (1975, 1980) and Clarridge et al. (1977).

The 1975 follow-up study is especially pertinent to the present research because, in addition to detailed information on educational and occupational experiences and attainments, considerable information was gathered in the interview on family formation and dissolution and fertility history up to the time of the interview. For a more complete discussion of these data and a review of results of the many articles and monographs based on them, see Sewell and Hauser (1977, 1980); Sewell, Hauser, and Wolf (1980): Hauser, Sewell, and Clarridge (1982): and Hauser, Tsai, and Sewell (1983).

THE WLS MEASURES OF INTELLIGENCE AND FERTILITY

The two main variables used in our research are measured intelligence and fertility. Because the measurement of these variables is crucial to any conclusions drawn from the study, their sources, characteristics, and limitations must be carefully examined.

The mental ability measurements are based on scores on the Henmon-Nelson Test of Mental Ability (revised 1954), which was administered to Wisconsin ninth- and eleventh-grade students in public, private, and parochial schools in the spring of 1957. The testing was done under the supervision of the Wisconsin Testing Service and machine scored at its headquarters. The raw scores, percentile ranks, and IQ scores for each student were recorded alphabetically by name of student for each high school in the state. Test scores were available from this source for all students in the WLS. (In a relatively small but un-

known number of cases where students had not taken the Henmon-Nelson test—usually in cases where they had transferred from out of state into a Wisconsin high school for their senior year—psychometrically equivalent scores on other standardized tests were substituted for Henmon-Nelson Test scores.) The Henmon-Nelson IQ scores are normed to a mean of 100 and a standard deviation of 15.

The Henmon-Nelson test, which at the time was one of the most widely used group tests of intelligence, was designed to measure the mental ability of students in grades 7 to 12 inclusive. In the words of the tests's authors (Henmon and Nelson, 1954, p. 3), "Tests of this kind have often been called 'tests of intelligence' and perhaps that term is as appropriate as the term 'mental ability.' Actually both terms leave something to be desired since such tests do not measure all of 'intelligence'; neither do they measure all kinds of 'mental ability.' Rather they measure the kinds of ability necessary for academic work and it is for this purpose that they will be found most useful."

There are three equivalent forms of the test, all identical in difficulty and construction. Each form of the test consists of 90 items arranged in a "spiral omnibus" pattern in which item types are mixed together with a steady rise in the difficulty of the items. The item types include geometric analogies, arithmetic problems, proverb interpretations, general information, synonyms, and antonyms. The test is "speeded" in that respondents are given 30 minutes to complete it. Thus, scores depend in part on reading skills. Validity coefficients, reported by Henmon and Nelson (1954, p. 6), based on the correlation of scores on their test with scores on other well-

known group tests of mental ability, including the Otis Self-Administering Test, the Terman Group Test, the Kuhlmann-Anderson Test, and the American Council Psychological Examination, are in the range of 0.780 to 0.858. Also, the means and standard deviations of the Henmon-Nelson Test and those of other well-known tests, including the tests mentioned above, are quite similar. All available evidence indicates that the Henmon-Nelson Test is highly reliable in the large and diverse Wisconsin sample. For example, Henmon and Nelson report a split-half reliability of $r = 0.887$ and an alternative forms reliability of 0.89 for two small samples of eleventh-grade students. Hauser et al. (1983, p. 40), taking into account the variance in IQ in the WLS, estimate these reliability coefficients to be 0.916 and 0.953 respectively. We have computed a test-retest correlation of 0.814 from data reported in Hauser and Sewell (1984) for a random sample of 336 Wisconsin men who took the Hemon-Nelson test in both their freshman and junior years.

The basic measure of fertility used in this study is the number of children ever born alive to the respondent. This is measured in the same way for males and females. It is derived from a complete roster of children born to ever-married female respondents or fathered by ever-married male respondents and includes the dates of all live births (including those born out of wedlock), the age of the respondent in century months (months elapsed since January 1, 1900) at the time of the birth of each child, and the date of each marriage and each marital dissolution. Never-married persons (i.e., never-married by age 35) were assigned a fertility of zero. From this information a number of measures of fertility

may be computed. The primary measure used in this study is the number of children born alive to the respondent before the 35th birthday, which in most cases is the same as the number of children ever born up to the date of the interview. This measure of necessity does not represent the ultimate fertility of our sample members, because the women in our sample still had, on average, at least five or six more years during which they could bear more children. For the men the possible period is of course even longer and depends largely on the age of their mates. What we know with reasonable certainty is how many children our respondents had up to approximately age 36. Consequently, to the extent that we wish to talk about completed fertility, it is necessary to produce estimates based on the 1975 data and other known characteristics of the sample and of the population it represents. These techniques, their assumptions, and their possible errors and biases will be discussed at the point of which they are used in the analysis.

It should be pointed out that the sample on which our analysis is based, although randomly drawn from the large and heterogeneous population of Wisconsin high school graduates, is not a complete age cohort because it excludes persons of the same age who did not graduate high school. We estimate that approximately 25 per cent of the complete age cohort did not graduate (Hauser, 1971; Sewell and Hauser, 1975, Appendix D). As a consequence, when we attempt to extend our findings to cover all persons in the age group, we must rely additionally on estimates of the measured intelligence, fertility, and mortality of the dropouts. Unfortunately, the available information on which to base these estimates is not nearly as adequate as that for the WLS sample members.

Differential Reproduction by IQ in the Wisconsin Longitudinal Study

As already mentioned, our basic fertility variable is the number of children ever born alive to the respondent before the 35th birthday, denoted as F(35). A minor problem with this variable is that not all respondents reached age 35 by the time of the 1975 survey. There were 28 males and 34 females who did not reach 35 by the time of the survey, and for some of these individuals F(35) may be underestimated by one child or possibly even two children in the unlikely event of multiple births. But the number of such individuals is so small in relation to overall sample size that the bias introduced into our aggregate-level estimates of F(35) is negligible. The 62 individuals who did not reach 35 by the time of the survey were retained in the analysis.

Another problem is that fertility information was not collected from never-married persons, who constituted 5.0 per cent of women and 5.7 per cent of men in the 1975 reinterview survey. As mentioned earlier, we assigned never-married persons a fertility of zero. It is thought that very few persons in this cohort who did not marry by age 35 had illegitimate children, so that the bias introduced by the zero-fertility assumption for never-married persons is probably very small, too. (In this regard, it may be noted that in the Wisconsin sample, average intelligence is slightly higher for never-married persons than for ever-married persons.)

In examining the relationship between fertility and IQ, we have considered F(35) as the dependent variable and IQ as the independent variable, as

shown in Table 1. To tabulate F(35) by IQ, it is convenient to group respondents into IQ deciles. The decile breaks, shown in the second column of Table 1, were obtained from the 1957 survey, using data on all the original respondents. The numbers of respondents in each decile are not precisely equal, because many respondents have the same IQ scores, and because those with the same score must be treated as a block when grouping respondents by IQ. This makes it impossible to define IQ intervals with precisely the same number of respondents in each interval. Although the IQ decile breaks are derived from the 1957 survey, the sample N's (number of respondents) in Table 1 are based on the 1975 survey, in which the fertility questions were asked. Because of sample attrition from deaths and nonresponse between the two surveys, the sample N's in Table 1 are smaller than

the sample N's in the 1957 survey. Table 2, discussed later, provides the N's for each decile for the original respondents in the 1957 survey.

Table 1 shows values of F(35) by IQ decile and sex, and these values are graphed in Figure 1. Figure 1 also shows values of the total fertility rate and the intrinsic rate of natural increase, which will be discussed in more detail later and can be ignored for the moment. The table and figure show striking differences by sex. The relationship between IQ and fertility is negative for both sexes, but much more so for females than for males. For females, the highest fertility is achieved in the second IQ decile, and fertility falls fairly regularly at higher deciles. The pattern for males is more erratic. Overall, the relationship between IQ and fertility for males is only slightly negative. Fertility is highest in the fourth decile, with the third and sec-

TABLE 1

MEAN NUMBER OF CHILDREN EVER BORN BY 35TH BIRTHDAY, $F(35)$,
BY IQ DECILE AND SEX, DERIVED FROM THE WLS[a]

IQ DECILE	IQ RANGE	FEMALES			MALES		
		F(35)	SE	N	F(35)	SE	N
1	67–81	2.76	0.065	436	2.36	0.058	418
2	82–87	2.92	0.066	438	2.45	0.064	365
3	88–92	2.83	0.062	509	2.46	0.057	437
4	93–96	2.81	0.060	444	2.50	0.060	407
5	97–100	2.73	0.057	515	2.29	0.057	428
6	101–103	2.70	0.059	501	2.27	0.062	411
7	104–108	2.70	0.066	434	2.26	0.057	438
8	109–112	2.55	0.057	542	2.19	0.052	454
9	113–120	2.50	0.056	513	2.37	0.054	488
10	121–145	2.29	0.057	456	2.07	0.057	470

[a]WLS denotes the Wisconsin Longitudinal Study. SE denotes the standard error of the estimate of $F(35)$. The standard errors include a finite population correction (fpc) of 0.82, based on a sampling fraction of 1/3 (Cochran, 1963). An F-ratio test of the 10 sample means of $F(35)$ indicates significant differences among IQ categories; $p < 0.01$ for females and $p < 0.05$ for males.

    Although there is good reason to believe that the relationship between IQ and fertility is intrinsically nonlinear, since fertility tends to drop off at both the upper and lower ends of the distribution of intelligence, one may nevertheless force a linear fit for the purpose of further significance testing. This is done by regressing $F(35)$ on IQ for each sex separately from individual-level data. Results yield coefficients of IQ of -0.00580 (SE $= 0.00148$, one-sided $p < 0.0002$) for males and -0.01057 (SE $= 0.00162$, one-sided $p < 10^{-6}$) for females. Both coefficients are highly significant, as is the difference between the male and female coefficients (one-sided $p = 0.0143$).

