International Journal of Sociology and Social Policy
SES VERSUS IQ IN THE RACE-IQ-DELINQUENCY MODEL
Robert A. Gordon

SES VERSUS IQ IN THE RACE-IQ-DELINQUENCY MODEL

by Robert A. Gordon,
Department of Sociology, Johns Hopkins University

What explains large and persisting differences in criminality between certain groups and why do those differences assume the particular magnitudes that are observed? In sociology and criminology, group differences in socio-economic status are generally asserted to be the fundamental cause, even when plausible alternative hypotheses are advanced. In this article, I demonstrate that black-white differences in rates of juvenile delinquency are explained successfully by a model based on IQ and that familiar socio-economic variables do not perform as well as IQ in that model. Finally, through a variety of analyses, I show that even on those few occasions when socio-economic variables approach the success of IQ, it is only because they function as surrogates for IQ.

Background

Ten years ago, I determined that a model based on the IQ means and standard deviations of blacks and whites would account almost perfectly for observed black-white differences in the sex- and race-specific proportions of individuals who qualify as official delinquents, according to some particular criterion, at least once by age 18.0 (Gordon, 1976). Such proportions are known as *lifetime prevalence* rates (here, to age 18.0), but it is convenient to refer to them simply as *prevalence* rates.

The model was based on the assumption that a modest but nevertheless dependable inverse association between IQ and delinquency existed within each race at the individual level, and that consequently the between-race difference in delinquency rates could be explained entirely by black-white differences in the two IQ parameters specified above, assuming that the major latent trait measured by IQ tests was distributed normally. It was thought that such an explanation would prove successful if the presently unknown IQ-specific delinquency rates within sex were roughly similar in both races. If the IQ-specific rates were similar enough, the difference between black and white prevalence rates would be mainly a reflection of the fact that blacks were more concentrated than whites in segments of the IQ range where delinquency rates were highest. Such a condition would mean that the relatively modest inverse association between IQ and delinquency within race was no guide at all, and even misleading, concerning the importance of IQ for explaining major delinquency differences at the aggregate level between blacks and whites.

Those major black-white differences at the aggregate level are of great practical significance. For example, black/white delinquency prevalence ratios are typically on the order of 3 or 4 to 1, thus corresponding closely to male/female prevalence ratios of 3, 4, or 5 to 1 within each population (Gordon, 1973: Table 3). Phenomenologically, such ratios mean that blacks would be experienced as more delinquent than whites to almost the same degree that males are experienced as more delinquent than females; the odds ratios (Reynolds, 1977: 35-36) in each case are roughly the same.

More concretely, a 3:1 black/white ratio for delinquency implies that as a community changes demographically from approximately all white to half black-as many cities in the U.S. have changed in recent decades-its delinquency rate would double. And, as a matter of fact, virtually all of the dramatic upsurge in arrest rates of youthful violent offenders that occurred between 1960 and 1977 can be encompassed by percentage increases of only 76% to 140% (Gordon, 1986a). These are the magnitudes underlying what has been experienced and described as "a crime wave of epic proportions" (Silberman, 1980: 367). Although these facts alone are not sufficient to show that the difference between black and white crime rates are largely what accounted for the upsurge, they do provide a sense of the scale of the black-white difference and of its relevance to current public concerns. Evidence to complete the argument, in the form of strong associations between racial composition and crime rates, has been presented by Gordon (1976) and Laub (1983). In view of its scope and obvious ramifications, the national problem of group differences in crime rates is actually far more urgent than the more ambitious one of preventing crime altogether (Gordon, 1986b).

Theoretical arguments concerning how and why low IQ contributes to delinquency have been considered elsewhere (Gordon, 1975, 1976, 1986a). However, it can be noted here that in addition to accounting for black-white delinquency differences on the outcome side, the model also explains in a general way (1) the widely observed inverse relation between social class and official delinquency (for references, see Gordon, 1976: note 1), (2) the surprising lack of variation within race in prevalence of delinquency according to size of place across most of the urban-rural continuum, and (3) the unusual stability of prevalence rates within race over time (Gordon, 1976). The first of these outcomes reflects the direct association between parental social class level and average IQ of offspring and parents within race, and the second and third reflect the relative constancy of IQ parameters within race from one community to another as well as over time.

Important developments related to the empirical context out of which the race-IQ-delinquency model emerged have been reported by Laub (1983). Using victimisation data instead of official delinquency data, Laub has confirmed and extended to blacks and to adults four propositions that I had earlier derived from a meta-analysis of delinquency prevalence rates for white juveniles (Gordon, 1976). First, race-specific rates vary little over most of the urban-rural continuum with no trend that would be indicative of "urbanism"; second, the general impression of such a trend is the product of confounding between size of place and racial composition; third, rates do decline at the rural extreme (for which theoretical reasons have been given in Gordon, 1975, 1976); and fourth, the point at which that decline begins occurs in the neighbourhood of community size 10,000.

Other relevant developments and emerging criticisms concerning the role of IQ in the model have recently been reviewed (Gordon, 1986a). Briefly, new evidence supporting the construct validity of the model has been concerned with the following issues: (a) the normality of the latent trait underlying IQ scores of both blacks and whites (Gordon, 1984); (b) the validity of IQ scores for both blacks and whites (Gordon, 1984, 1985, 1987; Jensen, 1980, 1985; Wigdor and Garner, 1982); (c) the validity of large black/white ratios in official crime statistics, which are corroborated by data from victimisation surveys based on crimes in which the victim could identify the offender's race (Hindelang, 1978, 1981; Langan, 1985; Laub, 1983); (d) the existence and consistency in several studies of modest negative correlations (in the low to high .20s) between IQ and delinquency in populations whose mean IQ would fall close to 100 (Gordon, 1986a: Table 1); (e) the meaning of some differences between blacks and whites in the IQ-delinquency correlation for individuals, which can reflect differences in location of the two races on the IQ continuum as distinct from differences between them in IQ-specific delinquency rates; and (f) the stability over time and accuracy of the estimates of the black and white IQ parameters used in testing the model (see Table 1, below).

Respect for the release date of the original work has precluded, until now, my referring to perhaps the strongest evidence of all of the construct validity of the model. Langan and Greenfeld (1985) have reported lifetime prevalence rates of adult imprisonment for inmates admitted to state prisons in 1973 and 1979, using the first through fourth admissions in order to vary the severity of the criterion with respect to recidivism. The race-IQ model fits their sex- and race-specific data extremely well, and so 12 more data points have been added to the four published thus far (for the latter, see Table 2, below). The extension of the model to adults, and to points in time that add more than a decade to the span previously explored, represents a major replication as well as an important generalisation of past findings.

One particular criticism of the model that I have encountered reflects confusion over the role of socio-economic status in determining both IQ and delinquency. At least, that is a likely interpretation of the criticism, which consisted only of the dismissive reaction "Of course!" As indicated elsewhere, the alternative possibility that this vague criticism was motivated by the existence of some race-IQ-delinquency model in the literature, unknown to me prior to my own description of that model, can be ruled out (Gordon, 1986a).

The usual meaning of such a criticism in social science debates is that one's causal model has been mis-specified (misconceived), and that the supposed causal connections under consideration are merely spurious by-products of some other, more fundamental, relation that is widely recognised and properly understood by virtually everyone except the person being criticised. In view of the well-known correlations with socio-economic status (SES) that are positive for IQ and negative for official delinquency, and given the theoretical predilections of social scientists, the alternative model most likely to be entertained in this case is one based on SES. For example, social class has recently returned to vogue as an explanation for the plight of the black underclass (Wilson, 1978). There is a sense in which SES could fairly be described as the master variable of sociology, just as culture might easily qualify as the master variable of cultural anthropology. Such master variables are often overworked as explanations. Accordingly, I have undertaken an examination of the empirical basis for supposing that SES rather than IQ might really be responsible for the success of the race-IQ-delinquency model. Such an hypothesis might seem reasonable to criminologists who were unfamiliar with the literatures on status attainment and behaviour genetics or with some of the more penetrating interpretations of the relation between IQ and SES (e.g., Gottfredson, 1985, 1986a).

Results concerning the individual level of analysis have already been presented (Gordon, 1986a). Briefly, correlations at that level are far too low to support the conjecture that IQ is somehow a spurious surrogate for SES at the aggregate level on which the model is tested. Correlations between child's IQ and the classic single indicators of parental SES, such as educational attainment, are only in the low to mid .30s at best. Such low correlations provide no basis for supposing that IQ and SES would be interchangeable in testing the model.

Although it is possible to combine parental background variables in multiple regression equations so as to obtain correlations with child's IQ as high as about .55, the inclusion of some of those variables compromises the operational and theoretical definitions of parental SES. For example, one of the best predictors of child's IQ is mid-parent IQ (the mother-father average), but few theorists would want to equate that predictor with SES. Furthermore, even the best sets of background predictors fare badly when the children are early adoptees rather than natural offspring. In such comparisons, the IQ variance accounted for by a set of eight predictors can drop from 30.9% for natural children to only 7.5% for the adoptees, even though the IQs of legal parents were included among the predictors (e.g., Scarr and Weinberg, 1978: Table 4). Much of the usual within-race correlation between parental background and child's IQ, therefore, seems due to shared heredity rather than to shared SES. The impression that the connection between IQ of child and parental social class or SES is strong derives from the practice common among sociologists of presenting data in aggregate form and of analysing those data by non-correlational methods, for example, examining mean IQs by parental social class, as White's (1982) review and meta-analysis showed. In contrast to sociologists, behaviour geneticists have concluded that "we know of no specific environmental influences nor combinations of them that account for as much as 10 per cent of the variance in IQ" (Plomin and DeFries, 1980: 21-22). This statement covers the subset of influences represented by SES.

On the aggregate level, the linear correlation between parental SES and measures of children's cognitive achievement has been found to average .73, based on 18 examples (White, 1982: Table 5). Although this correlation is much higher than ones at the individual level, it is still too low to establish interchangeability in aggregate models. Furthermore, there are severe complications due to intergenerational regression toward different population means when blacks and whites are both considered. Recent Scholastic Aptitude Test (SAT) data showed, for example, that white college applicants from families with income under $6,000 had slightly higher median scores on both parts of the SAT than black applicants from families with income of $50,000 and over (Arbeiter, 1984). Thus, although the rank correlation between aggregate SES and SAT values is perfect within each race, the same SES value does not translate into the same mental ability score for both races, and by a wide margin. Such findings are typical (for additional examples, see Gordon, 1976: 265).

At the next stage in the causal sequence, I also examined studies in which correlations with delinquency of both IQ and SES could be compared (Gordon, 1986a: Table 1). There was no basis in the results for assuming that SES would be superior to IQ in testing the model, for the correlations between IQ and delinquency were at about the same level as, or higher than, those involving SES and delinquency; this is consistent with the well-known results of Hirschi and Hindelang (1977) concerning the relative strength of these two predictors at the individual level. It was revealing to note that in the one study where the correlation between IQ and delinquency was extremely small, the accompanying correlations between SES and delinquency were even smaller. The smallness of all effects was probably related to the unusual community in which this study had been conducted, evidently a "college town," where the mean IQ was 108.The lesson is clear-one should never draw conclusions as to the relative strength of the IQ correlation unless the SES correlation has been presented simultaneously for the same set of data. In another study, when juvenile and adult arrests were combined, improving reliability, the correlation between intelligence and criminality reached -.35 and clearly dominated that of -.12 with parental background SES (McGarvey et al., 1981: Table 2).

The question of the interchangeability of SES indicators for IQ at the level of aggregate statistics employed in the race-IQ-delinquency model is ultimately an empirical one. Although the individual-level data are useful for establishing that the competing variables are not interchangeable at that level, that is, not perfectly correlated, there remains the possibility that they could still function interchangeably in the model at the aggregate level. For example, although a perfect correlation at the individual level is a sufficient condition for establishing a perfect correlation between means, it is not a necessary one and so the means of two variables could be perfectly correlated regardless of their individual-level correlation. If someone knew that competing variables functioned completely interchangeably with IQ it might justify saying "Of course" even though the more familiar individual-level data provided no basis for such an intuition. This article examines that remaining possibility.

I will exhibit the model as it works when based on IQ and then substitute SES indicators for IQ in order to see how well the same model works on the kinds of variables that have long been central to sociological theory. Where the data permit, I will also explore a variant for the model that does not depend on the assumption of normality for SES. When necessary, I will distinguish between the two versions of the model by referring to them as the *original* and *variant* models, respectively.

The average aggregate level correlation of .73 that White (1982) found between parental SES and children's cognitive performance provides no good hints as to the likely outcome of inserting SES variables into the model, not only because the dependent variable is now different, but also because White's correlations were linear ones, whereas the original model has a non-linear feature related to its dependence on the normal distribution and the variant model is apt to be non-linear too unless SES happens to be distributed rectangularly. Similarly, the import of the relatively poor performance of SES background as a predictor of delinquency in the review by Loeber and Dishion (1983: Table 10) is limited for present purposes because, although their dependent variable was delinquency and hence the relevant one, their results bore only on the question of differences within-race, not between-race. All of these considerations apply also to the implications of Thornberry and Farnworth's (1982) failure to find any correlation between SES background and delinquency; in addition, the within-race variance of childhood SES, which they did not report, may have been reduced by out-migration of middle-class whites from Philadelphia during the period concerned (e.g., see Gordon, 1976: 237-40).

As will be seen, on a few occasions the original model does achieve a good fit when based on SES variables. On other occasions, the variant model leads to a good fit. What interpretation to attach to these occasions will be a matter of concern throughout this article as well as its final topic.

The Race-IQ-Delinquency Model Proper

Prevalence rates to be explained. There are two major sources for representative race-specific prevalence rates of delinquency that refer to a common criterion. Both pertain to entire geopolitical communities. One set, reported by Gordon and Gleser (1974), applies to the criterion of having acquired a juvenile court record in Philadelphia during the period 1949-1954. The other set, reported by Gordon (1973), applies to having been committed to a training school in the United States in 1964. These estimates were derived by the synthetic or hypothetical cohort method, in which age-specific first-occasion rates are summed up through the final age at risk (i.e., to 18.0).

This method offers many advantages over the real (longitudinal) cohort method insofar as estimating prevalence is concerned. It is quicker and therefore more timely, it is less expensive, and it is relatively easy to implement as part of the routine collection of data conducted by any justice bureaucracy. Race-, age-, and sex-specific prevalence rates so generated ought to be essential features of the data produced by all modern criminal justice systems, because such rates provide critical information for assessing policy and interpreting developments over time. The effects of age composition (Farrington, 1986; Gordon and Gleser, 1974: Table 3; Gordon, 1976: 252; Gove, 1985; Hirschi and Gottfredson, 1983; Rowe and Tittle, 1977; Wilson and Herrnstein, 1985: chap. 5) and race composition are so powerful that changes in those compositions over time or from one place to another can render comparisons based on the usual forms of justice system data almost uninterpretable, a condition that opens the door to opportunism of one kind or another whenever policy is debated. It is gratifying to note that a major step in the recommended statistical direction has recently been taken in the United States (Langan and Greenfeld, 1985).

The Philadelphia juvenile court rates reported by Gordon and Gleser (1974) were 17.86% (males) and 3.35% (females) for whites, and 50.86% (males) and 15.82% (females) for blacks. The training school rates provided by Gordon (1973) have been adjusted slightly here so as to reflect a minor refinement in their estimation that has been described elsewhere (Gordon, 1986 (a). The new rates are 1.01% (males) and .23% (females) for whites, and 4.00% (males) and .82% (females) for blacks. These rates yield an average black/white ratio of 3.8:1, and raise the question of why the racial differentials frequently assume magnitudes of this particular order.

Note that testing the model empirically on these prevalence data entails confronting a number of important variations: (a) two delinquency criteria that differ considerably in the average severity of the records they imply, and hence also in the prevalence rates of the juveniles whom they qualify as delinquent; (b) two jurisdictional spheres that differed greatly in homogeneity, as well as in ecological make-up and area; (c) two well-separated points in time (at least one decade); and (d) two sexes, who typically differ drastically in their rates for a particular criterion, as well as in the nature of their offences.

How the IQ model is tested. The model assumes, and thus tests for, some systematic correspondence between the IQ-specific prevalence rates within both races. However, that assumption cannot be tested directly using aggregate rates. Instead, the test is accomplished indirectly, by means of a proxy model that is not at all plausible, but which should be sensitive to the correctness of the model that is really assumed.

The proxy treats the data as though there were a *critical IQ,* below which everyone becomes delinquent and above which no one becomes delinquent. The prevalence rates then represent the proportion of each population that falls below the critical IQ for that population, thereby identifying its critical IQ for the purpose of assessing goodness of fit. If all assumptions of the real model are correct, the proxy model should yield a good fit, even though it must invoke an unrealistic step function for the relation between IQ and delinquency rates.

Goodness of fit is assessed by comparing the critical IQs implied by the prevalence rates of blacks and whites, given their known IQ parameters, which are supplied to the proxy model separately as constants. For the original model, the prevalence rates are simply converted to normal deviates, and those z values are used, in conjunction with the IQ means and standard deviations, to determine the hypothetical critical IQ in each case. To the degree that the critical IQs of blacks and whites for a given criterion of delinquency and a specific sex match, the prevalence rates are said to be commensurate with the IQ parameters (i.e., the difference in IQ distributions) of the two populations in question. For convenience, I refer to this property as *IQ-commensurability* (Gordon, 1980a). It is unlikely that the property of IQ-commensurability would be present in the prevalence rates unless there was some kind of fundamental correspondence, not necessarily amounting to equivalence, between the IQ-specific delinquency rates of blacks and whites. When normality holds strictly for both populations, the original and variant models are equivalent, and so there is no need to distinguish between them at this point.