    The F-ratio test and the significance tests for regression coefficients make no allowance for finite population corrections.
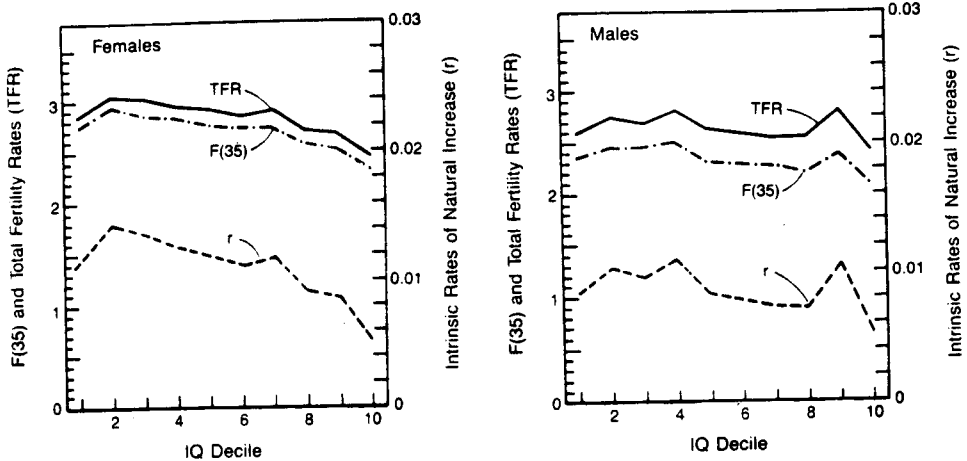
FIG. 1.—$F(35)$, total fertility rates (TFR), and intrinsic rates of natural increase (r) by IQ decile and sex in the WLS.

ond deciles close behind. Fertility declines at higher deciles, except for the ninth decile, which has fertility not much lower than the peak fertility of the fourth decile. For both females and males, fertility is lowest in the tenth or highest decile. The observed differences in fertility by IQ are statistically significant, as discussed in the footnote to Table 1.

Note that there is no decile for either sex where $F(35)$ drops below two children per woman. Thus, fertility is above replacement. For females, the mean number of children by the 35th birthday ranges from 2.29 in the tenth decile to 2.93 in the second decile, with an average of 2.68. For males, it varies from 2.07 in the tenth decile to 2.50 in the fourth decile, with an average of 2.32. The overall fertility of males is lower than that of females partly because, in the WLS sample, men tend to marry about two years later than women, so that their fertility by age 35 is less complete.

Further measures of differential reproduction by IQ decile and sex are presented in Table 2. The first measure of interest is the total fertility rate (TFR) for the WLS cohort. The total fertility rate is defined as the expected number of live births of both sexes that a woman or man would have if she or he experienced observed age-specific birth rates under hypothetical conditions of zero mortality. In the present context it is calculated by summing age-specific birth rates in 5-year age groups between ages 15 and 50 and multiplying the sum by 5.

Estimation of the TFR requires extrapolation of the fertility of WLS respondents to ages beyond 35. Our extrapolation procedure required estimates of age-specific birth rates for age groups 15–19 through 30–34, which were reconstructed without difficulty from the birth histories of the WLS respondents. Age-specific birth rates in 5-year age groups for women were calculated as the average number of births per woman per year over the five years of each age group, and age-specific birth rates for men were calculated as the average number of births per man per year over the five years of each age group.

TABLE 2
MEASURES OF REPRODUCTIVITY AND SELECTION BY IQ DECILES AND SEX, DERIVED FROM THE WLS[a]

SEX AND MEASURE	1	2	3	4	5	6	7	8	9	10	Total
Females											
IQ range	67–81	82–87	88–92	93–96	97–100	101–103	104–108	109–112	113–120	121–145	. . .
Mean IQ	75.30	84.75	90.22	94.44	98.06	102.01	105.55	110.11	116.25	126.68	100.43
$F(35)$	2.75	2.93	2.84	2.81	2.73	2.71	2.71	2.55	2.50	2.29	2.68
TFR	2.85	3.04	3.02	2.93	2.89	2.83	2.88	2.68	2.65	2.42	2.82
GRR	1.39	1.48	1.47	1.43	1.41	1.38	1.40	1.31	1.29	1.18	1.37
e_0	75.00	75.25	75.25	75.25	75.50	75.50	75.50	75.50	75.75	75.75	75.50
NRR	1.36	1.45	1.44	1.39	1.38	1.35	1.37	1.28	1.26	1.16	1.34
r	0.01198	0.01475	0.01403	0.01302	0.01232	0.01157	0.01215	0.00951	0.00901	0.00551	0.01140
T	25.47	25.17	25.79	25.51	26.00	25.81	26.02	25.99	25.98	26.35	25.82
Doubling time (yrs)	57.85	47.00	49.39	53.23	56.26	59.90	57.04	72.87	76.96	125.91	60.81
W	0.931	1.000	0.982	0.956	0.939	0.921	0.935	0.873	0.862	0.788	0.917
N (No. of respondents)	506	505	562	510	569	546	494	583	553	495	5,323
Males											
IQ range	67–81	82–87	88–92	93–96	97–100	101–103	104–108	109–112	113–120	121–145	. . .
Mean IQ	74.45	84.86	90.17	94.50	97.99	101.99	105.56	110.08	116.27	127.18	100.78
$F(35)$	2.36	2.45	2.46	2.50	2.29	2.27	2.26	2.19	2.37	2.07	2.32
TFR	2.60	2.75	2.68	2.80	2.61	2.57	2.53	2.54	2.80	2.39	2.62
GRR	1.33	1.41	1.37	1.43	1.34	1.31	1.29	1.30	1.44	1.22	1.34
e_0	68.30	68.82	69.08	69.34	69.34	69.60	69.86	70.13	70.39	70.92	69.60
NRR	1.25	1.32	1.30	1.35	1.26	1.24	1.23	1.23	1.36	1.17	1.27
r	0.00807	0.01013	0.00946	0.01090	0.00822	0.00777	0.00733	0.00735	0.01073	0.00544	0.00850
T	27.61	27.68	27.34	27.80	28.30	28.10	27.98	28.53	29.03	28.70	28.13
Doubling time (yrs)	85.91	68.40	73.26	63.58	84.30	89.20	94.58	94.25	64.59	127.34	81.55
W	0.923	0.979	0.960	1.000	0.927	0.916	0.904	0.905	0.995	0.858	0.935
N (No. of respondents)	525	427	498	467	501	472	501	522	559	522	4,994

[a] $F(35)$ is defined in Table 1. TFR denotes total fertility rate, defined as the expected number of live births of both sexes that a woman or man would have if she or he experienced observed age-specific birth rates under hypothetical conditions of zero mortality. GRR denotes gross reproduction rate, defined in the same way as TFR, except that it pertains to births of just one sex (the same as that of the parent). e_0 denotes life expectancy at birth. NRR denotes net reproduction rate, defined as the expected number of live births of the same sex that a person just born would have if she or he experienced observed age-specific birth and death rates; if observed age-specific birth and death rates remain constant over time and there is no migration, the NRR is also the long-run multiplicative growth factor by which the population increases or decreases over each succeeding generation. r is the intrinsic rate of natural increase, defined as the population growth rate that would result under the same long-run assumptions about unchanging fertility and mortality and no migration. T denotes the length of a generation, defined as the time required for a population growing at intrinsic rate r to increase by a factor equal to the NRR. The doubling time is the time required for a population growing at intrinsic rate r to increase by a factor of 2. W is relative fitness, defined as the ratio of the generational growth factor for the ith IQ decile to the generational growth factor for the most rapidly growing decile, based on the intrinsic rate of natural increase for each decile separately; for females and males, respectively, $W_i^f = e^{T^f(r_i^f - r_2^f)}$ and $W_i^m = e^{T^m(r_i^m - r_4^m)}$.

Fertility after age 35 was extrapolated using a demographic model of age-specific fertility developed by Luther (1982). Luther's method fits a truncated Pearson Type III curve—the incomplete gamma function $k\,x^2\,e^{-cx}$—to observed age-specific birth rates in 5-year age groups below 35 years of age. There is a "half parameter" that utilizes information at later ages, but it assumes only one of two values, and Luther has indicated which value to use in the case of developed countries such as the United States. Luther developed and refined his model in the process of testing it with 79 observed schedules of age-specific birth rates taken mainly from a compendium of demographic data compiled by Keyfitz and Flieger (1971), and he ultimately obtained excellent fits to the test data. Details of how the model was applied to fit the WLS data are given in a pre-publication version of this paper (Retherford and Sewell, 1986).

We applied Luther's model to extrapolate fertility beyond age 35, by sex and IQ decile. The model yielded excellent fits to the WLS data (Retherford and Sewell, 1986). Errors in the TFR estimates due to fitting errors are probably very small, since less than 7 per cent of the total fertility of the WLS cohort, as determined from Luther's procedure, occurs after age 35. (Period fertility estimates from published sources during the past two decades indicate a percentage that is 1 or 2 percentage points higher than 7 per cent, but one expects period estimates to exceed cohort estimates of this percentage during times of rapidly declining period fertility. The period total fertility rate, calculated for each calendar year from age-specific birth rates for that calendar year, fell substantially in the United States between 1957 and

1975 (U.S. Department of Commerce, Bureau of the Census, 1983, p. 64).)

As shown in Figure 1, the estimated TFR's are slightly higher than the values of F(35), and they vary in much the same way by IQ decile. Note that the TFR's for men are lower than those for women. As noted earlier, this is also true for F(35). There is, of course, no reason why the TFR's for the two sexes must be the same, since the wives of the WLS men are for the most part not in the WLS sample. Moreover, the wives of the WLS men are approximately two years younger, on average, than the WLS women, since, as mentioned earlier, men tend to marry women about two years younger than themselves. Viewed from the perspective of family fertility, the completed fertility of families of WLS men therefore lags the completed fertility of families of WLS women by about two calendar years. Because of this time lag, and because fertility in Wisconsin (as in the rest of the nation) declined sharply between 1957 and 1975, one expects the TFR's for men to be lower than the TFR's for women in the WLS. It is also possible that reporting of offspring for men may be somewhat less than complete in the birth histories, especially births illegitimately fathered who have never lived with the father.

It is a simple matter to calculate gross reproduction rates (GRR) from the estimated TFR's. The GRR for women is defined as the expected number of live girl births that a woman would have if she experienced observed age-specific birth rates for women under hypothetical conditions of zero mortality. The GRR for men is similarly defined, except that it is based on boy births and the age-specific birth rates for men. The

GRR for women is estimated by multiplying the TFR for women by the ratio of female births to all births, i.e., by 0.488 under the assumption that the sex ratio at birth is 1.05 male births for every female birth. The GRR for men is estimated by multiplying the TFR for men by the ratio of male births to all births, i.e., by 0.512. Gross reproduction rates are included in Table 2 mainly for comparison with net reproduction rates (NRR), which will be discussed later.