Implications of recent simulations. Except for the limiting case of the implausible step function in the proxy model, there is no simple mathematical relation between the property of IQ-commensurability and IQ-specific delinquency rates. However, a limited number of recently completed simulations have clarified the nature of the empirical relationship in more general cases, so that two important conclusions can be drawn. First, plausible sets of IQ-specific delinquency rates that are identical for blacks and whites are insufficient to account for the observed near-matching of critical IQs referred to as IQ-commensurability. The nature of the insufficiency is such that IQ-specific delinquency rates must be *higher* for blacks to produce perfect matching of critical IQs.

This first conclusion is important, because it opens the door to a variety of effects that would normally be anticipated on both statistical and sociological grounds. Because of regression to different population means, we would expect the midparent IQ of black parents to average about 5.3 points lower than that of white parents even if their children were matched exactly on IQ (Gordon, 1976: 266-67). Hence, if parental IQ contributes to delinquency over and above the IQ of children, "blacks and whites would have separate probability functions of delinquency over the range of IQ" (p. 268). Allowing for such a possibility is realistic, as I have indicated earlier, because "the low IQ child-low IQ parent dyad is fraught with potential for poor socialisation outcomes" (p. 265; see also Gordon, 1975). On more purely sociological grounds, we would also expect contextual effects conducive to delinquency to be present to a greater degree within the black population, where the density of delinquents is about four times greater than among whites, and where other variables thought to play a role in delinquency are known to be present to a greater degree within the black population, where the density of delinquents is about four times greater than among whites, and where other variables thought to play a role in

delinquency, are known to be present to a greater degree than in the white population (Gordon, 1986a).

The failure of such expectations to be borne out in critical IQs that indicate a "surplus" of black delinquents (i.e., the failure to find higher critical IQs for blacks than for whites) has been puzzling, and had led to the conjecture that such effects either are not as potent as was generally thought or are already reflected somehow by the property of IQ-commensurability (Gordon, 1986a). The simulation results clear up these mysteries by revealing that IQ-commensurability itself entails a built-in "surplus" of black delinquents in the form of higher IQ-specific delinquency rates for blacks.

Second, the simulations indicated that the average difference between black and white IQ-specific delinquency rates required for attaining IQ-commensurability can be substantial in absolute terms, that is, when expressed as the usual percentage difference. The percentage differences at those points in the IQ range where they are maximum can range from about 5% to 33% (judging from simulations thus far), depending on the criterion of delinquency and hence on the level of delinquency concerned. Over all IQ points (from 50 to 140), the averages of such black-white differences amount to about 60% of the maximum values in the more plausible simulations. Thus, the degree of separation to be anticipated between the black and white IQ-specific delinquency functions is not negligible; this fact complicates the task of establishing their mutual correspondence now that the easy case in which the two functions are simply identical must be excluded. Without the prospect of support from sets of IQ-specific delinquency rates that are distinctive by virtue of being the same for blacks and whites, the claim of theoretical meaningfulness for IQ-commensurability must rest heavily on the replicability and generality of that property and on its specificity to IQ.

But for the curious fact of IQ-commensurability, there would be no basis for distinguishing the residual surplus of black delinquents at each IQ point from a "race effect" requiring an explanation in its own right. However, the fact that black-white differences in prevalence of delinquency manage to remain commensurate with differences in IQ parameters between the two populations, in conjunction with the causal priority of IQ over most of its numerous non-genetic correlates, strongly suggests that the newly identified "surplus" of black delinquents is itself a form of IQ effect, but one that is mediated indirectly and perhaps diffusely at the population level rather than only through the IQ of the offender (see Gordon, 1986a). For example, virtually all of the information needed to define expected black-white differences in parental effects, based on regression to different population means, and differences in peer effects, based on the milieus afforded by the density of peers at each point in the IQ range, is contained in the IQ parameters themselves. Any remaining information would depend merely on certain general empirical facts, such as the parent-child IQ correlation and the tendency of people to associate with others whose IQs are not too different from their own.

Thus, the commensurability of delinquency prevalence rates with IQ parameters may be telling us something about the fundamental cause of group differences in delinquency as well as about the global and pervasive manner in which IQ affects populations (e.g., Cattell, 1938). After all, both the proportion of a normal distribution located below a critical IQ and the entire distribution itself are fully determined by the IQ mean and standard deviation. Consequently, it would not be surprising if global properties that were best understood as functions of the full IQ distribution also proved to correspond closely to the proportion lying below a given IQ cutoff. In the case of official delinquency, the discovery of IQ-commensurability is enhanced in the U.S. by the fact that justice system budgets and policies are rendered reasonably uniform for both races because that system is largely the cultural product of just one of them-and the majority race at that. Otherwise, such global differences between populations might also interfere with the comparability of race-specific delinquency rates, for example, through their impact on criminal justice system funding.

Establishing that the property of IQ-commensurability itself entails a surplus of black delinquents at each IQ also reconciles that property with certain empirical observations that did not seem to fit well with the hypothesis that IQ-specific delinquency rates were precisely identical for blacks and whites. First, Hirschi's (1969: Table 17; Hirschi and Hindelang, 1977: Table 1) data showed higher delinquency rates for blacks than whites with mental test scores controlled (see Gordon 1976: Table 8). The simulations reveal that such differences are not necessarily inconsistent with IQ-commensurability. Second, the finding by Hirschi and Hindelang (1977: 574) that "both family status and IQ are independently related to [official] delinquency" within racial groups also becomes more consistent as the IQ model accords a greater role to the surrounding IQ context. Third, negative correlations reported between IQ and delinquency have been weaker for blacks than whites (Hirschi and Hindelang, 1977). Part of that difference is attributable to the simple fact that blacks and whites are concentrated in different segments of the IQ range (Gordon, 1986a). The remainder was traced to the higher IQ-specific delinquency rates of the blacks. Now that a black-white difference in IQ-specific rates is seen to be intrinsic to the property of IQ-commensurability, the remaining difference in correlations no longer represents a potential threat to the IQ-commensurability model. Thus, the empirical generalization represented in IQ-commensurability knits together a number of other, seemingly unrelated, observations concerning the relations between black and white delinquency rates. Such a property merits serious theoretical attention.

The sources of the IQ parameters. Table 1 contains all of the existing major estimates of the black and white IQ parameters required for testing the model. There are now many more, as is evident from the citation dates, than were available for testing such a model when it was first presented (Gordon, 1976: Table 7). Although some minor adjustments have been introduced into the estimates actually used in the present test of the model in Table 2, so as to better reflect the data in Table 1, the effects of those changes on the results are almost unnoticeable (and hence not exhibited here). That this should be so is evident from the fact that there is remarkably little disagreement among the estimates of the black-white difference, even over long periods of time during which many of the social differences between blacks and whites in the United States have narrowed markedly (Gordon, 1980b). For comparison, the last row of Table 1 displays the parameter values used in 1976. Other IQ parameters required by the model, such as the black standard deviation and values for Philadelphia proper (i.e., the Northeast), have been selectively based on values in Table 1 that converge, with an eye also to black sample size, quality of data, and possible artifacts such as, for example, ceiling effects in the test battery of the Department of Defense (DOD, 1982). More details have been provided elsewhere (Gordon, 1986a).

Table 1: Major Estimates of the Black-White Difference in IQ or *g*, Nationally and in the Northeast, and of the Black Standard Deviation, ***Expressed in Units of the General White Standard Deviation***

Source	National data			Testing period	North or Northeast data		
	Mean B-W difference in white SD units	Mean black IQ, S-B Scale a	Black SD/ White SD		Mean B-W difference in white SD units	Mean black IQ, S-B Scale a	Black SD/ White SD
(Shuey, 1966: 503)b	1.04	84.8	-	1921-44	-	-	-
	1.11	83.6	-	1945-65	-	-	-
	-	-	-	1921-65	.87	87.6	-
(HARYOU, 1964: Tables 30 and 38)c	-	-	-	1962	.90	87.0	-
(Coleman et al., 1966; NCES, 1966)d							
Reported	1.10	83.8	-	1965	.97	85.9	-
Reweighted	1.18	82.4	-	1965	-	-	-
(DHEW, 1976: Table 2)e	1.08	84.1	.818	1966-70	-	-	-
(Jensen and Reynolds, 1982: Table 1; Kaufman and Doppelt, 1976: Table 4)f	1.12	83.4	.906	1972	.76	89.3	.870

(DOD, 1982: Tables C-1, C-5)g	1.21	82.0	.795	1980	1.15	82.9	.771
(Wolfgang et al., 1972: Table 4.5)h	-	-	-	1950-62	.97	85.9	-
(Gordon, 1976: Table 7)i	1.12	83.4	.787	1965	.96	86.0	.756

a Whenever the general white SD is available IQ estimates are expressed on the scale of the 1937 Stanford-Binet normative sample, on which whites had a mean of 101.8 and a standard deviation of 16.4 (Terman and Merrill, 1960: Fig. 4).

b The data are based on elementary school children, because such samples are usually the most representative. The regional data are for what Shuey terms "the North".

c I have reweighted the Grade 6 Otis IQ school mean for 20 Central Harlem elementary schools according to their enrollments, and have retained only the 16 schools whose pupils were at least 90% black. Grade 6 data were preferred over Grade 3, because older children usually test more reliably and because the Grade 3 mean was atypically high, i.e., 90.6.

d Based on an averaging of Verbal and Nonverbal ability tests at Grades 6 and 9, which seem to have the less distorted score distributions. The reweighted data reflect a more accurate weighting of regional data for arriving at the national estimate. The regional data are for the Metropolitan Northeast.

e Based on the Vocabulary and Block Design subtests of the WISC, and a national sample of 6,768 adolescents.

f Based on the WISC-R standardization sample, which included 305 blacks nationally and 57 in the Northeast. Note the smaller size of these black samples, particularly in the Northeast. Because of that size, other estimates of the black SD are preferred.

g Based on the AFQT battery and a national sample of 7,831 blacks and whites, 18 through 23 years of age. The race differences may be inflated slightly by ceiling effects that restricted the white SD. The regional data are for Pennsylvania, New York, and New Jersey (Middle Atlantic).

h Based on a large sample of black school children in Philadelphia and knowledge that the Philadelphia Verbal Ability Test yields IQs 10 points too high.

i My earlier estimates, based on data at that time. The regional black mean was rounded-off from the mean reported by Wolfgang, Figlio and Sellin (1972), and adjusted downwards by 10 points (see note h). The black SDs were based on the large Southeastern black Stanford Binet sample of Kennedy, Van de Riet and White (1963: Table 38), and the assumption of homoscedasticity in the North and South. This is probably still the best estimate of the black SD in the North. The national black mean was based on Shuey (1966).

The variable measured in Table 1 is understood to be the general factor, *g*, that represents about 80% of the total score variance in standard cognitive test batteries (e.g., Jensen, 1980: 219). Jensen and Reynolds (1982: 432-33) have demonstrated that the mean difference between blacks and whites on factor scores based on the general factor, expressed in standard units, is identical to the mean difference based on Full Scale IQ of the Wechsler Intelligence Scale for Children-Revised (Wechsler, 1974). Differences between blacks and whites on virtually all cognitive tests on which whites exceed blacks have been shown to be virtually entirely due to g, and to vary as a direct function of the tests' *g*-loadings, thus bearing out a relationship known as "the Spearman hypothesis" (Jensen, 1985; Gordon, 1985). Moreover, all of the classic subtest profile differences that have been found to differentiate criminals from non-criminals on Wechsler tests (for typical references, see West and Farrington, 1973: 92-93) can be shown to be simply a reflection of differences in subtest *g* loadings. The *g* factor scores from various standard batteries composed of different subtests, finally, have been shown to correlate on the average about .94 with one another before being corrected for attenuation and .98 after (Gordon, 1987: 129). Therefore, all of the estimates in Table 1, even though they derive from a variety of test batteries, some not labelled as measures of IQ, really do apply to the same variable, as their consistency itself would suggest. The stability of the black-white difference over time should give pause to anyone who optimistically assumes that long-term gains reported in IQ test raw scores (Flynn, 1984; Tuddenham, 1948) necessarily represent changes in *g*.

The parameters in Table 1 have been expressed in standard units, namely, the white standard deviation, so as to render them properly comparable. Those standard units can be transformed to any IQ scale by multiplying them by the white standard deviation of that scale or whatever standard deviation is appropriate in the case of recent normings. For convenience, and because that scale functions as a standard for many IQ tests, the actual tests of the model in Table 2 are expressed in the IQ metric of the 1937 Stanford-Binet normative sample, where the white standard deviation was 16.4 and the white mean was 101.8 (Terman and Merrill, 1960: Fig. 4).

Testing the model and quantifying goodness of fit. Table 2 summarises tests of the model on the available delinquency prevalence rates and provides data for judgeing the outcomes. The first two columns list the IQ parameters and the critical IQs of the proxy model that are implied by the two sets of delinquency prevalence rates reported above. Table 2 also serves as the paradigm for further tests of the original model to follow, in which IQ is replaced by classic measures of SES. The variant model to be explored departs from the model in Table 2 by obtaining critical SES values analogous to critical IQs directly from observed percentile distributions rather than from the normal integral. Differences between the two models will be discussed further when considering their results.

Table 2: Critical IQs Implied by Four Important Estimates of Race and Sex Specific Prevalence Rates Under the Simplified Assumption that Everyone Under a Certain IQ Becomes Delinquent According to the Criterion of Official Delinquency in Question

	Blacks	Whites	B-W difference: In parameters	B-W difference: In critical IQ	B-W difference: Between group variance	Percentage reduction in between group variance[h]
Philadelphia parameters:						
IQ mean	86.0a	101.8c	-15.8		62.41	
IQ standard deviation	12.5b	16.4c				
Philadelphia, court record prevalence:d						
Boys (implied critical IQ)	86.3	86.7		-0.4	.04	99.9
Girls (implied critical IQ)	73.5	71.8		1.7	.72	98.8
United States parameters:						
IQ mean	83.8e	101.8c	-18.0		81.00	
IQ standard deviation	13.1f	16.4c				
Nationwide, training school prevalence:g						
Boys (implied critical IQ)	60.9	63.7		-2.8	1.96	97.6
Girls (implied critical IQ)	52.4	55.3		-2.9	2.10	97.4
Summary measures of goodness of fit:						
Algebraic mean of critical IQs				-1.1		
Absolute mean of critical IQs				2.0		
Mean reduction in variance						98.4

a Based on mean of 95.9 for Philadelphia blacks (Wolfgang et al., 1972: Table 4.5) and knowledge that the Philadelphia Verbal Ability Test yields IQs 10 points too high, as well as on compatability with other regional estimates.

b Mean of black SD from Middle Atlantic states (DOD, 1982: Tables C-1, C-5) and from Southeastern states (Kennedy et al., 1963: Table 38), which are close to each other in value.

c Based on 1937 Stanford-Binet normative sample (Terman and Merrill, 1960: Fig. 4).

d Based on prevalence rates of Gordon and Gleser, 1974: Table 3.

e Based on frequently reported black-white difference of close to 1.1 white SD.

f Based on most consistent, large sample estimates of black SD as equal to almost exactly 0.8 white SD (which is 16.4). See DHEW (1976: Table 2) and DOD (1982: Tables C-1, C-5).

g Based on prevalence rates of Gordon (1973: Table 2), slightly refined so as to take account of small changes in population between 1956 and 1964, but still using an assumed interval of two months between referral and disposition, which has now been found typical for Maryland (personal communication, Joseph Szuleski, Chief of Research and Analysis, Juvenile Services Administration, 1978).

h The between group variance reflects the use of $N = 2$ as a divisor. This can be justified in either of two ways: first, because a descriptive statistic rather than an estimate is involved, second, because the statistic is actually based on very large N, for which N-1 would be only trivially different from N.

The remaining four columns in Table 2 assess the fit of the model. The simplest statistical comparison is that between the difference in the IQ parameters (i.e., the means), on the one hand, and the difference in the critical IQs, on the other hand. Although it is possible to produce a different but related set of comparisons, based on differences in observed and expected prevalence rates, say, for blacks, given the IQ parameters and the critical IQ implied by the white prevalence rate, I have from the outset chosen to base the comparison on critical IQs. The reason for this choice was that IQ was far more likely to represent an interval scale than the proportions embodied in the rates. New evidence, free of the well-known difficulties afflicting these issues, that the latent trait measured by IQ tests is indeed normally distributed in both the black and white populations and that, consequently, IQ is indeed an interval scale has recently been presented (Gordon, 1984). The availability of an interval scale in the form of the IQ metric makes comparing the goodness of fit between sets of prevalence rates that fall at widely separated points in the percentage range more meaningful. Such sets occur whenever they depend on criteria of delinquency that differ markedly in implied severity, for example, training school commitment vis-a-vis juvenile court record.

Ordinarily, goodness of fit is expressed in terms of a dependent variable, with the aim of testing the statistical significance of departures from perfect fit. When samples are huge, however, statistical significance becomes less informative, although the utility of quantified comparisons among different versions of a model remains undiminished. Pairs of critical IQs (or, below, of critical SES values) provide a rational basis for making such comparisons, despite their being couched in terms of the independent variable.