The NRR and remaining measures in Table 2 require estimates of mortality by IQ decile. To estimate mortality we used data on sample attrition between the 1957 and 1975 WLS surveys. The 1975 survey identifies respondents who died between 1957 and 1975. Deaths include 118 men and 56 women. There were 4,994 men and 5,323 women to start with in 1957. Therefore, proportions dying between the two surveys are $118/4{,}994 = 0.02363$ for men and $56/5{,}323 = 0.01052$ for women. To a close approximation, each of these proportions equals the proportion dying between ages 18 and 36, denoted here as $q(18)$.

There are not enough deaths in the sample to allow direct estimation of mortality for each IQ decile. Moreover, the age range of the deaths, 18–36, is incomplete. Therefore, mortality by IQ decile had to be estimated indirectly. The following simple procedure was used: First, for each sex separately, we divided the sample into two halves, below and above the median IQ for the sample. The lower and upper halves had mean IQ's of 88.83 and 111.95 for women and 88.27 and 112.53 for men. Next we computed the proportion dying between the 1957 and 1975 surveys for each of the two halves, as already described. A complication was that the lower IQ half of the sample had more grade repeaters and was therefore somewhat older on average than the upper IQ half. The lower and upper halves had average ages in 1957 of 18.00 and 17.81 for women and 18.14 and 17.90 for men. Thus, the proportions dying between the two surveys are not exact estimates of $q(18)$, which refers to the proportion dying between exact ages 18 and 36. Adjustments for these age discrepancies proved to be negligible (Retherford and Sewell, 1986).

Next, the two-point formula for a straight line was used to estimate the relationship between $q(18)$ and IQ, for each sex separately, based on the values of $q(18)$ and mean IQ for the upper and lower IQ halves of the sample. The equation for females is

$$q(18) = 0.01360 - 0.00003 \, IQ \qquad (1)$$

and the equation for males is

$$q(18) = 0.03873 - 0.00015 \, IQ. \qquad (2)$$

Estimated standard errors of the IQ coefficients in these equations are 0.00010 for females and 0.00015 for males. These estimates assume that the IQ difference between the upper and lower IQ halves is constant over repeated samples, and they include a finite population correction of 0.82 (see note to Table 1). The standard error of the difference between the male and female IQ coefficients is 0.00018. Comparison of these standard errors with the coefficients themselves indicates that levels of statistical significance of the coefficients and of the male-female difference between them are low. Our estimates of mortality differences by IQ and by sex are accordingly somewhat imprecise.

From Equations (1) and (2) and the values of mean IQ for each IQ decile, values of $q(18)$ were generated for each IQ decile, for each sex separately. The values of $q(18)$ were then matched to Coale-Demeny West model life tables (Coale and Demeny, 1966) to obtain estimates of complete model life tables, in which the tables for males and females are presented at intervals of 0.25 year of female life expectancy. In each case the life table chosen was the one with a value of $q(18)$ closest to the observed value, as estimated from Equation (1) or (2). Thus, the values of life expectancy in Table 2 are accurate only to within 0.25 year, sufficiently accurate for our purposes. The estimated mortality differentials by IQ are small, especially for females. The difference in life expectancy between the first and tenth IQ deciles is about 0.8 year for females and 2.6 years for males.

Given age-specific birth rates and a life table for each IQ decile, it is a simple matter to calculate the net reproduction rate. For each sex separately, the NRR is defined as the expected number of live births of the same sex that a person just born would have if she or he experienced observed age-specific birth and death rates over a lifetime. If observed age-specific birth and death rates remain constant over time and there is no migration, the NRR is also the long-run multiplicative growth factor by which the population increases or decreases over each succeeding generation. The NRR is similar to the GRR, except that age-specific mortality is taken into account in the calculations. Thus, the NRR is somewhat smaller than the GRR. Given the negative association of mortality with IQ, the relationship between the NRR and IQ is less negative

than the relationship between the GRR and IQ. But because mortality is so low in this population, this difference is small.

From the age-specific birth rates and the life table for each IQ decile, one may also calculate the intrinsic rate of natural increase, r (not to be confused with the product-moment coefficient of correlation). This is the population growth rate that results hypothetically in the long run if observed age-specific birth and death rates remain constant over time and the population is closed to migration. A property of r is that it does not depend on the age distribution of the original population as long as some persons in the original population are of reproductive age. The variation in r with IQ resembles the variation of fertility with IQ, as shown in Figure 1. Values of r by sex for the whole sample are 0.01140 for females and 0.00850 for males (calculated from individual-level data, not by averaging decile values). The difference between the sexes arises because of the one-sex demographic model that underlies the calculations. In real populations, of course, the two sexes have the same long-run population growth rate. Thus, the one-sex calculations are not entirely consistent.

From r, one can calculate the length of a generation, T. Once the growth rate converges to r under hypothetical conditions of constant age-specific birth and death rates, the population becomes stable. A constant growth rate, r, is one of the attributes of a stable population. The growth factor over a time period of length t is then e^{rt}. The generation length T is defined as the time required for the population to increase by a factor equal to the NRR under hypothetical conditions of stability. Thus, T is ob-

tained by solving the equation $e^{rT} = $ NRR, so that $T = (1/r)\ln$ NRR. T tends to increase slightly with IQ, but the variation is small and erratic. For the entire sample, T is 25.8 years for females and 28.1 years for males.

One may also calculate population doubling times from r, and these are also shown in Table 2. The doubling time is calculated by solving the equation $e^{rt} = 2$ for t, so that $t_d = (1/r)\ln 2$. The doubling time varies inversely with r.

Following Bajema (1963), we may compute another measure of interest, relative fitness, W, which compares growth over a generation in the various IQ deciles. A difficulty here is that the generation length, T, varies from one decile to the next. This difficulty is overcome by arbitrarily taking T to be the generation length over all deciles combined. Again, each sex is considered separately. Let $r*$ denote the value of r for the most rapidly growing IQ decile, which in Table 2 is the second decile for females and the fourth decile for males. From the standpoint of reproduction, the second decile for females, or the fourth decile for males, is the most "fit," and its growth rate is referred to in the literature (somewhat inappropriately in this case) as the "optimal" growth rate. Decile-specific indices of relative fitness, W_i, are then defined as $e^{r_i T}/e^{r*T} = e^{[T(r_i-r*)]}$, which is the ratio of the growth factor for the ith decile to the growth factor for the most rapidly growing decile. An overall index of mean relative fitness, W, is defined as $e^{[T(r-r*)]}$. For both females and males, it is evident that the tenth IQ decile, with IQ 121 and over, is the least "fit" in the Darwinian sense, with values of W of 0.788 and 0.858 for females and males respectively. Again, the one-sex model yields results that are in some respects incon-

sistent and unrealistic, but they do provide useful indices of sex differences.

THE IQ SELECTION DIFFERENTIAL FOR GRADUATES

One of the objectives of our research was to obtain an estimate of the IQ selection differential from the WLS data. The IQ selection differential, S, is defined as what the generational change in mean IQ would be if, hypothetically, each child in the birth histories had the same IQ as the mean of its parents' IQ's. Because we know the IQ of only one parent, we estimate S first for each sex separately and then average male and female values. (For further discussion of the general concept of the selection differential for any trait, see Falconer, 1966, 1981, or Crow, 1986.) In the calculation of S, the assumption of equality of IQ between child and parent is introduced because we have no information on the IQ's of respondents' children. This assumption is, of course, unrealistic, a point to which we shall return later. The calculation of the IQ selection differential is done first just for graduates in the WLS sample. Later it is redone with school dropouts added in. The data for dropouts are much weaker than the data for WLS graduates, and that is why graduates are first treated separately. The separate treatment of graduates in calculating the IQ selection differential is, of course, quite artificial, since they are by no means a breeding population that is separate from dropouts.

Our method of calculating the IQ selection differential makes use of the intrinsic rate of natural increase. As already mentioned, calculations are done separately for each sex. Let IQ_i denote the mean IQ of the ith IQ decile. If N_i is the population size of the ith IQ decile to

begin with, then, if each decile grows at the intrinsic rate r_i in Table 2, the population size of each decile at the end of one generation is $K_i = N_i e^{r_i T}$. Then the IQ selection differential is calculated as

$$S = (\Sigma K_i I Q_i) \ (\Sigma K_i)$$
$$- (\Sigma N_i I Q_i)/(\Sigma N_i) \qquad (3)$$

where the summation is across IQ deciles. The first of the two terms on the right is the estimated mean IQ of the offspring generation, and the second is the estimated mean IQ of the respondent generation. Note that the first and second terms on the right are the same, except for the factors $e^{r_i T}$ embodied in $K_i = N_i e^{r_i T}$. The factors $e^{r_i T}$ differ only slightly from the net reproduction rates $NRR_i = e^{r_i T_i}$. The difference is that T_i in the formula for NRR_i is replaced by T for the total sample (calculated as the average of the male T and the female T in the total column of Table 2), so that the length of generation is the same regardless of the category under consideration.[1]

Table 3 shows results. Our base estimates of the IQ selection differential, which are shown in the first row, are −0.79 for females and −0.25 for males. The selection differential is small, less than one IQ point decline in a generation, and about three times as rapid for females as for males. Of course, these are one-sex calculations, except that a common value of T is used. Given sexual reproduction, one expects that in reality the IQ selection differential would be about the same for either sex, except

for aspects of intelligence that are sex-linked. The IQ selection differential for both sexes combined is calculated as the simple average of the female and male values—about one-half of an IQ point decline in a generation, as shown in the table. The purpose of doing the calculation for each sex separately is not only to circumvent the difficulty that we know IQ for only one parent, but also to get some sense of the relative importance of female differential fertility by IQ and male differential fertility by IQ to the overall IQ selection differential.

The entire sequence of calculations to this point can be based alternatively on IQ quartiles. This was done in order to assess roughly the impact of alternative groupings of respondent IQ's. The results of the calculation based on quartiles, shown in the second row of Table 3, are virtually identical to those based on deciles. Thus, the estimates of the IQ selection differential do not appear to be sensitive to the way in which the data are grouped.

We also tested to see how much the results for males depend on how Luther's age-specific fertility model is fitted. As mentioned, Luther's model is a truncated Pearson Type III exponential curve. Luther applied it to age-specific fertility of women and truncated the curve at age 43 or 45, depending on the level of certain parameters. This cutoff age is denoted by u in the model. For the United States, a value of 43 is found to be the appropriate value of u, and this is the value that underlies the value of the IQ selection differential of −0.25 for males in the first row of Table 3. However, this value may not be appropriate for males. The third and fourth rows show the effect of basing Luther's fit on larger values of u for males. We tried $u = 45$ and $u = 50$. The higher cutoffs had

---

[1] In the pre-publication version of this paper (Retherford and Sewell, 1986), sex-specific values of T were erroneously used in the calculation of S for each sex. This error has been corrected in the present published version, which uses the average of the male T and the female T in the calculation of S for each sex.