For a summary measure of goodness of fit, one can employ the mean absolute difference between critical IQs, which in Table 2 equals 2.0 IQ points. This represents a slight improvement in fit over the model's first test in 1976, when the corresponding figure was 2.1 IQ points. Virtually all of the improvement is due to the minor refinement introduced into the estimation of the training school prevalence rates that was mentioned above, and not to refinements in the IQ parameters themselves (see Gordon, 1986a).

Table 2 also provides another statistic, the between-group variance, which can be calculated here by taking one-half of the difference between means or between critical IQs and squaring that (implicitly and correctly using N rather than N-1 in the denominator of the standard deviation). The availability of the between-group variance leads naturally to the statistic in the final column of Table 2, which represents the percentage reduction in between-group variance achieved by the model. Recall that perfect fit would be defined as no difference between the critical IQs in any pair. Since in that case the model would have reduced the initial large difference between the black and white IQ means to zero, it is reasonable to take that initial difference, which is not the same for all regions of the U.S., as the point of departure by comparing the final differences in critical IQs with the initial differences in IQ means. The reduction in between-group variance incorporates the comparison within a single familiar statistic; it is especially convenient to be able to express the comparison via a single number when the magnitudes of initial black-white mean differences in standard units vary considerably over the measures of interest, as happens to be the case within the SES domain. In view of that variance in initial differences, however, the two measures of goodness of fit reflect slightly different aspects of the data, and so it is desirable to present both. In practice, they would only rarely lead to divergent interpretations of the results.

Although the mean absolute difference can always be extracted from the tables concerned with SES to follow, this article will focus on comparisons based on the percentage reduction in between-group variance, which has the additional advantage of being transformable (through its square-root, after dividing by 100) to the metric of ordinary correlation coefficients and vice versa. This feature makes it possible to compare goodness of fit based on simple linear regressions between aggregate IQ means and prevalence rates with the fit of the nonlinear model in Table 2. Such simple linear least-squares fittings, when applied to the IQ means and delinquency rates of interest here, do not account for more than 49% of the variance (Gordon, 1986a). In contrast, the four tests reported in Table 2 account, on the average, for 98.4% of the variance. The comparison with simple least squares models also highlights the fact that the model in Table 2 is fitted deterministically rather than statistically, that is, the constants are supplied to the model rather than being obtained from it opportunistically via least-squares solutions.

SES Instead of IQ in Similar Models: Theory

General procedure for the original model. The black and white means and standard deviations of any other variable can be substituted for the IQ parameters in Table 2. One would simply take the z (i.e., inverse normal) transform of each prevalence proportion (rate) as before, multiply it by the standard deviation of the population in question, and subtract (or add, for proportions over .5) the product from (or to) the mean of the population in question. The result would be the critical value of the variable in question, below which, according to the proxy model, everyone becomes delinquent (thereby generating the observed prevalence rate).

There are some important properties of this general procedure that can best be described by a few simple equations. More formally, for a particular independent variable the critical value for whites, Cdw, is found by subtracting the product of z_w *(the normal deviate corresponding to the white sex-specific prevalence rate) and the white standard deviation, S_w, from the white mean, M_w.* A corresponding equation determines the critical value for blacks, C_b. Thus:

$$M_w - z_w S_w = C_w \tag{1}$$

and

$$M_b - z_b S_b = C_b \tag{2}$$

The critical black-white difference, C_b - C_w, can be expressed by subtracting the left-hand side of Equation (1) from the left-hand side of Equation (2). With terms rearranged, this becomes

$$M_b - M_w - z_b S_b + z_w S_w = C_b - C_w \tag{3}$$

Equation (3) can be simplified (i.e., reparameterised) by dividing both sides by the white standard deviation, Sw, in effect restating the two differences in terms of white standard scores analogous to IQ for the variable concerned:

$$[(M_b - M_w)/S_w] - z_b (S_b/S_w) + z_w = (C_b - C_w)/S_w \tag{4}$$

Thus, SES variable differences in the standard score metric can be converted to the 1937 Stanford-Binet IQ metric of Table 2 simply by multiplying them by the white IQ standard deviation of 16.4, yielding a common metric that makes comparing the goodness of fit of SES and IQ variables more convenient. The simplification achieved by Equation (4) is more apparent if the conversion to white standard scores is left implicit (i.e., indicatd by asterisks), as in Equation (4a):

$$(M_b - M_w)^* - z_b (S_b/S_w) + z_w = (C_b - C_w)^*. \tag{4a}$$

Equation (4a) is informative concerning the numerous SES analyses to follow. Testing a succession of SES variables on a particular sex-specific pair of black and white prevalence rates will define an entire family of models. Within such a family, the two delinquency rates remain fixed, but the SES means and standard deviations usually vary. Consequently, the roles of constants and variables in the model are, from this perspective, now reversed. The parameters, i.e., means and standard deviations, that were constants from the perspective of a single IQ model now become variables, and the normal deviates that varied with the prevalence rates now become constants within the particular family of models.

Equation (4a) reveals that differences in the goodness of fit, as defined by the difference between critical values $(C_b - C_w)^*$, are determined within a family of models simply by the standardised racial difference in initial means and the racial ratio of the standard deviations of the variables concerned. Indeed, by regressing the difference $(C_b - C_w)^*$ on the corresponding mean differences and ratios of standard deviations in any such family, one can recover z_w as the regression constant, z_b as the regression weight for the ratio of standard deviations, and 1.0 as the regression weight for the mean differences $(M_b - M_w)^*$, subject only to rounding error. The multiple correlation would always equal 1.0 (within rounding error).

Viewing Equation (4a) in the context of regression, and thereby identifying two of the constants with regression coefficients, enables us (a) to anticipate certain systematic differences among families of such models and (b) to trace those differences to systematic changes in the relative size of the regression coefficients. The true "regression" weight for the mean difference is implicitly 1.0 in all families of model, but the true "regression" weight z_b for the ratio of standard deviations depends on the black delinquency rate in question, which varies from family to family. At a prevalence rate of 15.87%, z_b would equal 1.0, hence, both weights would be equal. However, as the black delinquency rate increased beyond 15.87%, up to 50%, z_b would continue to decrease, and it would remain below 1.0 in absolute value up to a hypothetical rate of 84.13%, after which it would again exceed 1.0 absolutely.

In reality, black delinquency rates here range only from about zero to about 50%, so 15.87% would represent the only point at which the crossover in relative magnitude of the two coefficients actually occurred. Given the rates under consideration, and the fact that all variables are stated in comparable units within any single equation (i.e., units of S_w), the anticipated changes in relative size of coefficients in Equation (4a) reveal that the ratio of standard deviations is more important per unit in most families of models than the standardised difference between population means for determining a good fit as defined by the difference in critical values $(C_b - C_w)^*$. A good fit as defined by the reduction in between-group variance is obviously closely related to the difference in critical values and so is similarly affected whenever the standardised initial difference $(M_b - M_w)^*$ between population means does not vary too much within a family of models. In terms defined above, after divisions by 2 have cancelled, and with asterisks now optional, the proportional reduction in between-group variance, R^2, can be expressed as

$$[(M_b - M_w)^2 - (C_b - C_w)^2]/(M_b - M_w)^2 = R^2. \tag{5}$$

Clearly, if the term $(M_b - M_w)$ does not vary much, variation in R^2 and hence goodness of fit are determined mainly by variation in the term $(C_b - C_w)$, and so the two measures of fit would tend to agree.

A major exception to the pattern of relative importance per unit that usually enables the ratio of standard deviations to be more influential than the difference between means occurs in the case of the juvenile court rates for boys in Philadelphia, where the high rate of 50.86% for blacks yields a z_b of almost zero. In that special case, with z_b effectively nullified, variation in goodness of fit within the corresponding family of models is determined almost entirely by variation in the standardised difference between black and white means, and the goodness of fit of any single model will depend almost entirely on the degree to which that standardised mean difference approaches a value that is optimum for fitting the rates of black and white boys.

Since the coefficient z_b of the ratio of standard deviations increases in absolute size as the black prevalence rates decrease from 50%, two conclusions follow. First, since prevalence rates for girls are always much lower than prevalence rates for boys for a given criterion of delinquency z_b will be much larger for girls than boys. Second, since prevalence rates for more severe criteria of delinquency are always lower within sex than prevalence rates for less severe criteria, z_b will be larger for more severe criteria.

The ultimate significance of these two conclusions stems from the fact that the coefficients also multiply any discrepancies from the hypothetical values of $(M_b - M_w)^*$ and (S_b/S_w) that would afford an optimum fit for both sexes or for two different criteria of delinquency. However, since the coefficient of $(M_b - M_w)^*$ is always 1.0, the effects of any such discrepancies from the optimum value of $(M_b - M_w)^*$ do not actually impact on goodness of fit in a multiplicative manner. Instead, discrepancies in the mean difference lead to constant errors that are unaffected by the size of the rate in question, and so they have the same effect on both sexes and on all criteria. The coefficient z_b, on the other hand, ranges from about zero to almost 3.0 in absolute value for the data in hand. As that coefficient reaches 1.0, it makes the goodness of fit depend equally on one more parameter, the black standard deviation, and as it ranges beyond 1.0 to 2.0 or 3.0 discrepancies from the optimum value of that additional parameter are multiplied, several fold, to spoil the fit. (The white standard deviation, of course, is already implicitly present elsewhere in Equation 4a.)

These considerations lead finally to two simple generalisations that are useful for understanding certain patterns in the goodness of fit: A good fit will usually be harder to achieve within a family of models pertaining to girls than within the corresponding family that pertains to boys and will also be harder to achieve for more severe than less severe criteria of delinquency. These are simply corollaries of the greater difficulty that the model encounters in fitting lower than higher prevalence rates (within the existing range).

The two generalizations are not perfect, however, because compensating effects can occasionally occur in which discrepancies from the hypothetical value of $(M_b - M_w)^*$ and (S_b/S_w) that would provide an optimum fit for both sexes or for two criteria of delinquency offset each other so as to benefit one sex or one criterion, which in some cases would be the girls or the less severe criterion. The paired families of models pertaining to the sexes for a particular delinquency criterion share the same values of $(M_b - M_w)^*$ and (S_b/S_w), and hence experience the same discrepancies from the hypothetical values that would yield the best overall fit. This sharing would hold within sex across different criteria of delinquency only when those criteria applied to populations (e.g., jurisdictions, times, etc.) that possess exactly the same IQ or SES parameters, which is *not* the case here. Consequently, the generalisation concerning the relative ease of fitting less severe criteria of delinquency is complicated further for the present data, although one would still expect it to hold to some degree. Both generalisations should be kept in mind when interpreting the outcomes of testing the model using different SES variables.

The discrepancies from optimum parameters mentioned above are usually not sampling errors, because samples were huge; rather, they mainly represent real differences between the observed values and whatever values would provide optimum fit, given the prevalence data. In other contexts, when the prevalence rates are not viewed as given for the purpose of comparing SES with IQ in the model, those rates need not be considered immune to sampling error and other sources of disturbance (see Gordon, 1976).

Possible objections and formal theoretical concerns. The objection could be raised to the original model that, unlike IQ, a new variable might not be distributed normally. This objection is not as telling as it might seem, because the normal distribution of a total test score such as IQ is no guarantee of the essential requirements that the latent trait it measures is also normally distributed and vice versa (see, for example, the discussion of arguments on this point in Gordon, 1984 or Jensen, 1980). Such a guarantee would hold only if IQ was already known to be an interval or a ratio scale.

One could maintain, with some force, that SES variables are observed directly and hence that no latent trait is involved, thereby converting the question of normality into one to be settled by direct observation of the SES variable itself. Even more forcefully, one could argue the exact opposite, namely, that SES is understood to be a latent trait or unobserved variable, in which case the parameters of the indicator variable could conceivably serve as reasonable estimates of the mean and standard deviation of the unobserved or latent variable, which might then be normally distributed even though the indicator itself does not seem to be. Such a lack of correspondence between distributions of the trait and its indicator could occur if the indicator is not an interval or a ratio scale, or because the indicator does not have sufficient range to express the underlying distribution. For example, in 1976 I based estimates of the black and white IQ means nationally and in the Northeast on ability test data derived from the Coleman Report (Coleman et al., 1966). Those estimates proved to be quite adequate, as Table 1 shows, even though the data departed conspicuously from normality due to floor and ceiling effects. Further, standard deviations based on the Coleman Report's data would also have provided serviceable estimates.

The two viewpoints concerning the theoretical position of SES both have merit, but the cogency of each depends mainly on whether SES is regarded as a manifest variable or as a hypothetical construct. Practically all theoreticians, with the possible exception of Bogue (1969: 436), treat SES as though it was not simply equated with the observed indicator, particularly those who view SES as a subjective assessment or prestige rating imposed on objective data (e.g., Coleman and Rainwater, 1978; Reiss, 1961). Accordingly, the question remains as to whether the parameters estimated from manifest SES data do enable the model to achieve a good fit, given the prevalence rates. This question can only be settled empirically-which is where we began. In this sense, SES variables are no worse off than IQ, since both must be empirically tested in the model. If the original model should provide a consistently good fit for SES, that in itself would be an indication that the latent variable is distributed normally. That, in fact, was the inference to be drawn from the success of IQ in the model even before more recent evidence concerning the interval scale properties of IQ became available (Gordon, 1984; 1986a).

However, "SES" often is used to label correlates of social class without necessarily either invoking or rejecting the notion of a latent prestige continuum. According to that usage, an "Of course" reaction inspired by hypotheses concerning SES would have to refer to the effects of, say, income or education directly. (It is difficult to see how the reaction could refer to combinations of both without raising the issue of the latent continuum.) Such a totally pragmatic viewpoint must be considered seriously, because it is the one that typically underlies governmental interventions, which often assume that increasing income or education *per se* will affect a dependent variable such as delinquency. In this case, the observed lack of normality of such variables might seem to disqualify them as being suitable for the model *ab initio;* it would also disqualify them as sources of inspiration for an "Of course" reaction unless the criticism was based on the presumption that normality was not required, but only a fortuitous combination of parameters such as was likely to be found in the case of SES variables. Such chance combinations may indeed occur. However, in the absence of underlying normality it is unlikely that a good fit could be maintained by chance as prevalence rates vary, for example, over sex and over different criteria of delinquency.

Testing the original model using SES variables affords us the opportunity to see just how easily an excellent fit can be attained. Conceivably, many combinations of means and standard deviations could lead to large reductions in the between-group variance, so that an excellent fit of the model in any small number of trials might be merely fortuitous. Or, if not accidental, perhaps an excellent fit is virtually assured whenever the parameters belong to a variable on which the white mean exceeds the black mean, as often occurs within the broad family of correlates of social class. These questions can best be set to rest by accumulating experience from a large number of trials.

Testing the variant model, which obtains critical values directly from observed percentile distributions instead of through the inverse normal transformation, enables us to determine if the fit for SES variables benefits from relaxing the IQ model's assumption of normality. Conceivably, obtaining commensurate critical values and thus a good fit by coincidence is easy whenever the critical values correspond correctly to the percentiles indicated by black and white delinquency rates. If so, errors in determining critical values introduced by imposing the inverse normal transformation on a variety of non-normal SES variables could obscure that fact by spoiling the easy fit. Such errors would provide IQ with an artificial advantage, due only to the fact that its critical values can be determined accurately by either method.

Of course, coincidence need not be the only basis for a good fit, and so other possibilities will have to be considered if the variant model should improve the fit of SES variables so that it rivals the fit of the original model for IQ. However, it is difficult to imagine that a good fit of the same prevalence rates to both normal IQ distributions and non-normal SES distributions through percentiles could be sustained indefinitely over a wide range of delinquency rates, coincidentally or not. By the same logic, one would expect that a consistently good fit not due to coincidence would be characteristic of some particular SES variable having a particular distribution rather than of all SES variables with their varying distributions.

Choice of SES variables and their implications. The model requires a continuous independent variable. Of the three classic indicators of socio-economic status-occupation, income, and education (Reiss, 1961: 83)-that requirement rules out occupation, unless occupational status has been expressed in the form of a continuous scale such as Duncan's (1961a) socio-economic index (SEI). Bogue (1969: 431) has argued that occupation is no longer the valid indicator of status it was through the nineteenth century, when society was more unidimensionally ordered, which would imply that the loss of such information for present purposes may be unimportant. Quite the contrary, according to Bogue, occupation is now a defective indicator of social status, while income and education are the most powerful stratifying forces. Prestige, to Bogue, is "nothing more than a subjective measurement of the income-education components" (1969: 436), except for a few elite statuses found among only the upper 10% of the population. Such a narrow exception would render prestige, to the extent it is independent of income and education, largely irrelevant to black-white differences in delinquency, because prestige would add little by explaining relatively small differences among the already extremely low delinquency rates in that rarefied stratum. Ecological studies indicate that official delinquency is concentrated mainly in the lowermost deciles of the SES distribution (Gordon, 1967). The findings of Coleman and Rainwater (1978: chap. 8) confirm Bogue's assertions concerning the independence of prestige from other indicators of SES in that uppermost stratum.

Bogue (1969: 438) favoured giving income and education equal weights in defining a person's ranking, since they were so highly correlated that differential weighting would usually have little influence on rankings. Duncan (1961a: 124) used age-adjusted percentages of individuals in the upper parts of the income and education distributions of occupations to predict aggregated subjective ratings of occupational prestige and obtained a multiple correlation of .91, where the two empirical regression weights were practically equal. Each of the predictors singly correlated .84 or .85 with prestige. Interestingly, Duncan's correlation of .72 between the income and education characteristics of jobs can be accounted for entirely by the loadings of the two variables on a prestige factor, that is, by the product of the two correlations with job prestige cited above. Thus, occupational prestige (or whatever it reflects) exhausts everything that income and education have in common and they share no other common factor orthogonal to prestige in Duncan's data. Duncan used his regression equation for a subset of 45 occupations with known prestige ratings to create socio-economic status scores for all occupations-his socio-economic index. This index will be employed below in order to tap the occupational aspect of SES.

Thus, the importance of income and education is stressed by all approaches to SES whether based on objective measures or subjective ratings. Yet, by asking individuals what social class meant to them, Coleman and Rainwater (1978: 219-20) found that income and jobs ranked first and second in importance, and education trailed far behind, although its relative importance was doubled if high-income jobs (for which education encounters ceiling effects) were assigned only the weight that their empirical frequency warranted. But they also concluded, in agreement with other researchers (but see Gottfredson, 1985), that education was the most important objective resource for attaining status and income, and so is closely linked to status from a causal perspective, as distinct from the layman's phenomenological one. Actually, Coleman and Rainwater's interpretations are never as far from Bogue's as a simple summary might suggest, because they, too, noted "considerable ambiguity about the role of job or occupation in the American public's attribution of social standing" (1978: 47).

Obviously, no procedure that uses SES variables singly can hope to capture whatever hypothetical advantages they may hold when combined in some optimum way. Nevertheless, there is much consensus that both income and education are deeply implicated in determining SES, and each can be regarded as a purer measure than the other of certain unique aspects of status, as well as a potential cause, in its own right, of pragmatic effects. In view of their high correlation and equal weighting in aggregated level predictions of occupational prestige, the chief disadvantage in not being able to combine income and education consists mainly of a slight loss in validity, in as much as the relative importance of their small unique contributions can still be investigated by employing each of them separately in the model and comparing the outcomes. Comparing Duncan's correlations ($R = .91$ versus $r = .84$ or .85) shows the loss in validity for measuring *occupational* prestige to amount to about 11% of the total variance. Conceivably, the loss is smaller than that when *income* and education are used as measures of prestige not defined *a priori* by occupation.

The unique aspects of status captured by income and education as indicators can be characterised in various ways. Individual income for a population can be viewed as a reflection of the status of its members as either earners or consumers. Earning status embodies the concepts of "economic worth" and "bargaining power," whereas consuming status reflects the concepts of "livelihood" and "standard of living" (Bogue, 1969: chaps. 13 and 14). In contrast, family income blurs the meaning of the income variable with respect to a particular population's status as earning individuals, because comparatively more families in that population may have more than one earner to offset low individual earning status. However, family income is obviously the more accurate index of status as consumers (cf. Bogue, 1969: 404). Depending on which form of status is more relevant to delinquency, one or the other kind of income could provide the better fit.

Educational attainment is principally the main indicator of general technical and cultural status (Bogue, 1969: 434), but it is also the SES background variable most highly correlated at the individual level with parental IQ and, to a lesser degree, with the child's IQ (Scarr and Weinberg, 1978: Table 2). The higher correlation gives educational attainment a potential advantage for serving as a surrogate for IQ.

If there is reason to say "Of course" to the race-IQ-delinquency model because black-white differences in income or education are the real determinants of the model's success, some or all of the tests to follow should yield signs of that fact. Duncan's SEI represents a special case that cuts both ways, and so results based on it must be interpreted within a wider context.

Further comments introducing the variant model. To maintain the distribution-free nature of the variant model, all tests based on it employ medians rather than means as measures of the initial difference in SES between blacks and whites. Consequently, measures of goodness of fit for the variant model are expressed with reference to median differences only (see Equation 5).

Since those same medians are used, as one option, in testing the original model, and are reported in Tables 3 to 5, below, there will be no need to list them again for the variant model. Reporting parallel results from applying the variant model to the SES variables in Tables 3 to 5 can be postponed, therefore, until they have been collected in summary Table 7, below. There, the original and variant models can be compared conveniently. Such a between-model comparison is considered to be the one of main interest concerning the variant model.

Being based on the same medians, initial SES differences between blacks and whites are always the same for the original and variant models. Therefore, it is also sufficient to rely on the reduction in between-group variance alone for comparing the two models in Table 7. Obviously, the other measure of fit, $(C_b - C_w)$, provides no additional information whenever the initial difference, $(M_b - M_w)$, is the same for both models (see Equation 5).

SES Instead of IQ in Similar Models: Data

Education and individual income as determinants of juvenile court record prevalence in Philadelphia. The 1949-1954 delinquency data of Gordon and Gleser (1974) showed little variation between those years in the population-specific prevalence of delinquency. Consequently, data from the 1950 census concerning 1949 male income and 1950 education in Philadelphia can be used to derive SES parameters for blacks and whites that can be used for testing the original model on Gordon and Gleser's prevalence rates. Table 3 presents the income and education parameters and the outcomes of the tests. Additional details concerning exact sources and the manner of deriving the parameters from the census tables are provided in the table itself.

As is usually the case with census data, the highest category of education or income was reported in open-end form (e.g., "all over $10,000"). In order to estimate means and standard deviations for such data, one usually employs the midpoints of the bounded intervals and one assigns a somewhat arbitrary value to the final category on the basis of judgement (e.g., Duncan, 1961b: 142). Education presents less of a problem in arriving at such a judgement, because the upper tail of years of schooling does not extend indefinitely. For another purpose, Duncan (1961b: 142) assigned $20,000 to all 1949 incomes over $10,000, but I have assigned the much lower value of $11,500 to those incomes on the assumption that differences in incomes over a certain level are not especially relevant to differences in delinquency. The percentages of individuals falling into such upper categories are small, so mistakes in selecting midpoints for the open-end intervals do not weigh as heavily into the overall estimates as they might otherwise. The model was also tested using Duncan's $20,000 figure, and in all cases the fit was even worse than that in Table 3.

Bogue (1969: 440) recommended using $20,000 *in 1960* dollars as the ceiling for a prestige scale, on the grounds that it represented a standard of living to which the general population actually aspired. He considered more stratospheric incomes largely irrelevant for most people. Using a value for the highest income category in 1949 that is far below the equivalent of $20,000 in 1960 is, therefore, in keeping with Bogue's approach. For 1949, the equivalent figure was $16,089, according to the

Table 3: Tests of Original IQ-type Models Using Years of Education Completed by Persons Age 25 and Over in 1950 and Using Income for Males Age 14.0 and Over in 1949 in Philadelphia, Instead of IQ, to Account for Race Differences in Prevalence of Juvenile Court Record Delinquents, 1949-54.

	Blacks	Whites	B-W difference: Raw (1)	B-W difference: Raw÷ white SD (2)	B-W difference: Expressed in "IQ" metric[a] (3)
Philadelphia education parameters, in years:					
Observed mean	8.02	9.41	-1.39	-.41	-6.6
Observed median	8.19	9.28	-1.09	-.32	-5.2
Observed standard deviation	3.37	3.43	-	-	-
Implied critical educational background					
Boys (based on means)	8.09	6.25	1.84	.54	8.8
Girls (based on means)	4.64	3.13	1.51	.44	7.2
Boys (based on medians)	8.26	6.12	2.14	.62	10.2
Girls (based on medians)	4.81	3.00	1.81	.53	8.7
Philadelphia male income parameters, in dollars:					
Including males over 14.0 without income					
Observed mean	1,618	2,656	-1,038	-.49	-8.0
Observed median	1,695	2,634	-939	-.44	-7.2
Observed standard deviation	1,301	2,125	-	-	-
Excluding males over 14.0 without income					
Observed mean	2,029	3,111	-1,082	-.55	-9.0
Observed median	2,073	2,937	-864	-.44	-7.2
Observed standard deviation	1,135	1,968	-	-	-

Implied critical income, including "without income"					
Boys (based on means)	1,646	700	946	.45	7.3
Girls (based on means)	315	-1,237	1,552	.73	12.0
Boys (based on medians)	1,723	678	1,045	.49	8.1
Girls (based on medians)	392	-1,259	1,651	.78	12.7
Implied critical income, excluding "without income":					
Boys (based on means)	2,053	1,299	754	.38	6.3
Girls (based on means)	892	-495	1,387	.70	11.6
Boys (based on medians)	2,097	1,125	972	.49	8.1
Girls (based on medians)	936	-669	1,605	.82	13.4

Source. Years of school completed from U.S. Bureau of the Census (1952: Table 65). Income from U.S. Bureau of the Census (1952: Table 87).

Note. To obtain means and standard deviations, years of school categories were coded at the midpoints of their intervals, except for years 12 and 16, which were coded as 12.0 and 16.0, respectively, because persons in those categories are concentrated at the lower bound. Income categories were coded at the midpoints of their intervals too, except for "$10,000 and over," which was coded as $11,500. Excluding from the base persons without income, 1.9% of whites and .2% of blacks were in the highest income category.

a The difference in column (2) was converted to the IQ metric in column (3) by multiplying it by 16.4, which represents the standard deviation of IQ among whites in the 1937 Stanford-Binet normative sample (Terman and Merrill, 1960: Fig. 4).

Consumer Price Index (U.S. Bureau of the Census, 1983: Table 796):

To circumvent some problems in estimating means from such data (which Bogue [1969: 440] considered preferable to the median) all original models using regular census data have been tested a second time with medians serving as the measures of central tendency. There are also dual tests for income that reflect options in the census data concerning the inclusion or exclusion of males without any income.

The education of persons age 25 and over is conveniently limited to individuals who have mainly completed their educations, and largely to those old enough to have had a first child who was in turn old enough by 1950 to have become, officially, a delinquent (i.e, at least 8 years of age). Although the income data apply to individuals age 14 and over and therefore extend further downward in age than the education data, the fraction too young to have had a delinquent child by 1950 must be small, especially when males without income have been excluded as one option. In any case, the presence of young childless individuals with possibly atypical low incomes, probably dependents of others, would not detract from interpreting income data as a reflection of a population's generalised earning status rather than as a measure of its resources for consumption. Bogue (1969: 441), for example, considered "the socio-economic profile for a nation [to be] exactly equal to its overall income distribution."

Table 3 reveals that all differences in critical values for both SES variables in numbered column (1) were positive. Given that critical values for whites have been subtracted from those for blacks, the positive sign implies that as far as these variables are concerned there was always a surplus of black delinquency. That is, the black-white differences in delinquency were much larger than the black-white differences in overall education or income would lead one to expect. Elsewhere (Gordon, 1976, 1986a), I have discussed the theoretical implications of a surplus (or lack of surplus) of black delinquency concerning the role of contextual effects and of other variables thought to contribute to race differences in delinquency. Unlike the SES variables in Table 3, IQ typically left no unexplained black surplus, as can be seen in Table 2, where the mean of black-white (signed) differences in critical IQs was -1.1, that is, negative. The lack of unexplained black delinquency in the case of IQ obviates the need at the population level for additional explanatory variables independent of IQ.

Table 3 expresses the black-white differences in units of the relevant white standard deviation (column [2]), while column (3) converts column (2) to the IQ metric by multiplying entries in column (2) by 16.4 (the white IQ standard deviation). Comparisons between IQ and the SES variables based on differences in critical values are facilitated by this transformation to a common metric. As Table 3 shows, the fit of these SES models was consistently poorer than that achieved by IQ in Table 2.

Some readers may puzzle over the implications of negative entries for income and education that occasionally appear as critical values in Tables 3, 4, 5, and 6. Such entries, on their face, are not an appealing feature. I suggest that they be viewed merely as formal, abstract numbers, because it is possible for the model to achieve a good fit even when the critical values happen to be negative.

Education and individual income as determinants of the prevalence of training school commitment nationwide. We are at a slight temporal disadvantage in testing SES on the training school data for 1964, because much of the necessary SES data must be based on the 1960 census. However, according to arguments and data presented elsewhere (Gordon, 1976; Gordon and Gleser, 1974), population-specific prevalence rates do not fluctuate much if delinquency criteria remain stable. Low rates especially, such as those for training school commitment, may fluctuate very little in absolute value; on the other hand, even a small error in this part of the percentage range tends to have a relatively large impact on goodness of fit (Gordon, 1976: 263).

If, as seems to be the case, prevalence rates for stable criteria really do fluctuate little, certain models based on SES are at a severe disadvantage, because real income (especially when based on family income) and nominal educational attainment have changed drastically in recent decades, and with those changes the relations among the black and white SES parameters have also changed, sometimes substantially (on family income, see U.S. Bureau of the Census, 1979: Table 15). Between 1960 and 1981, for example, the difference between whites and blacks 25 to 34 years old in percentages completing high school declined from 25.3% to 11.0% (U.S. Bureau of the Census, 1982: Table 6-3). However, relative change has not been the case for IQ, as Table 1 shows. Special theoretical efforts would be required to support the claim that SES is relevant to delinquency when delinquency rates seem unaffected even by major changes in SES variables of an absolute as well as of a relative nature.

For example, absolute improvements in SES for blacks and whites can be accommodated by the delinquency model better if SES is construed as a latent trait reflecting relative differences rather than as a manifest trait embodied directly in income or education. Such an interpretation is rendered tenable by the work of Coleman and Rainwater (1978: 72-75), which found indications of invariance in certain relations between education when expressed as a standard score and the income and prestige associated with education, even though major changes had occurred over time in the absolute amount of education involved (see also Blau and Duncan, 1967: 121). (For a theory that might explain such constancies, see Gottfredson, 1985.) However, depending on their magnitudes, relative changes between the standings of blacks and whites could still pose problems for models based on SES, unless delinquency rates changed commensurately.

Table 4 continues to display, with no exceptions, a series of extremely poor fits for education and male income in the original model. It departs from Table 3 by showing not a surplus but a large deficit of black delinquency for the education variable; this reflects the fact that Southern blacks are now included, so that the adult black-white education gap is larger even though the point in time is 1960 instead of 1950. When large enough to escape being interpreted as merely a minor departure from perfect fit, a deficit of black delinquency is as embarrassing to the hypothesis of contextual effects as a perfect fit between IQ-specific delinquency rates of blacks and whites would be if one accepts the etiological significance of the variable concerned. The contextual hypothesis suffers because the deficit throws into question the relevance of other causes that are presumed to be more present among blacks than among whites, just as would a lack of surplus in IQ-specific rates of black delinquency. Consequently, results showing much lower critical values for blacks than whites prove awkward for any theory based on a correlation between the variable concerned and delinquency, because they fail to make a good case for that variable by demonstrating a good fit while simultaneously closing the door to a likely source of compensating explanations involving contextual effects.

Family income in 1964 as a determinant of the prevalence of training school commitment in 1964. Family income for blacks and whites has been reported for only certain years. Fortunately, one of them was 1964. Although family income is not a good index of the SES of the main earners of a population, it does have a special role in certain theoretical formulations concerning delinquency. For one thing, it applies only to the population of individuals residing together who are related by blood, marriage, or adoption, and so it focuses more sharply on households that might contain juveniles. Also, it applies better than individual income to theories that depend on consumption status as a precursor of delinquency. However, family income is especially vulnerable if prevalence for a stable criterion does not change much within race, because real median family income, in 1974 dollars, has increased substantially during recent periods. Between 1947 and 1974, for example, real family income increased by factors of 2.2 for blacks and 1.9 for whites (U.S. Bureau of the Census, 1979: Table 15). In the event that prevalence remained stable, a

Table 4: Tests of Original IQ-type Models Using Years of Education Completed by Persons Age 25 and Over in 1960, and Using Income for Males Age 14.0 and Over in 1959, in the United States, Instead of IQ, to Account for Race Differences in Prevalence of Commitment to a Training School Nationwide in 1964

	Blacks	Whites	B-W difference: Raw (1)	B-W difference: Raw÷ white SD (2)	B-W difference: Expressed in "IQ" metric[a] (3)
National education parameters, in years:					
Observed mean	7.92	10.36	-2.44	-.70	-11.5
Observed median	8.15	10.81	-2.66	-.77	-12.6
Observed standard deviation	3.93	3.47	-	-	-
Implied critical educational background:					
Boys (based on means)	1.04	2.30	-1.26	-.36	-6.0
Girls (based on means)	-1.51	.53	-2.04	-.59	-9.6
Boys (based on medians)	1.27	2.75	-1.48	-.43	-7.0
Girls (based on medians)	-1.28	.98	-2.26	-.65	-10.7
National male income parameters, in dollars:					
Including males over 14.0 without income					
Observed mean	2,176	4,276	-2,100	-.60	-9.9
Observed median	1,729	3,921	-2,192	-.63	-10.3
Observed standard deviation	2,048	3,479	-	-	-
Excluding males over 14.0 without income					
Observed mean	2,587	4,719	-2,132	-.64	-10.4
Observed median	2,254	4,337	-2,083	-.62	-10.2
Observed standard deviation	1,980	3,357	-	-	-

Implied critical income, including "without income":					
Boys (based on means)	-1,410	-3,806	2,396	.69	11.3
Girls (based on means)	-2,740	-5,584	2,844	.82	13.4
Boys (based on medians)	-1,857	-4,161	2,304	.66	10.9
Girls (based on medians)	-3,187	-5,939	2,752	.79	13.0
Implied critical income, excluding "without income":					
Boys (based on means)	-880	-3,079	2,199	.66	10.7
Girls (based on means)	-2,166	-4,795	2,629	.78	12.8
Boys (based on medians)	-1,213	-3,461	2,248	.67	11.0
Girls (based on medians)	-2,499	-5,177	2,678	.80	13.1

Source. Years of school completed by whites from U.S. Bureau of the Census (1964: Table 173). Years of school completed by blacks from U.S. Bureau of the Census (1963: Tables 9 and 19). Income for both blacks and whites from U.S. Bureau of the Census (1964: Table 218).