TABLE 3

ESTIMATES OF THE IQ SELECTION DIFFERENTIAL, S, FOR GRADUATES ONLY. DERIVED FROM THE WLS

METHOD	Selection Differential[e]		
	Females	Males	Total[f]
Based on IQ deciles[a]	−0.79	−0.25	−0.52
Based on IQ quartiles[b]	−0.77	−0.21	−0.49
Based on variants of Luther's fitting procedure[c]			
$u = 45$	. . .	−0.26	. . .
$u = 50$	. . .	−0.27	. . .
Based on hypothetical assumption of no mortality differences between IQ deciles[d]			
1960 mortality level	−0.81	−0.33	−0.57
1970 mortality level	−0.81	−0.33	−0.57

[a] This is the base estimate to which all the others are compared. Each of the other estimates differs from the base estimate in the manner indicated by the row heading. See text for explanation of how the base estimate was calculated.

[b] IQ quartile breaks are 67–90 for Quartile 1, 91–100 for Quartile 2, 101–110 for Quartile 3, and 111–145 for Quartile 4.

[c] The base estimate in the first row of the table was calculated with $u = 43$.

[d] Life tables by sex for Wisconsin for 1959–61 (whites only) and 1969–71 (all races) were used alternatively. See U.S. Department of Health, Education and Welfare (1965) and National Center for Health Statistics (1975).

[e] The IQ selection differential, S, is defined as what the generational change in mean IQ would be, if, hypothetically, each child in the birth histories had the same IQ as its parent in the WLS sample.

[f] Values in the total column are simple averages of values in the previous two columns for males and females separately.

the effect of slightly increasing the estimate of the selection differential for males, from −0.25 to −0.26 or −0.27. Hence our base estimate of −0.25. obtained by applying Luther's methodology without alteration. errs slightly on the conservative side.

We also tested for sensitivity to mortality estimation errors by redoing the calculations under the assumption of no mortality differences by IQ. We used a single life table for all IQ deciles. In the first simulation, the Wisconsin life table for 1959–61 was used. and in the second simulation, the Wisconsin life table for 1970 was used. Since mortality is inversely associated with IQ, the artificial suppression of mortality differences increases the IQ selection differential. But this effect is very small for females. for whom S increases from −0.79 to −0.81. The effect is larger for males, for whom S increases from −0.25 to −0.33. Whether the 1960 or 1970 life table is used makes virtually no difference in the results. The results for males are more

sensitive to elimination of mortality differentials by IQ than the results for females. mainly because mortality differences by IQ are larger for males than for females.

One can also assess the accuracy of the estimate of the IQ selection differential from the point of view of sampling variability and statistical precision. This cannot be done directly, because of the complicated estimation procedure. part of which involves extrapolation to older ages and the use of demographic models. For purposes of exploring statistical precision, it is useful to consider an alternative approach to calculating the IQ selection differential that could have been used had the data been more complete, as follows:

Let F denote the number of surviving offspring in a respondent's completed family. Then S may be calculated alternatively as

$$S = [\Sigma(F \, IQ)]/[\Sigma F] - [\Sigma IQ]/N \quad (4)$$

where summations range over individ-

ual adults of both sexes with the total number of adults denoted by N. As a rough approximation, we may use $F(35)$ in this formula, which is about 93 per cent of $F(50)$. When this is done, S turns out to be -0.72, compared with the value of -0.52 for both sexes combined in the first row of Table 3. The finding that S is more negative than before is expected, because differential mortality by IQ is no longer taken into account, and because the fraction of fertility occurring after 35 may be greater among more intelligent persons, who tend to start childbearing at a later age.

If one is willing to treat ΣF as a constant in Formula (4), the variance of S, $\mathrm{Var}(S)$, can be approximated without difficulty as 0.36.[2] This must be regarded as a minimum estimate, since fertility is truncated at 35, ΣF is treated as a constant, and variability in the mortality estimates is necessarily ignored. The t-ratio for S is then $-0.73/0.36 = -2.03$, yielding a one-sided p-value of 0.02. Assuming a standard error of 0.36, a value of S of 0.52, as in Table 1, yields a one-sided p-value of 0.07. This is only slightly higher than the usual 5 per cent level.

[2] The calculation formula for $\mathrm{Var}(S)$ was derived as follows: Let M denote ΣF. Then Equation (4) becomes

$$S = [\Sigma(F\,\mathrm{IQ})/M] - [\Sigma\mathrm{IQ}]/N$$
$$= \Sigma(F\,\mathrm{IQ}/M - \mathrm{IQ}/N)$$

Then

$$\mathrm{Var}(S) = \mathrm{Var}\,[\Sigma(F\,\mathrm{IQ}\,M - \mathrm{IQ}\,N)]$$
$$= \Sigma[\mathrm{Var}(F\,\mathrm{IQ}\,M - \mathrm{IQ}\,N)]$$
$$= N\,\mathrm{Var}(F\,\mathrm{IQ}\,M - \mathrm{IQ}\,N)$$

In this last formula, the expression within parentheses, which we may denote by Z, can be evaluated for each individual in the sample, with M now approximated by the sample value of ΣF. It is then a simple matter to calculate $\mathrm{Var}(Z)$ directly from the sample data. Finally, the estimate of $\mathrm{Var}(S)$ from the last formula is multiplied by a finite population correction of 0.82 (see note to Table 1).

IQ SELECTION DIFFERENTIAL FOR GRADUATES AND DROPOUTS TOGETHER

Henceforth we shall use the term "dropouts" to mean all persons who failed to graduate high school, regardless of whether they ever went to school in the first place. We shall use the phrase "graduate cohort" to mean the graduates in the WLS, and the phrase "complete cohort" to mean graduates and dropouts together, where the dropouts include all nongraduates who, by virtue of their year of birth, were potential classmates of the graduates.

Incorporation of dropouts into the calculation of the IQ selection differential involves several preliminary steps: (1) estimation of the proportion of dropouts in the complete cohort; (2) estimation of the fertility and mortality of dropouts; and (3) estimation of the mean IQ of dropouts.

Proportion of dropouts.—Our estimate of the proportion that dropouts are of the complete cohort is derived from the 1960 census of population. The average age of the graduate cohort at the time they were first studied in April, 1957, was 17.98 for females and 18.09 for males. For the present purpose, we consider for simplicity that they were 18.0 years of age. In the census, age is recorded in completed years of age, so that 18.0 is the average of 17 and 18 completed years of age. Those who were 17 and 18 in April 1957 were 20 and 21 three years later at the time of the census in April, 1960. Table 4 shows the percentage of dropouts among 20-year-olds and 21-year-olds by sex in the 1960 census for the state of Wisconsin. Average values for these two ages are 21.5 per cent for females and 28.0 per cent for males. These are our estimates of the proportion that dropouts are of each sex in the Wisconsin cohort. These dropout

percentages for Wisconsin are comparatively low. Comparable figures for the United States as a whole, also derived from the 1960 census, are 34.0 per cent for females and 37.8 per cent for males (Hauser, 1971, Table 4).

TABLE 4

PERCENTAGE NOT COMPLETING AT LEAST FOUR YEARS OF HIGH SCHOOL BY SEX AND SELECTED AGES: WISCONSIN, 1960

Age	Females	Males
15	99.7	99.7
16	99.4	99.6
17	92.6	95.3
18	44.6	57.0
19	22.5	30.8
20	21.8	27.4
21	21.2	28.6
22	21.8	29.4
23	23.1	30.4
24	25.3	30.7

SOURCE: Derived from U.S. Department of Commerce, Bureau of the Census (1960), Tables 101, 102.

It should be noted that the census data pertain only to the noninstitutional population, so that the census probably underestimates the proportion of dropouts because of the omission of institutionalized persons, who probably have higher dropout rates than noninstitutionalized persons. But institutionalized persons in 1960 for the United States as a whole were well under 1 per cent of the population (U.S. Department of Commerce, Bureau of the Census, 1983, pp. 31, 56), so that their omission introduces only a very slight bias in our results. Moreover, this bias is negligible when compared with the crudity of our estimates of the mean IQ of dropouts, discussed later. Therefore, we have ignored institutionalized persons in our calculations. Since the institutionalized population includes, among others, mentally retarded persons of below-average fertility, the effect of ignoring institutionalized persons is probably to exaggerate slightly our estimate of the IQ selection differential for the complete Wisconsin cohort, which is also discussed later.

Census data also allow a rough assessment of the racial composition of the dropouts. For this purpose, we used 1970 census data, since the 1960 data were not ideally suited. As mentioned, the Wisconsin cohort was approximately aged 30 in 1970. To base our estimates of composition on larger numbers, derived from a 1/100 Public Use Sample tape for Wisconsin from the U.S. 1970 census, we considered persons aged 28–32 in 1970. Among such persons, dropouts are defined as those with less than 12 completed years of education. The dropouts are 91.6 per cent white, 6.5 per cent black, and 1.9 per cent other races. These estimates, insofar as they pertain to the complete WLS cohort, may be biased by migration in and out of Wisconsin between 1957 and 1970, to the extent that this migration was selective by race. This bias is probably small, but we have not investigated it.

Fertility of dropouts.—The fertility of dropouts was estimated by applying the own-children method of fertility estimation (Cho, 1973; Cho et al., 1986), which is a reverse-survival technique, to 1970 and 1980 census data for the State of Wisconsin. A 1/100 Public Use Sample tape was used for 1970, and a 1/1000 Public Use Sample tape was used for 1980.

Fertility estimates derived by the own-children method may be tabulated by characteristics of women, and in particular by number of completed years of education, which was asked in the 1970 and 1980 censuses. As already men-

tioned. dropouts are defined as those with less than 12 completed years of education. The application of the own-children method is straightforward in the case of women. To estimate the fertility of male dropouts, however. it was necessary to modify the method to obtain age-specific birth rates for men instead of women. This was necessary in order to be able to tabulate the fertility estimates by men's characteristics. the characteristic of interest being completed years of education. The modification of the method involves matching children to fathers instead of mothers within the same household. The own-children fertility estimates for men are less precise than those for women, mainly because children of divorced parents are more likely to be living with their mothers than their fathers, so that matching is less complete and less accurate for men than for women.