Note. To obtain means and standard deviations, years of school categories were coded at the midpoints of their intervals, except for years 12 and 16, which were coded as 12.0 and 16.0 for whites and 12.0 and 16.5 for blacks (for whom the latter category was open-ended); the category "17 or more" years of school was coded as 17.5 for whites alone, as no such category was available for blacks in the tables consulted. These decisions reflect the concentration of persons at the lower bound of certain closed intervals such as 12 and 16 years of education. Income categories were coded at their midpoints too, except for "$10,000 and over," which was coded as $13,000 because this code was close to the mean income in 1959 of the "male experienced labor force" within the range from $10,000 to $20,000; as depicted by Bogue (1969: Fig. 14-1), that segment would have a mean of approximately $13,043. Excluding from the base persons without income, 7.2% of whites and .4% of blacks were in the highest income category.

a The difference in column (2) was converted to the IQ metric in column (3) by multiplying it by 16.4, which represents the standard deviation of IQ among whites in the 1937 Stanford-Binet normative sample (Terman and Merrill, 1960: Fig. 4).

theoretician who preferred SES to IQ would have to retreat toward an argument based on relative rather than absolute deprivation, and even then the slightly differential changes for blacks and whites might pose problems.

Table 5 continues to show large positive differences between critical values, implying that original models based on income leave a large surplus of delinquency for blacks that is unexplained by racial differences in the parameters. However, the fit for family income is considerably better than it was for male income in Table 4, although still poor by IQ standards.

Education, income, and Duncan SEI scores for males from the 1962 OCG-I survey as determinants of the 1964 prevalence of training school commitment nationwide. The 1962 survey of "Occupational Changes in a Generation" (known as OCG-I since its replication in 1973) has reported three race-specific measures of aggregate SES background for a point in time, March 1962 (1961 in the case of annual income), that was nearer to 1964 than was the 1960 census. The OCG-I sample consisted of about 20,700 men between the ages of 20 and 64 who were included in households interviewed by the Bureau of the Census in the course of a Current Population Survey (Blau and Duncan, 1967: 10-14). The data analysed here were based on the slightly older age-group of men from 25 to 64 who were native-born members of the "experienced civilian labour force."

This older age-group contained 14,347 whites (including a negligible number of non-black non-whites labelled "Non-Negro Men" in some reports) and 1,394 blacks. For some analyses, parallel data were available separately for the major segment of the sample ("non-farm background") that excluded men whose fathers held farm occupations. Accordingly, paired tests of the model are presented, where possible, in which the sample has been defined by each option concerning the inclusion of men having farm backgrounds in childhood. Because means and standard deviations have been reported directly from OCG data, there is no need to consider medians (which were unavailable) or to be concerned about the values assigned to open-end categories. Since percentile or frequency distributions were not reported, only the original model can be tested with these SES variables.

Table 5: Test of Original IQ-type Model Using Family Income in 1964 to Account for Race Differences in Prevalence of Commitment to a Training School Nationwide in 1964

	Blacks	Whites	B-W difference: Raw (1)	B-W difference: Raw± white SD (2)	B-W difference: Expressed in "IQ" metric[c] (3)
U.S. family income parameters, in 1974 dollars:					
Observed mean	7,290[a]	12,101[a]	-4,811	-.60	-9.9
Observed median	5,921[a]	10,903[a]	-4,982	-.62	-10.2
Observed standard deviation	5,942[b]	7,980[b]	-	-	-
Implied critical family income, in 1974 dollars:					
Boys (based on means)	-3,115	-6,437	3,322	.42	6.8
Girls (based on means)	-6,973	-10,515	3,542	.44	7.3
Boys (based on medians)	-4,484	-7,635	.39	6.5	
Girls (based on medians)	-8,342	-11,713	3,371	.42	6.9

a Mean family income in 1964 dollars was reported for nonwhites and whites in U.S. Bureau of the Census (1979: Table 17). Each mean was converted to 1974 dollars by multiplying it by 1.587, which represents the ratio of Consumer Price Indexes (CPI) for the years 1964/1974 (see U.S. Bureau of the Census, 1983: Table 796). In addition, the mean for 1964 blacks proper, above, was estimated by assigning it the same proportional relation to the 1964 mean for nonwhites that the 1964 black median had to the 1964 nonwhite median family income. These two medians appear in U.S. Bureau of the Census (1979: Table 15; 1965: Table 13). Median family income for whites in 1964 was taken directly from U.S. Bureau of the Census (1979: Table 15), where it was expressed in 1974 dollars. That figure agrees closely with the same median, expressed initially in 1964 dollars, after adjustment by the CPI (cf., U.S. Bureau of the Census, 1965: Table 13).

b The standard deviation for black family income in 1964, expressed in 1974 dollars, was obtained from U.S. Bureau of the Census (1979: Table 15), by assigning the category "$15,000 and over" the mean value of $23,908, which was estimated in turn by dividing the share of total lack income residual to that category by the number of black families in that category. Such a result can only be a rational approximation, at best. Because only 8% of black families fell in that category, the procedure is likely to prove adequate. The standard deviation so obtained differs only slightly, and in the expected direction (it was smaller), from one obtained from the far more detailed categorization of 1964 family income for nonwhites in U.S. Bureau of the Census (1965: Table 13), where only .3% of nonwhites fell into the highest income category. The latter standard deviation, incidentally, expressed in 1964 dollars, can be converted to 1974 dollars by multiplying it by the CPI factor given above. The standard deviation for white family income was obtained from U.S. Bureau of the Census (1965: Table 13), where 1.2% fell in the highest category "$25,000 and over." That category was coded as $27,500. The conversion to 1974 dollars was as described. Similar data, already expressed in 1974 dollars. The conversion to 1974 dollars was as described. Similar data, already expressed in 1974 dollars, in U.S. Bureau of the Census (1979: Table 15), cannot be used for estimating the white standard deviation, because too many families, 28%, fall in the highest income category ("$15,000 and over") reported there. Use of that source leads to a white standard deviation that is far too small.

c The difference in column (2) was converted to the IQ metric in column (3) by multiplying it by 16.4, which represents the standard deviation of IQ among whites in the 1937 Stanford-Binet normative sample (Terman and Merrill, 1960: Fig. 4).

Table 6: Tests of Original IQ-type Model Using Years of Education Completed, Using Individual Income, and Using Duncan's Socioeconomic Index for Respondent's Current Occupation as of March 1962 for OCG-I Native Males in the U.S. Experienced Civilian Labor Force, Age 25 to 64, to Account for Race Differences in Prevalence of Commitment to a Training School Nationwide in 1964

	Blacks	Whites[a]	B-W difference		
			Raw (1)	Raw÷ white SD (2)	Expressed in "IQ" metric[b] (3)
National education parameters, in years:					
Including men with farm backgrounds					
Observed mean	8.34	11.15	-2.81	-.83	-13.6
Observed standard deviation	3.60	3.40	-	-	-
Excluding men with farm backgrounds					
Observed mean	9.35	11.70	-2.35	-.71	-11.7
Observed standard deviation	3.51	3.30	-	-	-
Implied critical educational background:					
Including men with farm backgrounds					
Boys	2.04	3.25	-1.21	-.36	-5.8
Girls	-.30	1.51	-1.81	-.53	-8.7
Excluding men with farm backgrounds					
Boys	3.20	4.03	-.83	-.25	-4.1
Girls	.92	2.35	-1.43	-.43	-7.1
National income parameters ($1,000):					
Excluding men with farm backgrounds					
Observed mean	3.3	7.1	-3.8	-.67	-10.9
Observed standard deviation	2.0	5.7	-	-	-
Implied critical income:					
Excluding men with farm backgrounds					
Boys	-.20	-6.14	5.94	1.04	17.1
Girls	-1.50	-9.05	7.55	1.32	21.7

National socioeconomic index parameters:					
Including men with farm backgrounds					
Observed mean	17.7	39.6	-21.9	-.89	-14.7
Observed standard deviation	15.3	24.5	-	-	-
Excluding men with farm backgrounds					
Observed mean	19.7	43.5	-23.8	-.97	-15.9
Observed standard deviation	16.9	24.6	-	-	-
Implied critical socioeconomic index:					
Including men with farm backgrounds					
Boys	-9.1	-17.3	8.2	.33	5.5
Girls	-19.0	-29.8	10.8	.44	7.2
Excluding men with farm backgrounds					
Boys	-9.9	-13.6	3.7	.15	2.5
Girls	-20.9	-26.2	5.3	.22	3.5

Source. Education and socioeconomic index parameters are from Duncan, Featherman and Duncan (1972: Tables A.1 to A.4). Income parameters are from Duncan (1968: Table 4-2); available only for men with nonfarm backgrounds. The OCG-I sample is described in Blau and Duncan (1967: 10-14).

a Whites included a negligible number of nonwhites other than blacks.

b The difference in column (2) was converted to the IQ metric in column (3) by multiplying it by 16.4, which represents the standard deviation of IQ among whites in the 1937 Stanford-Binet normative sample (Terman and Merrill, 1960: Fig. 4).

Considering the three OCG-I SES variables in Table 6, years of schooling (Duncan, Featherman and Duncan, 1972: Appendix) resulted in a slightly better fit than was the case in Table 4. Excluding men with farm backgrounds improved the fit still further, with the result that, in the case of prevalence for boys, the original model attained its best fit thus far in tests based on income or education. (As shown in Table 7, below, the reduction of variance in this instance amounts to 87.7%; the corresponding result for girls, at 63.2%, is much less impressive.)

Income (Duncan, 1968: Table 4-2) does its usual poor job in the model; once again it increased the between-group difference.

The special advantage of the OCG-I study for present purposes is that it affords an opportunity to employ Duncan's (1961a) socio-economic index in the model. That index was based on the regression of occupational prestige ratings on the education and income characteristics of jobs and so, more than any of the single SES indicators examined thus far, it combines features of all of them in one scale. Because it is a continuous variable, the index also provides a method by which the model can tap occupational SES more directly. In view of its special nature, it will be convenient to distinguish the SEI from the other variables by referring to it as "the index" and to them as "the single indicators" of SES.

Using Duncan's index, we arrive finally at results that rival the fit of IQ in the original model. Of all the SES variables considered thus far, the index led to the best average fit. For men from non-farm backgrounds, it closely approaches, when applied to the girls' delinquency rates, and actually equals, when applied to the boys' delinquency rates, the excellent fit attained for the same prevalence rates of IQ in Table 2. Before considering this result and the SEI further, let us pause for an overview and summary of all tests to this point, including deferred comparisons between the original and variant models.

Summary of results to this point comparing aggregate SES with IQ in the original and variant models. For convenience, Table 7 lists the percentage changes in between-group variance of all available tests of both models on variables in Tables 2 through 6 and classifies those changes according to whether they reflect reductions or increases in that variance. Any single change in variance represents a comparison that is internal with respect to the variable concerned, whereas any comparison between two such changes may be considered external, because it can involve two quite different variables, for example, SES and IQ. Let us consider the original model first, and then compare the variant model to it.

Table 7: Percentage Reductions or Increases in Between-group Variance Resulting from Using IQ, Years of School Completed, Male Income, Family Income, and Socioeconomic Index Parameters in Original and Variant forms of IQ-type Models to Account for Race Differences in the Prevalence of Delinquency

Variable (and table number, where applicable)	Percentage change[a]			
	Original model		Variant model[b]	
	Reduction	Increase	Reduction	Increase
IQ (Table 2)				
Boys (Juvenile Court, Philadelphia	99.9	-	-	-
Girls (Juvenile Court, Philadelphia)	98.8	-	-	-
Boys (Training School, nationwide	97.6	-	-	-
Girls (Training School, nationwide)	97.4	-	-	-
Years of education in Philadelphia in 1950 (Table 3)				
Boys (based on means)	-	77.8	-	-
Girls (based on means)	-	19.0	-	-
Boys (based on medians)	-	284.8	-	456.5
Girls (based on medians)	-	179.9	-	1017.5
Male income in Philadelphia in 1949 (Table 3)				
Boys (including "without income" and based on means)	16.7	-	-	-
Girls (including "without income" and based on means)	-	125.0	-	-
Boys (including "without income" and based on medians)	-	26.6	-	111.7
Girls (including "without income" and based on medians)	-	211.1	N/A	N/A
Boys (excluding "without income" and based on means)	51.0	-	-	-
Girls (excluding "without income" and based on means)	-	66.1	-	-
Boys (excluding "without income" and based on medians)	-	26.6	55.1	-
Girls (excluding "without income" and based on medians)	-	246.4	66.0	-
Years of education nationwide in 1960 (Table 4)				
Boys (based on means)	72.8	-	-	-
Girls (based on means)	30.3	-	-	-
Boys (based on medians)	69.1	-	99.2	-
Girls (based on medians)	27.9	-	99.9	-

Male income nationwide (Table 4)				
Boys (including "without income" and based on means)	-	30.3	-	-
Girls (including "without income" and based on means)	-	83.2	-	-
Boys (including "without income" and based on medians)	-	12.0	N/A	N/A
Girls (including "without income" and based on medians)	-	59.3	N/A	N/A
Boys (excluding "without income" and based on means)	-	5.9	-	-
Girls (excluding "without income" and based on means)	-	51.5	-	-
Boys (excluding "without income" and based on medians)	-	16.3	99.9	-
Girls (excluding "without income" and based on medians)	-	64.9	100.0	-
Family income nationwide in 1964 (Table 5)				
Boys (based on means)	52.8	-	-	-
Girls (based on means)	45.6	-	-	-
Boys (based on medians)	59.4	-	99.8	-
Girls (based on medians)	54.2	-	99.9	-
Years of education for men nationwide in 1962 (Table 6)[c]				
Boys (including men with farm backgrounds)	81.8	-	-	-
Girls (including men with farm backgrounds)	59.1	-	-	-
Boys (excluding men with farm backgrounds)	87.7	-	-	-
Girls (excluding men with farm backgrounds)	63.2	-	-	-
Male income nationwide in 1961 (Table 6)[c]				
Boys (excluding men with farm backgrounds)	-	146.1	-	-
Girls (excluding men with farm backgrounds)	-	296.3	-	-
Duncan's socioeconomic index nationwide in 1962 (Table 6)[c]				
Boys (including men with farm backgrounds)	86.0	-	-	-
Girls (including men with farm backgrounds)	76.0	-	-	-
Boys (excluding men with farm backgrounds)	97.5	-	-	-
Girls (excluding men with farm backgrounds)	95.2	-	-	-

aFor ease of interpretation, percentage changes for original models have been calculated directly from differences expressed in the IQ metric appearing in numbered column (3) of Tables 3, 4, 5 and 6, even though that increases the rounding error slightly. Percentage changes for variant models have been calculated in the same manner.

bResults for variant models are always based on medians. "N/A" indicates "not applicable" because of artifacts.

cFrequency distributions were not reported for these variables, and so variant models could not be applied.

Table 7 shows that the differences between blacks and whites in IQ parameters account, on the average, for 98.4% of the between-group variance after being used to explain black-white differences in the prevalence of delinquency. That much reduction in variance would require a linear correlation of .99. These summary figures provide a convenient standard for comparing the performances of IQ in the model with those of SES.

Of the 38 tests of the original model employing SES variables, 20 *increased* the between-group variance. Note that an increase refers not to the sign of the black-white difference in critical values, but to the absolute magnitude of that difference (and in relation to the original parametric difference in means or medians). Accordingly, in 53% of such internal comparisons SES variables actually *worsened* the fit by increasing the between-group variance. Any reduction at all in the between-group variance was attained in under half the cases.

When the comparison is external to the variable concerned, that is, with the goodness of fit achieved by IQ, SES variables did much worse 95% of the time, and the single indicators of SES did much worse 100% of the time. Even when attention is restricted to the 14 cases in which single indicators of SES led to some reduction in between-group variance, in no instance did they ever approach the excellent fit achieved by IQ in any of its tests. Such success was realised only by Duncan's index, and then only for men having non-farm backgrounds.

If we sift through the various SES variables in Table 7, there is little indication that any one kind of single indicator (i.e., excluding the Duncan SEI) performed consistently better in the model than the others, with the possible exception of family income, which always reduced the between-group variance in 1964. However, that result must be reconciled with the fact that male income usually worsened the fit everywhere.

Family income may appear more consistent than other SES indicators merely because it could be tested less often-and then on the nationwide prevalence rates of commitment to training school only. In the case of those delinquency rates, years of education in 1960 performed almost as well as family income; education achieved an average reduction of 50% in the between-group variance, as compared with the average of 53% achieved by family income. Although the results for education in 1960 were clearly much more variable than those for family income in 1964 (see Table 7), there is little reason to attach special importance to the role of family income in the model in view of these considerations.

Aside from the Duncan SEI, years of education for men in 1962, from OCG-I, produced the best performance with respect to any single set of delinquency prevalence rates by reducing the between-group variance an average of 73% for the training school criterion. However, the success in this case varied considerably once again, mostly between the sexes. One of the four reductions, at 87.7%, represents the highwater mark in goodness of fit for the single SES indicators in the original model. Even this high figure falls distinctly below the *lowest* of the corresponding results for IQ, which was 97.4%.