The own-children method is usually used to estimate age-specific birth rates for a particular year or time period. In the present instance. however. we are interested in estimating birth rates for a real cohort. If the time depth of the period estimates is long enough. the period estimates can be spliced together to yield estimates of age-specific birth rates for real cohorts. Our strategy was to consider a cohort centered at age 17 in 1957. 20 in 1960. 30 in 1970. and 40 in 1980. (Actual computations were based on persons aged 15–18 in 1957. 18–21 in 1960. 28–31 in 1970. and 38–41 in 1980.) Applications of the own-children method to the census of 1970 for Wisconsin yielded estimates of age-specific birth rates for this cohort for ages 15–19 during 1955–59, 20–24 during 1960–64, and 25–29 during 1965–69. Application to the census of 1980 for Wisconsin yielded estimates of age-specific birth

rates for this cohort for ages 25–29 during 1965–69, 30–34 during 1970–74, and 35–39 during 1975–79. The two overlapping estimates of the age-specific birth rate at ages 25–29 during 1965–69 were averaged to yield a final estimate for this age group and time period. For reasons of computational convenience, these estimates are centered on a cohort that is about one year younger than the WLS cohort, but this discrepancy introduces only very small errors in the fertility estimates, as will be seen shortly. Further details of how the own-children fertility estimates for the Wisconsin cohort were constructed are given elsewhere (Retherford and Sewell. 1986).

A difficulty is that the own-children method. as applied here, provides fertility estimates only through ages 35–39. It is necessary to estimate fertility also at ages 40–44 and 45–49. A simple ratio method was used to do this. The method involves a standard age-specific fertility schedule, which we somewhat arbitrarily chose to be the set of age-specific birth rates for the United States in 1970. This schedule indicates that the percentage of total fertility accounted for by age-specific fertility after age 40 is only 1.74 per cent. so in fact very little fertility is being extrapolated. In the standard schedule of age-specific birth rates, which we denote simply as $f(x)$. the ratio $f(40-44)/f(35-39) = 0.2555$ and the ratio $f(45-49)/f(35-39) = 0.0158$. We multiplied the own-children estimate of the age-specific birth rate for dropouts at ages 35–39 by each of these ratios in turn to obtain estimates of age-specific birth rates at ages 40–44 and 45–49.

The procedure just described for estimating age-specific birth rates for dropouts in the Wisconsin cohort is indirect enough that it seemed worthwhile to apply it to other education categories as
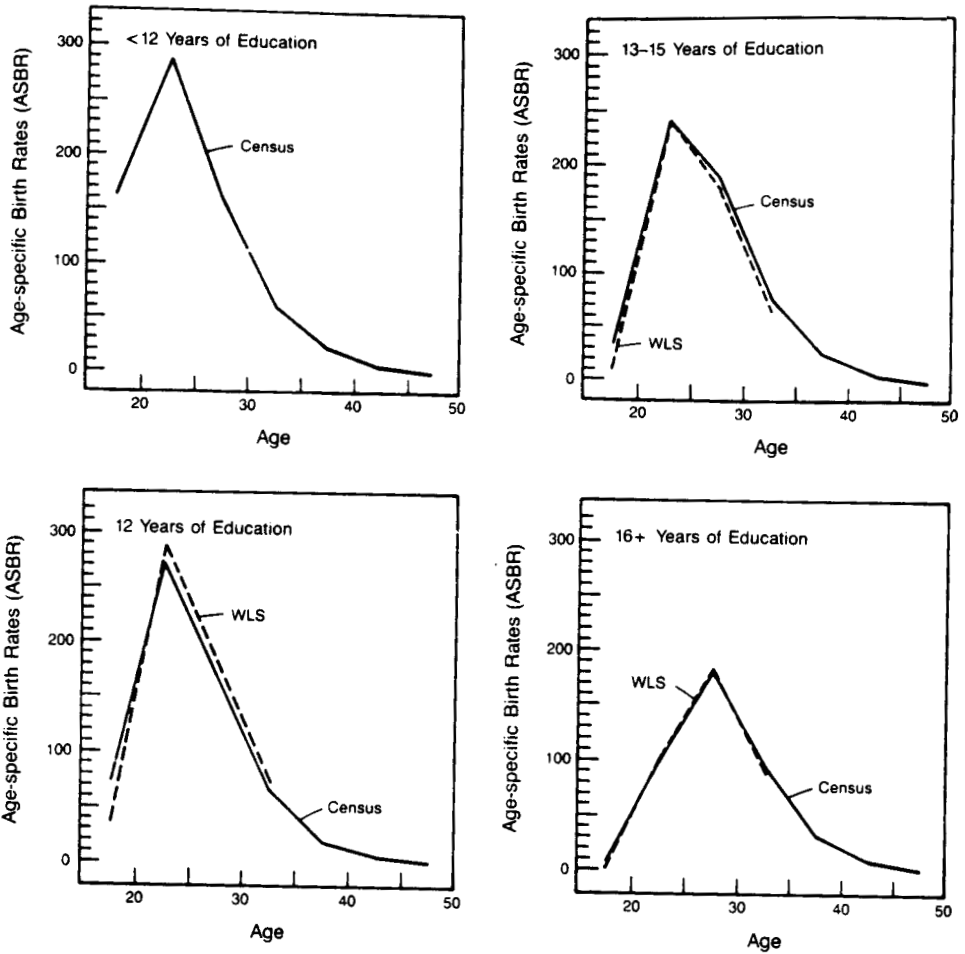
Fig. 2.—Estimates of age-specific birth rates by education for females in the Wisconsin cohort, derived alternatively from the WLS and from census data.

well, in order to compare the own-children estimates of age-specific birth rates to parallel estimates derived earlier from the WLS birth histories. The comparisons also provide a check on possible bias from migration in and out of Wisconsin between 1957 and 1980. Figures 2 and 3 show the results of these comparisons. The first panel in each figure shows the results for dropouts, for which there are no comparable estimates from the WLS. But for the other education categories, comparisons can be made for age groups 15–19, . . . , 30–

34. The agreement between estimates based alternatively on the own-children method applied to census data and the birth history method applied to WLS data is quite good. The comparisons suggest that our fertility estimates for dropouts in the Wisconsin cohort are reasonably accurate. The comparisons also suggest that migration in and out of Wisconsin introduces only minor distortions in the fertility estimates.

Table 5 shows a substantial sex differential in the fertility of dropouts. The TFR for female dropouts is about three-
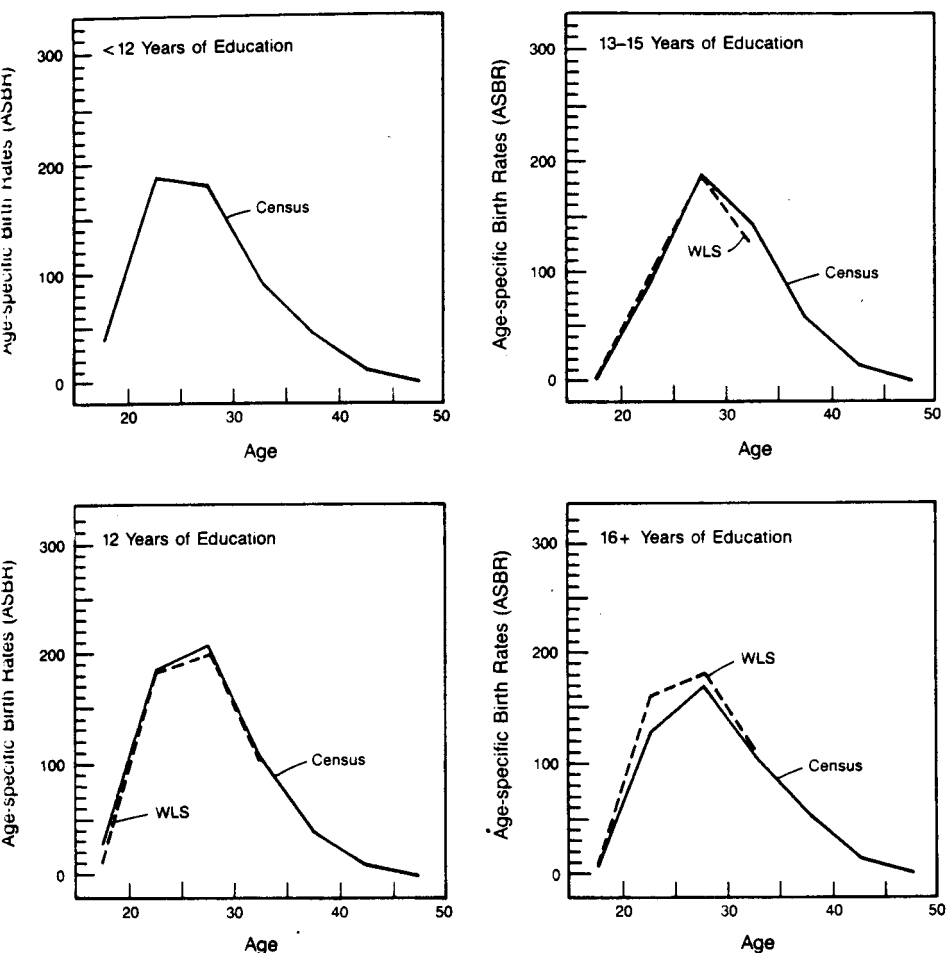
Fig. 3.—Estimates of age-specific birth rates by education for males in the Wisconsin cohort, derived alternatively from the WLS and from census data.

quarters of a child higher than that for male dropouts. Moreover, comparison with Table 2 shows that the TFR for female dropouts is considerably higher than the TFR for female graduates, regardless of IQ decile, whereas the TFR for male dropouts is only slightly higher than the TFR for male graduates. A plausible explanation of this sex differential in the fertility of dropouts is that girls frequently drop out of school to have children. This happens less frequently for boys, partly because children tie down women more than men,

and partly because boys tend to become fathers at an older age than that at which girls tend to become mothers.

Mortality of dropouts.—Kitagawa and Hauser (1973, p. 17) have estimated mortality by education for whites by sex in the United States in 1960. From their data we computed the proportion dying between ages 25 and 45 for graduates and dropouts, by sex. For a given sex, we denote these proportions as $q_g(25)$ and $q_d(25)$, respectively. Similarly, we denote the proportions dying between ages 18 and 36 for graduates and drop-

outs in the Wisconsin cohort as $q_g(18)$ and $q_d(18)$. With values of $q_g(18)$, $q_g(25)$, and $q_d(25)$ prespecified. $q_d(18)$ for the Wisconsin cohort was estimated as

$$q_d(18) = q_g(18)[q_d(25)/q_g(25)] \quad (5)$$

where the quantities with argument 18 pertain to the Wisconsin cohort and the quantities with argument 25 pertain to U.S. whites in 1960 from Kitagawa and Hauser. Equation (5) assumes that the known ratio $q_d(25)/q_g(25)$ from the Kitagawa and Hauser data closely approximates the unknown ratio $q_d(18)/q_g(18)$ for the complete Wisconsin cohort.