But for the relatively good fit of the OCG-I education data, the performance of education in the model would have seemed inconsistent and unpromising. Recall that although years of schooling in 1960 consistently reduced the between-group variance concerned prevalence of commitment to a training school nationwide (Tables 4 and 7), the reductions averaged only 50%, and the same variable in 1950 consistently increased the between-group variance concerning the juvenile court criterion of delinquency in Philadelphia (Tables 3 and 7). (Later, reasons for the generally superior performance of education in the case of the OCG-I data will be considered.)

Let us next shift from inspecting for differences among variables in Table 7 to inspecting for differences between statistics. For census data, tests of the model based on estimated means of SES variables appear to do as well as or slightly better than tests based on medians as far as goodness of fit is concerned. The inclusion or exclusion of males without income had no consistent effect. Except for the case of years of education in Philadelphia, SES variables always provided a better fit to delinquency rates of boys than girls.

The sex difference in fit is without substantive significance, however, and was anticipated. It merely reflects the greater difficulty the original model encounters in fitting pairs of prevalence rates as those rates depart from 50%, as was explained earlier. For the delinquency data in hand, that amounts to more difficulty in fitting lower than higher rates, and the girls' rates are in all cases lower than the boys' rates. The corollary expectation that in general it will prove more difficult to fit more severe than less severe criteria of delinquency is borne out only in the case of the four models based on IQ itself. All other such comparisons are apparently too confounded with substantial variations in the actual parameters of the SES variables concerned to enable such an artifact to shine through.

In any case, both logic and experience have shown that low prevalence rates in general and rates for girls in particular are especially challenging to the original model. There definitely is substantive significance, consequently, in the fact that IQ works so well in that model for both sexes as well as for several criteria of delinquency that yield rates widely separated in magnitude. Because prevalence rates for both sexes figure in all tests of the model, the consistency of fit for both sexes affords an additional basis for judgeing the consistency of the model when SES variables are involved, and hence for judgeing the potential causal relevance of the variables themselves.

Now let us consider the results in Table 7 from testing some of the same variables with the variant model. That model is itself subject to certain artifacts that intrude whenever black and white prevalence rates are extremely low and hence close together in value, as they are apt to be for girls and for the training school criterion (where the black-white rates are only .59% apart in the case of girls).

Under these circumstances, one artifact arises, for example, if the census category of males "without income" is included in the analysis. In that case, low prevalence rates often correspond to percentiles located within the zero income category, where all interpolations yield zero as their result, implying a perfect fit. In effect, all low percentiles yield the same critical value. The perfect fit in question is a trivial outcome, obviously, and so instances of it have been indicated in Table 7 by the entry "N/A" ("not applicable").

The same artifact could arise when low prevalence rates represent percentiles falling within a "no education" category. Practically speaking, however, the difference between one year and none of education is greater than that between one dollar and none of income, and so interpolations within the "no education" category have been treated as though that category stood for "zero to one year." However, this policy often leads to a second artifact.

The second artifact is similar to the first, but less severe. It appears when low rates are so closely-spaced that they inevitably lead to percentiles requiring interpolation within the same low census category of income or education for blacks and whites in order to identify critical values. Linear interpolation assumes a rectangular distribution within each category. If the census category is narrow and proportionately well-populated, therefore, small black-white differences between low prevalence rates produce critical values that are nearly identical. Under a rectangular distribution, small differences in percentiles amount to almost no difference at all in the variable, leading to an excellent, although not necessarily perfect, fit. This outcome is exactly opposite to the effect of small differences in low rates on the original model, for which they proved especially challenging because of the exponential nature of the normal density function.

Examples of the second artifact appear in Table 7 as reductions in between-group variance of 99% or more. Despite their impressive size, they are not considered important because they depend heavily on the categoric form in which census data are reported. Clearly, the variant model lacks sensitivity in the lower range of prevalence rates. This insensitivity can be illustrated by entering prevalence rates for blacks into the variant model that have arbitrarily been set equal to those for whites. For such data, there is no black-white delinquency difference to explain, and in the original model the data would yield a poor fit. In the variant model, however, the artificial data reduce fits of 99% or more in Table 7 by less than 1%, leaving a very good fit. *Results based on the higher prevalence rates that are well-separated for blacks and whites should be regarded, therefore, as being more diagnostic of the variant model's utility.*

In Table 7, the remaining results for the variant model are based on such higher prevalence rates. They show that whenever both models increased the between-group variance, the variant model performed even worse than the original model. Only two results, both involving 1949 male income in Philadelphia, indicated improvement for the variant model over the original model, but the fit of the variant model was mediocre even in those cases. According to these results, relaxing the normality requirement does not lead to non-artifactual improvements in fit for single SES variables that are either consistent or, when they occur, impressive.

Disaggregating age and also checking the fit of log-earnings from the 1962 OCG-I and 1973 OCG-II studies. All census variables examined thus far have applied to all ages, beyond some early cut-off age, of the population in question. The OCG-I data were no real exception, as they included ages up to 64, after which retirement often follows. Such inclusiveness is quite appropriate for the purpose of defining the SES of entire sub-populations as SES is usually conceived in discussions of black-white differences.

Although discussions of racial differences seldom do so, it is possible to specify a more detailed model of the way in which SES background relates to delinquency. The SES of some age cohorts of adults, for example, more recent cohorts or those more likely to contain the parents of the delinquents in a given period, could prove more relevant to delinquency than the SES of other age cohorts. Underlying such an age-specific model of the SES-delinquency relationship would be an awareness that SES parameters, especially measures of central tendency, can vary from one cohort to another because of age, cohort, or period effects, and hence variation from those sources could interact with race so as to affect the magnitude of black-white differences in SES and, ultimately, the goodness of fit of IQ-type models. Accordingly, those models might fit better for causally relevant age cohorts of adults than for the entire adult population. Approximately 77% of all fathers of sons born between the ends of World Wars I and II, for example, were concentrated in intervals of age that are only 20 years wide, even when the age cohorts of the sons have intervals that are as wide as 10 years and thus large enough to span the full period of being at risk for delinquency (see Featherman and Hauser, 1978: Table 2.6).

If successful, the class of models loosely specified here would constitute a more sophisticated basis for an "Of course" criticism. Although sophistication tends to undercut the criticism, which depends for its impact on the obviousness of the alternative explanation proposed, controlling for age is worth considering in its own right. An important difference between IQ and SES is that the relations between the black and white national parameters seem relatively invariant over age, cohort, and period for IQ (e.g., see Table 1), whereas they are known to vary for SES in some or all of these respects.

Readers familiar with the practice of transforming dollar income to natural logarithms (log-earnings) may also wonder how that transformation affects IQ-type models. Data reported by age cohort for earnings, log-earnings, and for the Duncan SEI (which has furnished the best fit thus far of any SES measure) from both the OCG-I and OCG-II studies enable us to address these new issues side by side (Featherman and Hauser, 1978: Table 6.1). The justification here for examining the fit of the model by combining 1964 prevalence rates with 1972 and 1973 SES parameters from OCG-II is that, under the tentative assumption that nationwide prevalence rates remained stable until that time, it represents an economical kind of "scouting ahead." The provisional nature of such an analysis needs no elaboration. Quite aside from the assumption of stability in prevalence rates, the OCG-II data provide a useful opportunity for replicating any cohort effects from OCG-I by examining different samples from essentially the same age cohorts 11 years later.

Certain sampling differences between the OCG-I and OCG-II studies must be noted. The most important is that the OCG-II data have not been reported with men having farm backgrounds in childhood excluded. Second, unlike its predecessor, OCG-II did not include military personnel living off base or in family quarters on base (Featherman and Hauser, 1978: 8-9). The first of these differences matters more than the second, because the exclusion of men with farm backgrounds not only led to a further improvement in fit (Tables 6 and 7), but also offered the only example in which the fit was nearly as good for an SES variable as that attained repeatedly by IQ. Although the analyses that follow may be handicapped by the omission of data reported in that form, any benefits from controlling for age should emerge nevertheless through comparing the fit achieved by the entire sample, even if far from perfect, with the fit achieved by specific age cohorts from that sample.

The results from the two national OCG surveys to be examined are based on data reported by the authors of OCG-II, who excluded all military personnel from *both* studies in order to render them more comparable. Consequently, the Duncan SEI parameters of the OCG-I sample employed here are extremely close to, but not necessrily identical with, those listed in Table 6 when men having farm backgrounds were included. The race-specific parameters themselves do not appear in Table 8, but are readily available (Featherman and Hauser, 1978: Table 6.1). Once again, the variant model cannot be applied because distributions were not reported for the OCG data.

The first rows of Panels A and B in Table 8 show summary measures of goodness of fit for the two OCG samples as a whole. Below those rows are corresponding data for four constituent age cohorts, each having an interval 10 years wide. The bottom rows contain the average measures for the cohorts. Comparing the bottom rows with the top rows allows us to determine easily whether there was a net improvement in average goodness of fit with age cohort controlled, as distinct from an interaction between cohort age and goodness of fit, which would be signalled by a much better fit in some cohorts than others whether or not there was a net improvement.

Although not necessarily mutually exclusive, the two kinds of age cohort effect would be suggestive of quite different forms of theory. A net improvement, especially if interaction were absent, would suggest a general effect of age cohort on goodness of fit, whereas an interaction, if interpretable, would suggest the greater relevance of some age cohorts than others to delinquency regardless of whether a net improvement were present. Because one sex of juveniles is easier to fit than the other, it is important to consider both in judgeing the success of the model when used in conjunction with a particular variable. Therefore, the goodness of fit statistics for boys and girls have also been averaged across sex in each row of Table 8. Such averages will be termed *boy-girl means.* It may help to keep in mind that a good fit is indicated by small critical differences and large reductions in between-group variance, i.e., with no plus sign indicating an increase in variance.

Potential replication is gained for three of the cohorts by comparing, where applicable, an age-specific row in Panel A (e.g., 25-34) with the next row down in Panel B (e.g., 35-44). Although earnings for both OCG samples have been expressed in terms of constant 1972 dollars, one does not anticipate as much stability for earnings over the occupational career as for the Duncan index, and hence replication must be judged accordingly. Occupational status of jobs has demonstrated great stability over time (see Blau and Duncan, 1967: 120-21). Therefore, if occupations of individuals did not change much in the aggregate in either content or status over the 11 years between OCG studies, stability of the SEI parameters within cohorts would be expected and those parameters would be equally valid for the stratification systems at both points in time. Under these assumptions, we would have two opportunities to estimate the fit of the SEI in the model for each of the three younger OCG-I cohorts.

Table 8 shows, yet again, that dollar earnings fits IQ-type models so badly as to offer no encouragement. Usually, the between-group variance was increased. The poor outcome was replicated in Panel B, both overall and in its gross pattern over cohorts. The pattern suggests the action of comparable age effects in both OCG studies. Although the least poor fit occurred in the 35-44 cohort in both studies, the changes in between-group variance involved almost no reduction, and so there is really nothing to consider by way of explaining race differences in delinquency in terms of race differences in earnings parameters.

The substitution of log-earnings, on the other hand, usually reduced the between-group variance substantially. In Panel A, the occurrence of the largest reductions in the two older cohorts is more suggestive of coincidence than of any good theory of delinquency. Even then, the best results barely approach the good fit realised by IQ. Moreover, although the best result in Panel A is to be found in the oldest cohort, in Panel B the oldest cohort displays the worst fit. In Panel B, the fit for boys is often excellent, but that for girls is never as good. Across panels, there is no replication in the details of the fit for different cohorts at the same age, and no replication for the same cohort at different ages. If cohort age does interact with goodness of fit for log-earnings, the effect is an inconsistent one at best that does not lend itself to good theory even if attention is limited to Panel B, where the fit of log-earnings was much better on the average than in Panel A. According to the top and bottom rows, neither earnings nor log-earnings showed any net improvement in fit as the result of controlling age in either Panel A or Panel B.

It could be argued that the youth of 1964 somehow managed to anticipate their future earnings more accurately than the 1961 earnings data in Panel A would have permitted, and that their delinquency rates were a response to the forecast. Superficially, such an argument could account for the substantially better fit in Panel B. Although the log-earnings data in Panel B follow the 1964 delinquency rates by eight years, that was too short a time to allow the 8-18 age cohort of 1964 youth to enter even the youngest age cohort of Panel B (25-34). To support such an interpretation, therefore, one would have to read the statistics in Panel B as proxies for those eventually manifested by the invisible 16-26 cohort. Note that a theory based on anticipation departs from viewing SES as a background variable entirely, because even the existing earnings background in 1964 would not have furnished an accurate forecast of 1972 aggregate earnings. The theory would have to ascribe to 1964 youth in general, some as young as 8 years old, the ability to forecast 1972 earnings, even though real earnings and the black/white ratio of median wages or salaries both increased during the interim (e.g., U.S. Bureau of the Census, 1979: Table 30). Since such a feat is usually beyond the capacities of trained economists, an *ad hoc* theory founded on anticipated earnings deserves no further attention on the basis of the data now before us.

Why does log-earnings improve the fit over that of earnings at both points in time, and what does that improvement imply concerning the role of SES? Answers to these questions require some familiarity with properties of the transformation. First, there is a substantial *positive* correlation of .94 between means and standard deviations of the 16 tabled observations concerning race-specific age cohorts when raw earnings are stated in constant dollars (cf. Featherman and Hauser, 1978: Table 6.1). Since whites earned more than blacks, that correlation is reflected in the fact that white standard deviations are also larger than black standard deviations. That, in turn, makes the ratio of standard deviations (S_b/S_w) in Equation (4a) small. Furthermore, because real earnings increased between 1961 and 1972 more for blacks than for whites, the ratio is always somewhat larger for the 1972 data (although still below 1.0), than for the 1961 data, just as the positive correlation would imply. Such a relation between means and standard deviations is typical for income.

Table 8: Goodness of Fit Statistics from Tests of Original IQ-type Models Using Earnings, Log-earnings, and the Duncan SEI from the OCG-I and OCG-II Surveys Within Age Cohorts to Account for Race Differences in 1964 Prevalence of Commitment to a Training School

Age group of OCG men and sex of delinquents	Earnings: Critical B-W difference (IQ metric)	Earnings: Variance reduction (percent)[a]	Log-earnings: Critical B-W difference (IQ metric)	Log-earnings: Variance reduction (percent)[a]	Duncan SEI: Critical B-W difference (IQ metric)	Duncan SEI: Variance reduction (percent)[a]
A. OCG-I: 1961 earnings and 1962 occupational status						
25-64 (entire sample)						
Boys	13.2	+33.4	-7.3	73.7	5.9	83.5
Girls	16.6	+110.5	-10.5	45.6	7.6	71.8
Mean	14.9	+72.0	-8.9	59.6	6.8	77.6
25-34						
Boys	10.3	27.3	-12.1	42.2	4.8	89.0
Girls	12.9	+13.2	-16.6	-8.2	6.2	81.7
Mean	11.6	7.0	-14.4	17.0	5.5	85.4
35-44						
Boys	8.4	.6	-13.3	38.8	5.2	86.5
Girls	10.8	.4	-17.6	+7.3	6.7	77.8
Mean	9.6	.5	-15.4	15.8	6.0	82.2
45-54						
Boys	16.9	+151.9	-5.3	85.3	6.7	79.1
Girls	21.3	+302.2	-8.0	66.3	7.8	71.4
Mean	19.1	+227.0	-6.6	75.8	7.8	71.4
55-64						
Boys	16.6	*179.7	-2.5	95.2	8.2	69.6
Girls[c]	20.7	+334.6	4.8	81.8	11.0	45.3
Mean	18.6	+257.2	-3.6	88.5	9.6	57.4
Mean of 4 cohorts:[b]						
Boys	13.0	+75.9	-8.3	65.4	6.2	81.0
Girls	16.4	+162.4	-11.8	33.2	8.2	67.1
Mean	14.7	+119.2	-10.0	49.3	7.2	74.1

B. OCG-II: 1972 earnings and 1973 occupational status

25-64 (entire sample)						
Boys	12.3	+138.0	-1.4	97.5	3.9	87.4
Girls	14.1	+211.8	-4.0	79.3	3.6	88.9
Mean	13.2	+174.9	-2.7	88.4	3.8	88.2
25-34						
Boys	8.0	+25.2	-.4	99.6	4.1	79.4
Girls	7.9	+21.2	-3.8	71.8	3.2	87.4
Mean	8.0	+23.2	-2.1	85.7	3.6	83.4
35-44						
Boys	8.6	7.0	-2.0	96.5	3.1	91.9
Girls	9.4	+10.1	-4.7	80.8	2.6	94.5
Mean	9.0	+1.6	-3.4	88.6	2.8	93.2
45-64						
Boys	15.1	+205.9	-.8	99.4	4.1	89.6
Girls	18.2	+343.2	-3.1	89.8	4.6	86.8
Mean	16.6	+274.6	-2.0	94.6	4.4	88.2
55-64						
Boys	15.4	+257.8	-4.0	82.8	6.4	75.9
Girls	18.3	+409.0	-7.6	37.1	7.9	63.3
Mean	16.8	+333.4	-5.8	60.0	7.2	69.6
Mean of 4 cohorts:[b]						
Boys	11.8	+120.5	-1.8	94.6	4.4	84.2
Girls	13.4	+195.9	-4.8	69.9	4.6	83.0
Mean	12.6	+158.2	-3.3	82.2	4.5	83.6

Source: Featherman and Hauser (1978: Table 6.1).
aA plus sign indicates an increase in the between-group variance.
bFor critical differences, the algebraic and absolute means are identical in this table.

Second, the transformation to log-earnings introduces instead a substantial *negative* correlation of -.62 between the means and standard deviations. As a result, the standard deviation of log-earnings always declines from OCG-I to OCG-II, and it is always smaller for whites than blacks rather than vice versa. This reversal of the size of the two dispersions often doubles the value of the ratio (S_b/S_w), which now attains values greater than 1.0. Related changes in $(M_b - M_w)^*$ are less pronounced, although the trend is definitely toward an increase in absolute value because of the smaller divisor, S_w.

As Equation (4a) showed, because of its greater coefficient the ratio of standard deviations is always more critical to a good fit than the mean difference when delinquency rates are as low as those for the training school criterion that is now operative. Moreover, we have just seen that the major impact of the log-earnings transformation is on that more critical term in the equation, and so any improvement in fit must be attributed primarily to that term. Since by IQ-model standards the term $(M_b - M_w)^*$ for log-earnings is always too small (as is the case for all SES variables), the increase in (S_b/S_w) helps to compensate for the deficiency. Therefore, the acceptability of the log-earnings transformation as the proper indicator of earnings status hinges mainly on the acceptability of its effects on the standard deviation.

The fact that the log-earnings transformation drastically reduces the standard deviation in relation to the mean (i.e., lowers the coefficient of variation) is not objectionable. There is even an advantage from that effect: Critical values no longer assume the unappealing negative values that they often displayed for income in Tables 3 to 6, since zero now lies many standard deviations from the mean. The standardised mean differences are, of course, increased somewhat by the proportionately smaller devisor, but not enough to compromise their intuitive meaning. The order of black and white means is always maintained, for example.

However, accepting log-earnings as the metric for earnings status simply because it generally fits the model better than dollar earnings and occasionally provides an excellent fit for boys requires that we also accept the interpretation that the variance of income status is greater among blacks than whites. This challenging interpretation flies in the face of the fact that for dollar earnings and the SEI the standard deviations are always much greater for whites than blacks in both OCG surveys (Featherman and Hauser, 1978: Table 6.1), which is typical. The quite opposite ordering of the dollar earnings and SEI standard deviations for blacks and whites also happens to be the same as that of the IQ standard deviations in Table 1. Thus, the relative dispersions of SES usually parallel those of IQ when blacks and whites are compared.

One rejoinder might be that there is another SES precedent that is more supportive: Except for the very oldest age cohorts, the standard deviations of years of schooling were always greater for blacks than whites within both sexes in Philadelphia in 1950 and in the nation in 1960 (U.S. Bureau of the Census, 1952: Table 65, 1964: Table 173). However, the relative magnitudes of black and white standard deviations may have been subject to an artifact of years of schooling, because the variance of educational attainment for the more recent cohorts of whites has been constrained by ceiling effects on the amount of schooling that is possible.

Drawing inspiration from the example of education nevertheless, one might contend that earnings status is itself subject to a natural ceiling effect because status ceases to increase with income after reaching a certain point. Such a view would be consistent, for example, with Bogue's (1969: 440) recommendation that the 1960 equivalent of $20,000 be regarded as the ceiling for a prestige scale. The 1972 equivalent would be $28,210 according to the Consumer Price Index. The highest white mean would be only 1.9 standard deviations below Bogue's ceiling in the earnings metric and 1.4 standard deviations below in the log-earnings metric. But other means, especially those for 1961, would be somewhat further from that ceiling, and so it is difficult to say whether all of them would be subject to any ceiling effect. Coleman and Rainwater (1978: 35) found no indication of a fixed ceiling up to

$360,000, although subjectively assigned status did increase with 1971 income at a clearly decelerating rate. The log-earnings scale captures some of the characteristics of income status as observed by Coleman and Rainwater, because it contracts as a fixed amount of dollars is added to each point in a distribution of incomes, but the scale does so incidentally rather than in conscious emulation of any empirically-derived function. Whether log-earnings would perform better or worse in the model than such a function is itself a question that can only be settled empirically, and the necessary data are not at hand.

Whatever one makes of log-earnings, the following facts should be kept in mind. First, the fit to the model was far from being equally good at the two different points in time, which would raise questions about any explanation. Second, the fit was better for the 1972 data, which means that it depends on the unverified assumption that 1964 delinquency rates held in 1972. Third, the fit was truly excellent in only two of four cohorts, and then only for boys. Thus, log-earnings never met the more *challenging demand of fitting the model successfully for both sexes simultaneously.* Its great superiority over earnings in the model should not be mistaken for an equivalence to IQ, therefore.

The remaining results in Table 8 are based on the Duncan SEI, whose performance was clearly the best and most consistent over cohorts and sexes in both panels. In general, the level of fit was as good as that achieved by the SEI in Table 7 when men having farm backgrounds were included, where the average reduction in between-group variance for boys and girls amounted to 81.0%. For men 25-64 in Table 8, the corresponding averages were 77.6% in Panel A and 88.2% in Panel B (see the top rows). Although none of these percentages approaches those for IQ, they are of special interest because they consistently represent a better fit than that achieved by any other SES variable, with the single exception of log-earnings in Panel B. (That exception was a qualified one, because it was not sustained in Panel A and it matched up the wrong years of delinquency and log-earnings.) Consequently, it is reasonable to suspect that Duncan's index differs from the other SES variables in some theoretically relevant way. Before considering why the fit of the SEI was generally superior, some final details must be attended to in Table 8.

In Table 8, the poorest fit for the SEI occurs in the oldest cohort in both panels, although different historical cohorts were involved each time. There is also replication of the fit for the same cohorts 11 years later, in that all three maintain the same rank order of variance reductions for both boys and girls in both panels. The three pairings over time also match fairly well in mean fit; the standard deviation of the difference in mean boy-girl fit between Panels A and B for the three replicated cohorts is only 3.1%. Although small, that figure remains larger than the entire range of fit for IQ models in Table 7, which was 2.5% Thus, the IQ models showed more consistency than the SEI cohort replications, even though the latter were relatively successful as SES variables go. The evidence for replication qualifies the nature of the interaction between cohorts and goodness of fit for the SEI, in that the interaction tends to be specific to cohorts rather than to particular ages or periods of OCG study. Again, there were no signs of a net improvement in fit as the result of controlling age. Finally, in Panel B the fit, atypically, was often better for girls than boys.

Explaining the remaining better results: The IQ-surrogate hypothesis. Let us examine further the two untransformed SES variables that have proved most consistently successful in the original model, namely, years of schooling for men in 1962, from OCG-I (Tables 6 and 7), and the Duncan index (Tables 6, 7, and 8). Stipulating that the schooling was for men establishes these two as the most consistently successful SES variables because, unlike log-earnings in Panel A, neither had a relatively unpromising set of trials. True, the trials for years of schooling for men were few in number, but that does not alter the description given; it merely qualifies that description. If we can understand the reasons for the better performance of these two rather straightforward SES variables in IQ-type models, it may shed light on the relation between black-white differences in SES and black-white differences in delinquency in general.

A puzzling aspect of years of schooling for OCG-I men in 1962 is that the variable was far more successful in tests of the model than years of schooling for both sexes together either in Philadelphia in 1950 or nationwide in 1960. The greater proximity in time to 1964 delinquency rates of the OCG-I data in comparison to the 1960 census data seems too slight to account for such a big difference, because a population's education parameters change slowly, mainly through replacement. The fit may also have benefited from having OCG-I means and standard deviations directly available, but if that is the key advantage it holds no theoretical message. A third possibility is that the model benefited from excluding men with farm backgrounds in half of the tests; this hypothesis has merit, but it must be far from the whole story, because even when those men were included the fit was well above average in the case of the OCG-I education data, which then produced a mean boy-girl reduction of 70.4% in the between-group variance (see Table 7).

A fourth, and ultimately most interesting, explanation is that restricting the measure to years of schooling for men alone provided the main advantage. This hypothesis was tested on the 1960 education data in Table 4 by disaggregating the sexes and reapplying the model to the education parameters for adults of each sex separately. Whereas the average reduction in between-group variance with sexes combined had been 50.0% (Table 7), it rose to 61.3% in the case of men's education but fell to 38.1% in the case of women's education. These results represent boy-girl means averaged over analyses based on means and medians. Perhaps it is no coincidence that the average, weighted by sample size, of the two sex-specific percentages is 49.3%, which is practically equal to the 50.0% obtained with the sexes combined. The average reduction in variance of 61.3% achieved by the men's 1960 education parameters now surpasses that of 53.0% achieved by family income in 1964, thus diminishing the relative importance of raw income's most successful fit even further.

Clearly, there is a big difference in the degree to which aggregate educational background fits the model, depending on which sex is used to estimate the parameters of years of schooling. Although the average variance reduction of 61.3% for 1960 men does not quite equal the overall average level of 73.0% achieved with data for OCG-I men in 1962, or even the 70.4% level achieved when attention is restricted to OCG-I data including men with farm backgrounds (the more appropriate comparison), the percentage difference between the two sources was at least halved in each case when 1960 years of schooling was based on men alone rather than on both sexes.

Accordingly, similar analyses reflecting disaggregation by sex were applied to years of schooling in Philadelphia in 1950. Here, of course, the prevalence rates to be fitted refer to the juvenile court criterion of delinquency. As Table 7 showed, with the sexes combined years of schooling increased the between-group variance in Philadelphia and hence provided an extremely poor fit. After disaggregating the sexes, the same variable performed somewhat better when based on men and considerably worse when based on women. The exact figures need not concern us, except to say that now the male data actually reduced the variance on several occasions while increasing it less on the remaining occasions (of means and medians).

Even though some of the results never produced an especially good fit, the consistent improvements realised through disaggregating educational attainment by sex raise the possibility that its parameters mimic those of IQ better when this background variable is based on men rather than on women. If that surmise is correct, education would achieve its best results by acting merely as a surrogate for IQ rather than as a measure of parental resources that impact on children in the course of socialisation. Both the surrogate hypothesis itself and the sex-specific education data from which it derives break sharply with the most common socialisation scenario. In that scenario, the combined educations of both parents would index resources available to their developing child better than the education of either parent alone. The hypothesis and its supporting data are also incompatible with a variant of that scenario in which the influence of the mother's education might even exceed that of the father's in view of the greater exposure that children have to mothers than fathers. In view of these considerations, and because there is no apparent reason why the education of men *qua* education should be so much more efficacious in explaining black-white differences in delinquency than the education of women, the IQ-surrogate hypothesis merits further attention.

The most common socialisation scenario is no straw man. It can be illustrated by certain data pertaining to IQ. Among white families, for example, the educational attainments of fathers and mothers correlate about equally with child's IQ in biological as well as adoptive families (Scarr and Weinberg, 1978: Table 2; Mercy and Steelman, 1982: Table 2; Sewell, Hauser and Wolf, 1977: Table 1). Such correlations are consistent with the interpretation that if both parents serve as resources in respect to IQ, they do so equally well. Other outcomes, such as perceived parental encouragement to attend college, college plans, occupational aspirations, educational attainment, and status of first and current occupations present a more complicated picture that involves an interaction with sex of child. For girls, the educational attainments of both parents again correlate about equally well with these outcomes, but for boys that of the father correlates a bit more strongly than that of the mother (Sewell et al., 1977: Table 1). But this interaction does not conform to the results obtained by disaggregating educational attainment either, because goodness of fit improved for prevalence rates of girls as well as for those of boys when educational attainment was based on men instead of women (see Table 9, below). Sometimes the improvement was even greater for the rates of girls than for the rates of boys (e.g., see the first row of Panel D in Table 9).

It was the fact that none of the various socialisation scenarios could account for the advantage of men's education over women's in the model that prompted the surrogate hypothesis in the first place. Additional data that lend independent support to that hypothesis come from some of the same sources. Scarr and Weinberg's (1978: Table 2) data for sets of both biological and adoptive parents show that parental education and parental IQ correlate with each other about .10 more strongly in the case of fathers than in the case of mothers. For correlations of the magnitude concerned, this difference accounts for another 9-10% of parental IQ variance and indicates that educational attainment distributions of men would be better surrogates for IQ distributions than the educational attainment of both sexes combined or of women alone. A quite similar sex difference in the correlation between eleventh-grade IQ and final educational attainment has been reported by Sewell et al. (1977: Table 1). Alexander and Eckland (1974: Table 5) have published regression coefficients that fit the same picture. Thus, differential correlation of education with IQ according to sex held in four samples. Elsewhere, Gottfredson (1984: Table 5) has reported that the median level of schooling of workers in 276 occupations correlated slightly more with the intelligence required in those jobs in the case of men than of women, even after age (representing historical factors relevant to the employment of women) was included in the equation. This difference, too, would be consistent with a sex differential in the *g*-loading of years of schooling.

In one of these studies, Sewell et al. (1977: 10) noting their standard deviation for years of schooling was smaller for women attributed its smallness to the fact that women were more likely to complete high school than men, but also were less likely to continue schooling beyond high school. Since the correlation of educational attainment with IQ can be regarded as an estimate of the former's *g*-loading, the lower correlation and smaller standard deviation for women may simply be two aspects of the same phenomenon, namely, that women's educational decisions have been less governed by differences in *g* than the decisions of men. Among low IQ students, females apparently tolerate the difficulty of schooling better than males up through high school, but among higher IQ students females pursue further schooling less often than males for a different set of reasons. Both sets of causes combine to affect the *g*-loadings of sex-specific educational attainment, as reflected in its correlations with IQ, and also the parameters of education, as reflected in its standard deviations. In Project TALENT, high schools' rates of students going on to college were less loaded on the *g* factor for girls than boys (Humphreys, Parsons and Park, 1979: Table 2).

If educational attainment does fit the model only because it acts as a surrogate for IQ, any important refinement of its role as surrogate ought to lead to further improvements in fit. We have already seen that one such refinement based on disaggregating by sex brought major improvement that would not have been predicted by the usual socialisation scenarios. Controlling for age, as was done in Table 8, ought to represent another such refinement, and hence another opportunity for testing the hypothesis, because mean levels of schooling for blacks and whites have increased steadily since the turn of the century, and that increase can be understood as a series of upward shifts within the educational system of the entire intelligence distribution (Gottfredson, 1985: 157). Such a refinement would correspond to standardising the parameters of schooling in a manner parallel to the way in which IQ standardises children's mental ability test scores for age. The main difference is that in the one case schooling decreases with age (of cohort) and in the other test scores increase with age (of children).

Table 8 showed that controlling for age of cohort does not automatically improve the average fit for SES variables. None of the three variables considered benefited in that way. In the case of years of schooling, the rationale for controlling age is based on the presumed constancy of mental ability distributions over time despite steady increases in the mean level of educational attainment. If IQ, rather than education *per se,* is the real cause of black-white differences on ability tests and in delinquency, therefore, improving the eligibility of education to act as a surrogate for IQ by controlling age should improve the fit of the model in a manner that is pervasive for the age cohorts. That is, the effect on goodness of fit should be one of net improvement rather than one of meaningful interaction. If such a general effect fails to appear at all, it would indicate that IQ may not be responsible for the success in fitting the model that education exhibited in the case of the OCG-I men.

By contrast, the rationale for controlling age in the case of the previous SES variables was quite different. According to that rationale, some cohorts, probably those nearer in age to that of parents of the contemporary adolescents, would prove more relevant to delinquency than others. The effect predicted by such a rationale is one of meaningful interaction with goodness of fit over cohorts rather than one of net improvement, because the improved fit in relevant cohorts would have to be at the expense of the fit in irrelevant cohorts if the average fit was to remain largely undisturbed. Thus, the new test of the IQ-surrogate hypothesis predicts an effect of a specific kind, under circumstances that are also rather specific with regard to the variable concerned (since that effect did not appear when age was controlled while other SES variables were being tested in the model). When testing hypotheses, such specificity is generally considered a strength.

In view of the already demonstrated advantage of basing educational background parameters on the years of schooling of men, it is desirable to maintain that advantage by controlling for sex when testing the IQ-surrogate hypothesis by introducing controls for age. The surrogate hypothesis suggests that both controls should each contribute independently to the goodness of fit of IQ-type models. Thus, holding age relatively constant should add to the improvement in fit already realised from basing education on men. Combining the two controls also affords us an opportunity to examine the effect of controlling sex in greater detail since, if the sex effect has been general over time, it conceivably could appear in every single age cohort.

The outcomes of employing sex and age controls together when testing the model with years of schooling from Tables 3 and 4 are displayed in Table 9. Distinctions in the table between men and women concerning years of schooling should not be confused with distinctions between boys and girls concerning rates of delinquency. As notes in Table 9 indicate, non-whites have been substituted for blacks in the 1960 data in order to retain the same detail in age classifications for both races. The cohorts terminate at age 74 because thereafter the classification becomes open-ended. The age range does not extend indefinitely, therefore, as it did in Tables 3 and 4, and so the fit for the entire age range may not correspond exactly to the results obtained in those tables, quite aside from the switch to using non-whites. In the case of the Philadelphia data, it should be noted that the older cohorts employ a wider age interval than the younger ones.

Table 9 contains evidence in several forms that supports the IQ-surrogate hypothesis for years of schooling. Comparing the top and bottom rows in Panels A and B, where the model is based on the Philadelphia data, shows that controlling age produces very substantial net improvements in fit whether judged by critical differences or by reductions in between-group variance. The conclusion holds regardless of whether educational attainment is based on men or women. This represents strong support for the IQ-surrogate hypothesis.