The value of $q_g(18)$ for the Wisconsin cohort estimated from Equation (5) was then matched to a Coale-Demeny Model West life table, which was taken as an estimate of the life table for dropouts. These calculations were done separately for each sex, with results shown in Table 5. Comparison with life expectancy values in Table 2 shows that life expectancy of dropouts falls short of life expectancy for persons in the lowest IQ decile by 1.25 years for females and 2.3 years for males.

TABLE 5

ESTIMATES OF INTELLIGENCE.
REPRODUCTIVITY. AND SELECTION
FOR DROPOUTS

Measure [a]	Females	Males
Mean IQ	86.00	86.00
$F(35)$	3.35	2.48
TFR	3.49	2.76
GRR	1.70	1.41
e_0	73.75	66.00
NRR	1.65	1.30
r	0.02119	0.00985
T	23.68	26.79
Doubling time (yrs)	32.71	70.39
W	1.181	0.971

[a] For definitions. see notes to Table 2. W exceeds unity for female dropouts only because the second IQ decile of graduates is retained as the basis of comparison for females.

Mean IQ of dropouts.—Our estimates of the mean IQ of dropouts are by far the weakest link in the chain of data necessary for calculating the IQ selection differential for graduates and dropouts together in the complete Wisconsin cohort. In fact we have virtually no data at all on IQ of dropouts in Wisconsin. Our estimates of mean IQ of dropouts are culled instead from a number of studies of other populations. All of these other populations comprise some segment of the U.S. population, but none pertains just to Wisconsin.

The best estimates of mean IQ of dropouts relative to mean IQ of graduates that we have been able to find are derived from Project TALENT (Wise et al.. 1977; Wise and Steel, 1980. p. 115). The TALENT samples included over 400.000 men and women from the high-school classes of 1960 through 1963 in the United States. covering grades 9–12 in 1960. A battery of tests was given to the respondents. and the results of these tests were summarized in a measure called general academic aptitude (GAA). Subsamples from these large samples are of particular interest to us here. Two subsamples, 101.777 ninth-graders and 99,849 tenth-graders, were followed up eleven years later (Wise et al.. 1977, p. 4). Dropouts were identified in the follow-up survey, making it possible to calculate mean GAA score for both graduates and dropouts.

The eleven-year follow-up response rate. however, was only 20 per cent in each case. It is highly likely that nonrespondents tended to be selected for low IQ, so that results based on a 20-per cent response rate may be severely biased. The authors of the study were aware of this problem and attempted to solve it through an intensive study of a sample of the nonrespondents. In order to cor-

rect for nonresponse bias, a representative sample of 4,029 nonrespondents in the ninth grade sample, which is the one in which we are interested, was selected for intensive follow-up. Of these, 87 per cent were found, and of these 90 per cent provided usable information, yielding an overall response rate of 78 per cent. On the basis of this sample of nonrespondents, a set of weighting factors was derived and applied to the original sample of ninth graders. The weights were designed to make the original sample representative on two variables, general academic aptitude and socioeconomic index. For our purposes, this choice was fortunate, since GAA is the variable in which we are interested.

Results showed that male dropouts scored 1.09 standard deviations below male graduates, and female dropouts scored 0.95 standard deviations below female graduates on the GAA measure (Wise and Steel, 1980, p. 115). The greater difference between graduates and dropouts for males than for females is consistent with the fertility differences between males and females among dropouts, noted earlier. It was found earlier that female dropouts have much higher fertility than male dropouts. We hypothesized that the main reason for this sex differential in fertility among dropouts was that girls were much more likely than boys to drop out of school because of the birth of a child. This means that girls are more likely than boys to drop out of school for reasons other than low academic ability. Therefore, the IQ difference between graduates and dropouts should be smaller for women than for men. The Project TALENT data confirm this expectation, but the sex differential in the IQ difference between graduates and dropouts is small. It seems that girls who drop out of school because of pregnancy tend to have about the same average IQ as male dropouts.

Another national-level study of interest is the study on Youth in Transition, which yielded data on dropouts between the tenth and twelfth grades (Bachman, 1970; Bachman et al., 1971). A sample of 2,519 tenth-grade boys was identified in 1967 and followed up again in 1970. There were 1,949 graduates and 326 dropouts, yielding a dropout rate of 17 per cent between the tenth and twelfth grades. This compares with a dropout rate between tenth and twelfth grades of 10.5 per cent in the Wisconsin cohort (calculated from Hauser, 1971, Table 6). The QUICK test, with a mean score of 108.5 and a standard deviation of 12.5, was used to measure intelligence. Dropouts scored 0.47 standard deviations below the graduates on this test (calculated from data in Bachman et al., 1971, Table C–3–6, p. 202). This difference is considerably smaller than the gap of about one standard deviation derived from the TALENT data for ninth-grade dropouts discussed earlier. One expects the gap between tenth-grade dropouts and graduates to be less than the gap between ninth-grade dropouts and graduates, but the observed difference is larger than we anticipated.

An earlier study reported by Wolfle (1954) is also of interest. Wolfle presents results for a cohort of 14-year-olds in 1949. The mental test scores are from the Army General Classification Test (AGCT), which apparently was administered to a large national sample of 14-year-olds in 1949. Unfortunately, the details concerning the original sample are not given, and it appears that various kinds of indirect estimates may have been used to fill gaps in the data. For example, regarding rates of progression to

higher grades of school, Wolfle states that "the rates shown are estimated for the year 1953 and have been derived from special studies conducted by the Commission [on Human Resources and Advanced Training] supplemented by census data and a number of reports of school dropout studies conducted by others" (Wolfle, 1954, p. 311). But most of these special studies and studies conducted by others are not cited, nor are the details of the methodology elaborated. Thus, it is not possible for us to assess the quality of these results.

Wolfle's results include a table that classifies persons by AGCT score (with scores grouped in 5-point intervals) and educational attainment, from which one can calculate mean AGCT scores for graduates and dropouts (Wolfle, 1954, p. 314). As in the Project TALENT study, dropouts are defined as those who leave school between the ninth and twelfth grades. The AGCT is designed to have a mean of 100 and a standard deviation of 20. When the mean AGCT scores are converted into standard scores, by subtracting the mean score and dividing by 20, it is found that dropouts score 1.12 standard deviations below graduates. There is no breakdown by sex. The estimate of 1.12 is in rough agreement with the corresponding estimates derived from the Project TALENT data.

None of the national-level studies says anything about dropouts who did not advance as far as the ninth grade. According to the 1 per cent Public Use Sample tape for Wisconsin in 1970, based on census data on number of completed years of education of persons aged 28–32 (average age of 30), about 28 per cent (24 per cent for females and 33 per cent for males) of dropouts in the Wisconsin cohort did not complete

ninth grade. Since those who did not complete ninth grade probably had lower mean IQ than those who dropped out after ninth grade, the differences in mean IQ between dropouts at all levels and graduates are very likely greater than the estimates derived from the studies just reviewed.

A conservative estimate of the difference in mean IQ between high school graduates and dropouts (those in an age cohort who do not graduate from high school) would seem to be about one standard deviation, and we take this as our base estimate. In the 1975 WLS sample, graduates have a mean IQ of almost 101, with a standard deviation of 15. Thus, our base estimate of the mean IQ of dropouts is 86. This estimate is obviously very rough.

Table 5 summarizes the estimates of intelligence, reproductivity, and selection for dropouts. The results in this table may be compared with those in Table 2 for graduates.

The IQ selection differential for graduates and dropouts combined for the Wisconsin cohort.—The IQ selection differential, S, for graduates and dropouts combined, can now be calculated in the following way: Let subscript i denote IQ decile in the WLS, and let subscript d denote dropouts. For each sex separately, define N as

$$N = (\Sigma_i N_i) + N_d \qquad (6)$$

where N_i denotes the number of WLS graduates in the ith IQ decile, and N_d denotes the estimated number of dropouts that one would have obtained in the WLS had the survey sampled all persons of graduation age instead of just high-school graduates.

For each sex separately, N_d was estimated as follows: First, we computed,

for Wisconsin, the ratio of high-school dropouts to high-school graduates among 20-year-olds in the 1960 census (from data in Hauser, 1971, Table 4). These ratios are 0.279 for females and 0.375 for males. Multiplying this ratio by the first of the two terms on the right side of (6) yielded an estimate of N_d. The IQ selection differential for graduates and dropouts combined was then calculated by Equation (3), as before, the sole difference being that the summations ranged over eleven terms instead of ten, the eleventh pertaining to dropouts. The term for dropouts utilizes the value of N_d calculated as just described, and N is redefined in accordance with Equation (6).

Results are shown in Table 6, which can be compared with earlier results in Table 3 for graduates only. In order to assess the sensitivity of S to possible error in the estimate of the mean IQ of dropouts, we have based the calculation not only on a mean IQ of dropouts of 86 but also on 81 and 92. Adding dropouts into the calculation of the IQ selection differential, S, increases the estimate of S (i.e., makes it more negative) substantially for females but very little for males. The reason for the greater effect on females is that female dropouts have considerably higher net fertility than female graduates, whereas male dropouts have only marginally higher net fertility than male graduates, as seen from Tables 2 and 5. Table 6 also shows that the estimate of S for graduates and dropouts together is sensitive to errors in the estimate of mean IQ for dropouts, especially for females. With the mean IQ of dropouts estimated at the middle value of 86, the average value os S for males and females together is −0.81, or about eight-tenths of an IQ point decline in a generation.

TABLE 6
Estimates of the IQ Selection Differential, S, for Graduates and Dropouts Together, by Sex, Derived from the WLS

Mean IQ assumption for dropouts[a]	Selection Differential		
	Females	Males	Both Sexes[b]
$IQ_d = 81$	−1.61	−0.31	−0.96
$IQ_d = 86$	−1.33	−0.28	−0.81
$IQ_d = 91$	−1.05	−0.25	−0.65

[a] IQ_d denotes the mean IQ assumed for dropouts. Our preferred alternative is $IQ_d = 86$.
[b] Values in the total column are simple averages of values in the previous two columns for males and females separately.

The value of S for both sexes combined can be written $S = 0.5(S^f + S^m) = 0.5\ S^f + 0.5\ S^m = -0.67 - 0.14 = -0.81$, where superscripts f and m denote females and males respectively. This decomposition of S into two sex-specific components indicates that the contribution of females to the IQ selection differential is almost five times greater than the contribution of males $[(-0.67)/(-0.14) = 4.8]$.

Generational Change in Mean Genotypic IQ

The phenotypic IQ of an individual (or phenotypic value, in the vocabulary of quantitative genetics) is, by definition, the same as measured IQ. The individual's genotypic IQ is the value that he or she would have, hypothetically, if reared in the average environment of the population from which the individual came. The genic (or additive) IQ is the transmissable component of IQ; animal breeders call this the breeding value. The expected genotypic and phenotypic IQ's of a child, if reared in the average environment, are the same. The expected genotypic IQ of a child is the average of the genic IQ's of its parents.

The IQ selection differential is not the same as the generational change in mean genotypic IQ. Individuals with higher than average phenotypic IQ typically have this higher value not only because of higher genotypic IQ but also because of being reared in better environments. Therefore, genotypic IQ tends to be lower than phenotypic IQ and genic IQ lower yet, although both still tend to be higher than the average IQ. The same principle works in reverse for individuals whose IQ is below average; in this case genotypic and genic IQ tend to be above phenotypic IQ but still below average IQ. Since only the genic value is biologically transmitted, children tend to be closer to the population mean than their parents. This tendency is called regression toward the mean.

As explained earlier, the IQ selection differential is the generational change in mean IQ if, hypothetically, the child's IQ is the average of its parents' IQs. It is computed as the difference between the mean IQ of reproducing individuals, weighted by the number of their children, and the unweighted mean IQ of reproducing individuals to begin with. If the selection differential, as estimated here, is -0.81 IQ point, then the expected reduction in genotypic IQ over the next generation will be somewhat less than 0.81 because only the genic values of the parents are biologically transmitted. Unfortunately it is difficult to say how much less.

Under conditions of artificial selection, as in plant and animal breeding experiments with randomized environments, the generational change, R, in a measured trait is estimated as

$$R = h^2 S \qquad (7)$$

where S is the selection differential for that trait and h^2 is the narrow-sense heritability. This heritability is the ratio of the genic value to the phenotypic value, both scaled as deviations from the population mean. In practice it is estimated from correlations between relatives (often half-sibs) reared in controlled or randomized environments. In the case of human beings, however, the environments of close relatives tend to be similar, so that human correlations typically confound genetic and environmental causes. Therefore, simple h^2 estimates tend to be biased; corrections of this bias require additional assumptions and introduce additional uncertainties.

With randomized environments the predicted phenotypic IQ of a child is the same as its predicted genotypic value. When there is genotype-environment correlations, as in human populations, this is no longer true. Ironically, because of positive genotype-environment interactions and covariances, the phenotypic IQ of a child is more predictable from the parental average IQ than its genotypic value; it can be estimated from direct measurements.

Empirically based estimates of h^2 for IQ, based on midparent-offspring regressions, vary between 0.4 and 0.6 (Plomin and DeFries, 1980; this range excludes estimates based on Burt's questionable data, which have been reviewed by Goldberger, 1979). Rao et al. (1982) more recently arrived at an estimate in the range of 0.31 to 0.34, based on a highly sophisticated statistical model but less than ideal data. Our educated guess is that h^2 for IQ is about 0.4, and that the generational change in mean genotypic IQ is accordingly about 0.4 times as large as the IQ selection differential, i.e., in the neighborhood of one-third of an IQ point decline for the Wisconsin cohort $[(0.4)(-0.81) = -0.32]$. It could, however, be consider-

ly smaller or larger, within the range tween zero and eight-tenths of an IQ int decline.

SUMMARY AND DISCUSSION

Previous studies of intelligence and nily size provide a mixed picture of ɔ relationship between these two variɔles. Many early studies based concluɔns on the commonly observed negaɔe correlation between measured IQ d sibship size. But this correlation has ɔen shown to be biased (insofar as it ptures the relationship between IQ d fertility), because sibship size, ɔich measures the fertility of the parental generation, contains no information about childless persons in the parental generation. The few studies that ɔvestigated this question found that the ɔgative correlation between IQ and mily size disappeared when childless ɔrsons in the parental generation were ken into account. These studies were ɔsed on samples that may not have ɔen representative of the entire popuɔion, but they nevertheless indicate ɔthodological reasons, having to do ɔth biased correlations, for basing conɔusions about the relationship between) and fertility on information on numɔr of offspring by IQ rather than numɔr of siblings by IQ.

Studies that have analyzed data on ɔmber of offspring by IQ show varying ɔsults. These studies may be divided ɔto two groups, those that examined ɔhorts whose reproductive experience ɔverlapped the post–World-War-II ɔby boom, and those that examined ɔhorts whose reproductive experience ɔgan toward the end or after the baby ɔom.

Most of the studies of cohorts whose ɔproductive experience overlapped the

baby boom show a slight positive relationship between IQ and number of offspring; however, none of these studies is based on a probability sample. One other study, however, is based on a series of NORC probability samples designed to be nationally representative of the United States. This study showed a negative relationship between IQ and number of offspring for all 5-year birth cohorts considered, both before, during, and after the baby boom. It also showed that differential fertility tended to converge during the baby boom (the correlation became less negative but never positive) and diverge again in the negative direction after the baby boom. Unfortunately, results from this study contain an unknown amount of bias due to nonresponse, which is probably concentrated among persons of less than average IQ. Moreover, the IQ measure is of unknown quality, and there may be problems of selectivity having to do with the retrospective nature of the data collection, which takes no account of differential mortality by IQ.

Studies of cohorts whose reproductive experience began toward the end or after the baby boom show a more sharply negative correlation between IQ and number of offspring. One of these studies was the one just mentioned that used NORC data. Another was based on the U.S. National Longitudinal Study of Labor Market Experience. However, the intelligence measure in this latter study is questionable, and there is a serious problem of nonresponse. Another set of studies compared fertility trends of elite groups known or presumed to have high intelligence (samples from *Who's Who*, the Forbes 400, Terman's high IQ group, and the American Mensa) with the fertility of the nation as a whole. These

studies, which suffer by not being based on representative samples as well as some other difficulties, also tend to confirm the pattern of convergence of negative differential fertility by IQ during the baby boom followed by divergence in the negative direction after the baby boom.

Taken singly, none of the earlier studies is very convincing. Taken together, however, they present a fairly consistent picture: Differential fertility by IQ in the United States appears to have been negative in this century. It probably became less negative during the baby boom (and positive for some samples). It seems to have diverged again in the negative direction after the baby boom. It is worth noting that fertility differentials by education in the United States have shown a rather similar trend (see, for example, Rindfuss and Sweet, 1977; Sweet and Rindfuss, 1983). However, educational level is an inadequate substitute for measured intelligence since most studies show the two variables to be correlated only at about 0.5 (Sewell et al., 1970).

Our own empirical results, based on the Wisconsin Longitudinal Study (WLS), confirm a small but statistically significant negative relationship between fertility and IQ for a large representative sample of the cohort of Wisconsin high school graduates in 1957, whose reproduction commenced at the end of the baby boom. Results are based on fertility to age 35, which was extrapolated by means of a demographic model to the end of the reproductive ages. A combination of sample data and demographic models was also used to compute mortality by IQ. The estimates of fertility and mortality by IQ allowed estimation of the intrinsic rate of natural increase by IQ and a number of derived measures. One of these derived measures is the IQ selection differential, which is what the generational change in mean IQ would be if, hypothetically, each child had the same IQ as the mean of its parents' IQs.

We calculated the IQ selection differential for each sex separately and for graduates and the complete cohort (graduates plus dropouts) separately. In reality, of course, given sexual reproduction, one expects the IQ selection differential to be about the same for either sex. The purpose of doing the calculation separately for each sex was not only to circumvent the difficulty that we know IQ for only one parent, but also to get some sense of the relative importance of female differential fertility by IQ and male differential fertility by IQ to the overall IQ selection differential. The rationale for the separate examination of graduates and the complete cohort is similar. An additional reason for examining graduates separately was that the results just for graduates are based on comparatively good data, whereas the results for graduates and dropouts together are based on a mix of comparatively good data and comparatively weak data. Our two-step procedure allowed a rough assessment of how much the results based on the comparatively good data for graduates were affected by addition of the comparatively weak data for dropouts. The effects of adding dropouts were substantial for females but slight for males.

In this regard, an unexpected finding from our analysis is that differential fertility by IQ is much more negative for women than for men. In retrospect, this finding is perhaps not surprising, given the ubiquitous finding in demographic

udies that in modern populations fer-
ity varies more sharply and negatively
y wife's education than by husband's
Jucation (Cochrane, 1979). Neverthe-
ss, we did not anticipate such a large
:x difference. The magnitude of the IQ
election differential is about three
mes larger for females than for males
i the WLS cohort of graduates only.
nd almost five times larger for females
ian for males in the complete cohort of
raduates and dropouts together. Re-
arding possible explanations of this sex
ifference, we have two hypotheses:
'irst, childbearing gets in the way of
igher education and careers much
nore for bright women than for bright
nen. And second, dull women are more
:kely than dull men to start having chil-
iren at an early age, whether by inten-
ion or because of an accidental out-of-
vedlock birth. We have not tested the
ıypotheses, which remain speculative.

For both sexes together, we estimate
he IQ selection differential to be about
)ne-half of an IQ point decline in a gen-
:ration for graduates only, and about
:ight-tenths of an IQ point decline in a
:eneration for graduates and dropouts
ogether. Both estimates must be
iewed with considerable caution, be-
:ause they are based partly on the use of
demographic models for estimation of
missing data, and because some of the
data on dropouts are weak and impre-
cise.

Subject to these limitations, our esti-
mate of the IQ selection differential of
about eight-tenths of an IQ point de-
cline in a generation may be viewed as
an upper bound of the generational
change in mean genotypic IQ for the
complete Wisconsin cohort. Genotypic
IQ is defined as the expected value of
measured IQ for an individual of a given

gene configuration, or genotype, under
the hypothetical assumption that the in-
dividual is raised in the average environ-
ment obtaining in the population. The
generational change in mean genotypic
IQ is less than the IQ selection differen-
tial, mainly because of regression to-
ward the mean, which means that in ac-
tuality children of extremely bright or
extremely dull parents tend to have
genotypical IQ's closer to the mean IQ
of the population than their midparent
genotypic IQ, and because of environ-
mental influences on measured IQ. We
cannot estimate the generational change
in mean genotypic IQ from our data, but
an educated guess, based on findings
from IQ heritability studies, is that it is
about four-tenths as large, or about one-
third of an IQ point decline in a genera-
tion in the complete Wisconsin cohort.

Our finding that mean genotypic IQ
may be declining is at variance with sev-
eral earlier studies, based on time series
of cross-sections, that indicate that
mean IQ has been increasing over time.
A possible explanation is that mean
genotypic IQ has been declining very
slowly, at a pace well below one IQ
point per generation, but that pheno-
typic, or observed, IQ has been increas-
ing because of offsetting environmental
improvements, especially in the general
educational level of the population. But
this explanation is speculative, and
more research is needed to resolve this
puzzle. In this regard, it would be desir-
able to undertake longitudinal studies,
based on large, nationally representa-
tive samples, in which detailed demo-
graphic information and comparable IQ
measurements for both parents and chil-
dren are obtained. To our knowledge,
such a study has never been done.

ACKNOWLEDGMENTS

REFERENCES

ANASTASI, A. 1956. Intelligence and family size. Psychol. bull. **53**:187–209.

ANGOFF, W. H. 1971. Scales, norms and equivalent scores. *In* R. L. Thorndike (ed.), Educational measurement. 2nd ed. American Council on Education, Washington, D.C.

BACHMAN, J. G. 1970. Youth in transition. Vol. II. The impact of family background and intelligence on tenth-grade boys. Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

BACHMAN, J. G., S. GREEN, and I. D. WIRTANEN. 1971. Youth in transition. Vol. III. Dropping out—problem or symptom? Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

BAJEMA, C. J. 1963. Estimation of the direction and intensity of natural selection in relation to human intelligence by means of the intrinsic rate of natural increase. Eugen. Quart. **10**:175–187.

BUROS. O. 1965. The sixth mental measurements yearbook. Gryphon Press, Highland Park, N.J.

CATTELL, P. 1930. Comparability of IQ's obtained from different tests and different IQ levels. School and Society **31**:437–442.

———. 1951. The fate of national intelligence: Test of a thirteen-year prediction. Eugen. Rev. **42**(3):138–148.

CHO, L. J. 1973. The own-children method of fertility estimation: An elaboration. *In* International Union for the Scientific Study of Population. International Population Conference. Liege 1973. Vol. 2. pp. 263–278.

CHO, L. J., R. D. RETHERFORD, and M. K. CHOE. 1986. The own-children method of fertility estimation. An East-West Center Book. East-West Center, Honolulu, Hawaii.

CLARRIDGE, B. R., L. L. SHEEHY, and T. S. HAUSER. 1977. Tracing members of a panel: A 17-year follow-up. *In* K. F. Schuessler (ed.), Sociological methodology, 1978. Jossey-Bass, San Francisco.

COALE, A. J., and P. DEMENY. 1966. Regional model life tables and stable populations. Princeton University Press, Princeton, N.J.

COCHRAN, W. G. 1963. Sampling techniques. John Wiley, New York.

COCHRANE, S. H. 1979. Fertility and education: What do we really know? Johns Hopkins University Press. Baltimore, Md.

CROW, J. F. 1986. Basic concepts in population, quantitative, and evolutionary genetics. W. H. Freeman, New York.

DUNCAN, O. D. 1952. Is the intelligence of the general population declining? Amer. Sociol. Rev. **17**:401–407.

FALCONER, D. S. 1966. Genetic consequences of selection pressure. *In* J. E. Meade and A. S. Parkes (eds.), Genetic and environmental factors in human ability. Oliver and Boyd, Edinburgh.

———. 1981. Quantitative genetics. 2nd ed. Longmans, New York.

FEATHERMAN, D. L., and R. M. HAUSER. 1977. The process of stratification: Trends and analyses. Academic Press, New York.

———. 1978. Opportunity and change. Academic Press, New York.

FLEMMING, C. W. 1925. A detailed analysis of achievement in the high school: Comparative significance of certain mental, physical, and character traits for success. Columbia University, Teachers' College, New York.

FLYNN, J. R. 1984. The mean IQ of Americans: Massive gains 1932 to 1978. Psychol. Bull. **95**:29–51.

———. 1987. Massive IQ gains in 14 nations: What IQ tests really measure. Psychol. Bull. **101**:171–191.

GOLDBERGER, A. S. 1979. Heritability. Economica **46**:247–327.

GULLIKSEN, J. 1950. Theory of mental tests. John Wiley and Sons, New York.

AUSER. R. M. 1971. Some notes on school retention and enrollment in Wisconsin. Memorandum N. 50 to Sewell-Hauser Project Staff, Feb. 17, 1971. Dept. of Sociology, University of Wisconsin. Unpublished.

AUSER. R. M., and W. H. SEWELL. 1984. Family effects in simple models of education, occupational status, and earnings: Findings from the Wisconsin and Kalamazoo studies. CDE Working Paper 84–29. Center for Population and Ecology, University of Wisconsin, Madison.

AUSER. R., W. H. SEWELL. and B. R. CLARRIDGE. 1982. The influence of family structure on socioeconomic achievement: A progress report. CDE Working Paper 82–59. Center for Demography and Ecology, University of Wisconsin, Madison.

AUSER. R. M., S. L. TSAI. and W. H. SEWELL. 1983. A model of stratification with response error in social and psychological variables. Sociol. Educ. **56**:20–46.

ENMON. V. A. C., and M. J. NELSON. 1954. The Henmon-Nelson Tests of Mental Ability. Manual for administration. Grades 7–12. Houghton Mifflin. Cambridge. Mass.

IGGENS. J. V., E. W. REED. and S. G. REED. 1962. Intelligence and family size: A paradox resolved. Eugen. Quart. **9**:84–90.

EYFITZ. N., and W. FLIEGER. 1971. Population: Facts and methods of demography. Freeman and Co.. San Francisco.

ITAGAWA. E. M., and P. M. HAUSER. 1973. Differential mortality in the United States: A study in socioeconomic epidemiology. Harvard Univ. Press. Cambridge. Mass.

UTHER. N. Y. 1982. Fitting age-specific fertility with a truncated Pearson Type III curve. Asian Pacif. Census Forum **8**(4):5–12.

ATIONAL CENTER FOR HEALTH STATISTICS. 1975. U.S. decennial life tables for 1969–71. Vol. II. State life tables: Wisconsin. DHEW Publication No. (HRA) 75-1151. Washington. D.C.

JLNECK. M. R. 1976. The determinants of education and adult status among brothers: The Kalamazoo study. Unpublished Ph.D. dissertation. Harvard University. Cambridge. Mass.

JLNECK. M., and B. WOLFE. 1980. Intelligence and family size: Another look. Rev. Econ. Statist. **62**:241–247.

OMIN. R., and J. C. DeFRIES. 1980. Genetics and intelligence: Recent data. Intelligence **4**:15–24.

ATCLIFF. J. M. 1934. An analysis of results obtained from different intelligence tests and from repeated examination with particular reference to the effect of practice. Unpublished Ph.D. dissertation. Harvard University. Cambridge, Mass.

RAO, D. C., N. E. MORTON, J. M. LALOUEL, and R. LEW. 1982. Path analysis under generalized assortative mating II: American IQ. Genet. Res. **39**:187–198.

RETHERFORD. R. D., and W. H. SEWELL. 1986. Intelligence and family size reconsidered. Working papers of the East-West Population Institute. No. 39. East-West Center, Honolulu. Also issued as CDE Working Paper 86-3. Center for Demography and Ecology, University of Wisconsin, Madison.

RINDFUSS. R. R., and J. A. SWEET. 1977. Postwar fertility trends and differentials in the United States. Academic Press, New York.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION. 1933. The intelligence of Scottish Children. Publications of the Scottish Council for Research in Education, V. University of London Press. Ltd.. London.

———. 1949. The trend of Scottish intelligence. University of London Press, Ltd., London.

SEWELL. W. H., A. O. HALLER, and G. W. OHLENDORF. 1970. The educational and early occupational attainment process: replication and revision. Amer. Sociol. Rev. **35**:1014–1027.

SEWELL. W. H., and R. M. HAUSER. 1975. Education, occupation, and earnings: Achievement in the early career. Academic Press. New York.

SEWELL. W. H., and R. M. HAUSER. 1977. On the effects of families and family structure on achievements. *In* P. Taubman (ed.), Kinometrics: Determinants of socioeconomic success within and between families. North-Holland. Amsterdam.

SEWELL. W. H., and R. M. HAUSER. 1980. The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements. *In* A. C. Kerckhoff (ed.), Research in sociology of education and socialization. Vol. 1. JAI Press, Greenwich, Conn.

SEWELL. W. H., R. M. HAUSER, and W. C. WOLF. 1980. Sex. schooling and occupational status. Amer. J. Sociol. **86**:551–583.

SWEET. J. A., and R. R. RINDFUSS. Those ubiquitous fertility trends: United States, 1945–79. Soc. Biol. **30**:127–139.

THORNDIKE. R. L. 1942. Two screening tests of verbal intelligence. J. Appl. Psychol. **26**:128–135.

THORNDIKE. R. L., and E. HAGAN. 1962. Ten thousand careers. Wiley and Sons, New York.

U.S. DEPARTMENT OF COMMERCE, BUREAU OF THE CENSUS. 1960. U.S. census of population, 1960. Vol. 1. Characteristics of the population, Part 51. Wisconsin. Washington. D.C.

——. 1983. Statistical abstract of the United States. 1984. Washington, D.C.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE. 1965. Life tables 1959–61. Wisconsin state life tables 1959–61. Vol. 2, No. 50. Public Health Service Publication 1252. Washington, D.C.

VAN COURT, M., and F. D. BEAN. 1985. Intelligence and fertility in the United States: 1912–1982. Intelligence 9:23–32.

VINING, D. R. 1982. On the possibility of the reemergence of a dysgenic trend with respect to intelligence in American fertility differentials. Intelligence 6:241–264.

——. 1986. Social versus reproductive success: The central theoretical problem of human sociobiology. Behav. Brain Sci. 9:167–187.

WALLER, J. 1971. Differential reproduction: Its relation to IQ test score, education, and occupation. Soc. Biol. 18:122–136.

WISE, L. L., D. H. MCLAUGHLIN, and K. J. GILMARTIN. 1977. The American citizen 11 years after high school. Vol. II. American Institutes for Research, Palo Alto, Calif.

WISE, L. L., and L. STEEL. 1980. Educational attainment of the high school classes of 1960 through 1963: Findings from Project TALENT. In A. C. Kerckhoff (ed.), Research in sociology of education and socialization, Vol. 1. JAI Press, Greenwich, Conn.

WOLFLE, D. 1954. America's resources of specialized talent. Harper & Brothers, New York.