Comparing the left and right halves of Panels A and B shows that disaggregating sex for years of schooling continues to yield a much better fit for men's than women's education even when age is also disaggregated. With minor exceptions, that conclusion holds in detail within the individual age cohorts. Specifically, the fit is better for men's education in 23 out of 28 possible pairs of observations, which is significant according to a one-tailed sign test ($p < .03$) after reducing the observations by half in recognition that means and medians are not independent. Many of the exceptions involve ceiling or near-ceiling effects on the goodness of fit, where there was little opportunity for men's education to improve on a good fit attained from usihg women's education. The men's and women's averages over cohorts in the bottom rows offer a convenient summary of such details.

On the average, the model seems to do a better job in Panels A and B when based on medians instead of means, but this impression is partly due to a few cohorts that had very large increases in variance when means were employed. In Panels C and D, means and medians perform about equally well, and the slight difference lies in favour of means most of the time.

When we turn to panels C and D, we find that the two measures of goodness of fit do not seem to agree as to whether controlling for age produces a net improvement in the 1960 data. The bottom row of Panel C, based on critical differences, shows a mixed set of outcomes, with two net improvements, one tie, and five average fits that are slightly or somewhat poorer than those in the top row.

In Panel D, on the other hand, five of the mean reductions of between-group variance in the bottom row reveal non-trivial net improvements in fit, and the remaining three cases portray only trivial changes in either direction. The overall net result in this panel, consequently, is one of improved fit due to controlling age.

Which statistic merits greater weight? Recall that the between-group variance was introduced for several reasons, one of which was that it took into account the initial difference in black-white means and hence evaluated the critical difference in relation to where it was coming from. This was thought to be an especially desirable feature when initial differences varied widely over a series of variables or over multiple tests of the same variable; under these circumstances critical differences of the same size may represent quite different reductions in the between-group variance. When critical differences are equal, the ones linked with greater initial differences will reflect the larger reductions in between-group variance (see Equation 5, above). When initial standardised differences do not vary from one test of the model to another, the critical differences alone will provide an adequate basis for making comparisons. In Panels C and D, the initial black-white standardised differences were indeed somewhat larger on the average for the individual age cohorts than they were for the entire 25-74 age group in the top rows. If expressed in the IQ metric,

Table 9: Goodness of Fit Statistics from Tests of Original IQ-type Models Using Years of School Completed in 1950 and 1960 by Men and Women Within Age Cohorts to Account for Race Differences in Prevalence of Philadelphia Court Record and Nationwide Training School Delinquents

Age	Education of men				Education of women			
	Based on means		Based on medians		Based on means		Based on medians	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
A. Philadelphia, 1950: Critical black-white difference (IQ metric)								
25-74 (entire range)	6.9	5.5	8.2	6.7	10.0	8.2	10.6	8.8
25-29	2.8	-1.0	1.8	-2.0	5.1	-1.5	1.7	-4.9
30-34	3.2	-1.1	.5	-3.8	5.6	.3	-.6	-6.0
35-49	3.2	.1	4.2	1.0	7.6	4.2	7.2	3.8
40-44	3.1	.6	5.4	2.9	6.2	3.1	8.7	5.6
45-54	5.5	3.6	7.6	5.7	8.1	6.4	8.9	7.3
55-64	9.0	8.3	7.2	6.6	10.0	9.6	7.6	7.2
65-74	7.1	6.9	4.0	3.8	7.4	7.1	3.4	3.1
Absolute mean of 7 cohorts:	4.8	3.1	4.4	3.7	7.1	4.6	5.4	5.4
B. Philadelphia, 1950: Between-group variance reduction (percent)a								
25-74 (entire range)	33.2	58.4	+29.3	12.6	+231.0	+125.0	+366.3	+225.3
25-29	95.2	99.4	98.3	97.9	76.4	98.0	98.5	87.6
30-34	93.2	99.2	99.9	93.5	67.5	99.9	99.9	86.3
35-39	92.9	99.9	86.1	99.2	7.5	72.2	23.6	78.9
40-44	93.7	99.7	71.0	91.4	55.0	88.7	+63.8	32.5
45-54	69.8	87.3	6.0	47.6	+19.9	24.4	+88.2	+24.7
55-64	+89.6	+63.4	22.6	35.7	+255.5	+225.1	3.3	14.2
65-74	25.8	30.2	87.7	90.0	13.5	21.5	91.8	93.4
Algebraic mean of 7 cohorts:	54.4	64.6	67.4	79.3	+7.9	25.7	23.6	52.6

C. Nationwide, 1960: Critical black-white difference (IQ metric)

25-74 (entire range)	-6.2	-9.8	-6.6	-10.2	-5.7	-9.8	-9.5	-13.6
25-29	-3.7	-7.1	-3.2	-6.5	-4.1	-8.4	-4.3	-8.6
30-34	-5.0	-8.5	-6.6	-10.1	-6.2	-10.8	-8.0	-12.6
35-39	-7.4	-11.2	-11.0	-14.7	-9.6	-14.6	-13.8	-18.8
40-44	-10.6	-14.9	-15.0	-19.3	-11.1	-16.0	-17.7	-22.6
45-49	-10.2	-14.0	-10.8	-14.6	-9.5	-13.6	-12.1	-16.2
50-54	-9.0	-12.3	-7.7	-11.0	-8.0	-11.6	-8.2	-11.8
55-59	-7.8	-10.9	-5.7	-8.7	-6.9	-10.1	-4.5	-7.7
60-64	-6.5	-9.3	-6.2	-9.0	-5.0	-7.8	-3.5	-6.3
65-69	-4.0	-6.0	-6.1	-8.2	-3.5	-5.8	-4.1	-6.4
70-74	-2.3	-4.3	-6.0	-8.0	-3.4	-5.8	-5.5	-7.9
Absolute mean of 10 cohorts:	6.6	9.8	7.8	11.0	6.7	10.4	8.2	11.9

D. Nationwide, 1960: Between-group variance reduction (percent)a

25-74 (entire range)	73.9	34.3	72.2	33.3	68.6	8.8	53.9	6.4
25-29	86.6	50.8	90.0	53.1	72.9	+15.5	71.3	+14.9
30-34	79.6	40.9	72.8	36.2	55.2	+36.8	47.5	+30.0
35-39	65.8	22.3	54.3	17.7	30.2	+61.8	22.8	+43.8
40-44	47.0	+5.0	37.4	+3.7	30.0	+44.5	20.7	+28.9
45-49	56.6	17.5	54.9	17.0	52.9	2.5	45.7	1.9
50-54	66.4	37.4	70.6	40.3	66.6	29.2	65.7	28.9
55-59	73.6	49.4	81.3	56.0	74.8	46.0	84.3	53.9
60-64	80.0	58.9	81.0	59.7	84.8	63.1	90.5	69.2
65-69	91.9	81.0	85.4	73.8	92.7	80.0	90.8	77.6
70-74	96.5	87.6	86.0	74.8	92.5	78.7	85.8	71.0
Algebraic mean of 10 cohorts:	74.4	44.1	71.4	42.5	65.3	14.1	62.5	18.5

Source. U.S. Bureau of the Census (1952: Table 65, 1964, Table 173).

Note: Unlike Table 4, the 1960 data here were based on nonwhites rather than blacks in order to retain the same detail for both races in the age classifications. The 1950 data were based on nonwhites as in Table 3. Very few nonwhites would have been other than black in Philadelphia in 1950. Years of school completed was coded as described in Tables 3 and 4.

aA plus sign indicates an increase in the between-group variance.

that difference would amount to about one IQ point, which represents an increase of about 10% in the magnitude of the mean difference. Thus, there was a genuine net improvement in fit for the 1960 data when age was controlled, even though it is not apparent from comparing critical differences.

The success of controlling age in Panel D was definitely less impressive than in Panel B, but this fact too must be interpreted in relation to an initial state of affairs that differs considerably between the two panels. As can be seen in the top row of each of the two panels, the fit was much poorer initially in Panel B than Panel D. This contrast can be traced back to the first tests of the 1950 and 1960 data in Tables 3 and 4; on these, the 1950 data consistently produced the worst fit observed for years of schooling. The 1960 data, however, produced reasonably good results (see Table 7). Disaggregating the sexes and using education for males alone improved the fit for both sets of data, but the fit for the 1950 data remained extremely poor. Consequently, it has been much easier to obtain an improvement in fit by controlling age in Panels A and B than in Panels C and D, because the 1950 data allowed more room for improvement in the first place.

More detailed evidence in support of that conclusion is at hand. Examining Panels B and D of Table 9 further, we find that controlling age failed to improve the average fit in all but one case whenever the initial reduction in between-group variance in the top row amounted to 68.6% or more. For the sole exception, the improvement was trivial. Whenever the initial reduction amounted to 58.4% or less, controlling age led to a tangible improvement. Thus, the effect of controlling age must be interpreted in light of the prospects for extracting further improvements in fit from the data. In Panel D those prospects were not as good as in Panel B, and the improvements that were observed are all the more significant for that reason.

This line of reasoning also accounts for the odd fact that four out of the five net improvements of any size in Panel D applied to delinquency rates of girls. Since a good fit is harder to obtain for girls, there was simply more room for improvement in the case of the girls' data. It also accounts for the fact that the only sizable improvements involving boys and girls together occurred on the extreme right side of Panel D, where education was based on women and where the central tendency of education was based on medians. That combination of characteristics gave rise to the poorest pair of initial fits for boys and girls in the panel, hence that pair had the most room for improvement in fit. The importance of the available room for improvement is further demonstrated in this case by the fact that the two columns of corresponding critical differences directly above, in Panel C, were the only ones to register a net improvement as measured by that statistic.

Understandably, considerations concerning constraints imposed by the initial room for improvement must now qualify an earlier interpretation in the case of three of the six major analyses in Table 8. The failure of log-earnings from OCG-I and the Duncan SEI from both OCG surveys to show net improvements in fit after age was controlled may have been due to the fact that their initial reductions in variance were all greater than 68.6% (to invoke our only empirical standard of lack of room). In the remaining three analyses, however, the initial fit was clearly poor enough to allow room for net improvements. Since none occurred, there remains some indication still that the benefit of controlling age may be restricted to years of schooling, which is the interpretation that best supports the IQ-surrogate hypothesis. Unlike earnings and the Duncan index, amount of schooling is usually fixed once and for all relatively early in life, and so, as a cohort ages, schooling becomes a datum that reflects historical rather than contemporary conditions to an ever increasing degree.

The left and right halves of Panels C and D once again show that men's education yields a better fit than women's education within individual age cohorts. The effect is sustained mainly within the six younger cohorts as well as in the oldest cohort; overall, it appears in 30 out of 40 possible comparisons, which is again significant according to the sign test ($p < .05$) after dividing the observations in half to allow for non-independence between means and medians. Once more, the men's and women's averages in the bottom rows of the panels summarise the sex effect conveniently.

The place to look for meaningful interactions of fit with age is on the men's sides of Table 9. Since the fit tends to be better for men's education, meaningful patterns have a better opportunity to emerge. In Panels A and B, the fit is noticeably better in the four younger cohorts, which would be the adult cohorts more relevant to delinquents according to most age-based theories. However, those happen to be the cohorts with age intervals of 5 rather than 10 years, which might make controlling for age more effective in their case.

To assess the extent to which the good fit of the younger four cohorts depended on a smaller age interval, two new cohorts were formed with intervals 10 years wide, embracing ages 25-34 and 35-44. For the new 25-34 cohort the reductions were 93.5% and 99.6% (means) and 99.9% and 93.8% (medians) for boys and girls, respectively. These figures represent trivial decreases in fit. The next two original cohorts, which did not attain as good a fit when based on medians as the first two, even improved slightly in fit when combined in a 35-44 cohort, where the reductions were 94.1% and 99.9% (means) and 92.1% and 97.3% (medians). Since the interaction was in no way diminished by combining the younger cohorts, it cannot be attributed to differences in cohort intervals.

Naturally, any sign of meaningful interaction in the 1950 data renders an inspection for interaction of the 1960 data all the more interesting. In Panels C and D, where all cohort intervals are five years wide, there is some interaction, but since the better examples of fit occur in both the younger and older age ranges and the worse examples occur in the mid-range the effect is neither meaningful as it stands nor especially corroborative of the 1950 pattern of interaction. The very best fits, for example, are in the two oldest cohorts.

Considered together, the mutually inconsistent 1950 and 1960 interactions provide no real support for the hypothesis that some age cohorts are systematically more relevant to black-white differences in delinquency than others in any causal sense. The absence of evidence for a consistent, meaningful interaction represents a further lack of support as well for the hypothesis that black-white differences in educational attainment account for race differences in delinquency by functioning intergenerationally as background variables are usually supposed to operate during socialisation. By a process of elimination, therefore, the hypothesis that any reductions in between-group variance accomplished by years of schooling stem merely from its role as a surrogate for IQ in the model becomes more attractive.

Besides the net gains in fit from controlling age and sex in Table 9, there are forms of indirect evidence favouring the surrogate hypothesis. One is that in all 34 combinations of race with cohort in Table 9 the standard deviation of years of schooling was always greater for men than for women. (These education parameters are not shown in the table, because they are too numerous.) The consistency over race and time indicates that the observation by Sewell et al. (1977) concerning standard deviations in their own data must hold in general. Recall that the sex differential in standard deviations was regarded, in view of known correlations, as a manifestation of a sex differential in the *g*-loading of that variable. Now the totally consistent relation between size of men's and women's standard deviations has occurred in conjunction with widespread improvements in fit from using men's rather than women's education in the model, and that model involved delinquency rates that are known to mark certain points in the IQ distributions of blacks and whites fairly precisely (Table 2). Such interrelated facts should not be allowed to languish in isolation.

Another indirect form of evidence establishing the greater *g*-loadness of men's education and hence supporting the IQ-surrogate hypothesis involves a deduction based on the Spearman hypothesis (Jensen, 1985). According to the Spearman hypothesis, black-white differences in representative samples (i.e., samples unselected on g) vary directly with the g-loading of the variable concerned. Consequently, if years of schooling is more g-loaded for men than women, black-white differences on that variable should prove greater for men than women. This expectation is borne out for raw means and medians in all 17 age cohorts of Table 9. For each statistic, the sign test is significant ($p < .001$). Note that this combination of psychometrics and demography also accounts for a consistent race-by-sex interaction with schooling within cohorts, which otherwise might elicit social psychological attempts to explain that phenomenon without taking g into consideration.

The greater success of men's than women's education in the model, the sex differential in size of standard deviations, the consistency of the race-by-sex interaction with the Spearman hypothesis, and the nature of the model itself are all tied together, empirically as well as theoretically, in ways that are difficult to account for without assuming that the men's extra variance must be g-loaded. Without that assumption, it is difficult to see how the improvements in fit gained from disaggregating men's from women's education could have occurred throughout the many cohorts. Empirical consistency within a set of complex relations is usually accepted as evidence for validity of the unifying inference. The inference here is that education succeeds in the model to the extent that it serves as a surrogate for IQ and not otherwise. Goodness of fit responds to variations in the degree to which education fulfils the surrogate role rather than to variations in education proper.

One final aspect of Table 9 concerns the goodness of fit of years of schooling vis-a-vis that of IQ. With all available refinements applied to years of schooling, how would that variable stack up against IQ if the evidence for the surrogate hypothesis did not exist? It makes sense to begin with the best example of fit for both boys and girls in Table 9.

As judged by reduction in variance, that example occurs in the 25-29 cohort on the left side of Panel B. There, for 1950 data based on medians, the mean boy-girl reduction amounts to 98.1%. This single result is definitely in the same league with those achieved by IQ. In fact, it is slightly superior to the outcomes for IQ when based on the more severe (and hence more difficult to fit) training school criterion of the 1960 nationwide data, which averaged 97.5% in Table 2. However, this best example does not apply to that delinquency criterion, and so the fact remains that IQ, with a mean boy-girl reduction of 99.35%, can still claim the best fit for the 1950 court record criterion to which the example does apply. We should keep in mind, too, that the example of excellent fit in question happens to represent the best of 68 age-specific attempts in Table 9 to fit years of schooling to the delinquency prevalence rates. If we consider only the Philadelphia data, IQ has performed better in a single try than years of schooling succeeded in doing in 28 age-specific attempts. Comparing the best education example from numerous trials with one IQ trial risks capitalizing on any chance factors that might favor education.

The next best-fitting models are located in the next three cohorts of Panel B, which have mean boy-girl reductions ranging from 96.4% to 96.7% if we count only the more favourable pairs of the dual results based on means or medians. Beyond these four consecutive cohorts the fit declined markedly; this contrast constituted the indication of meaningful interaction that was discussed above.

However, the same pattern of interaction was not replicated in Panel D, where none of the 40 age-specific boy-girl combinations of outcomes seriously rivaled the success of IQ. The best single outcome, with a mean boy-girl reduction of 92.05%, was found in the oldest cohort of all. Such a datum serves no theoretical purpose other than that of undermining the pattern of interaction in Panel B.

Moreover, the benefits of controlling cohort age were not confined to the four younger cohorts in Panel B, which suggests that focusing on the likely parents of 1950 adolescents or on young adults likely to serve as role models to adolescents was not the crucial consideration. The 45-54 cohort benefited too, and if we examine the statistics based on medians (among which the best fit of all was contained), there were improvements throughout, including a particularly large one for the oldest cohort. Many more such improvements in the older cohorts can be found on the right side of Panel B. Thus, even though there was a potentially meaningful interaction, the pervasiveness of the improvements suggests that the IQ-surrogate hypothesis rather than causal relevance to adolescence could well have been responsible for the better results in the four younger cohorts too.

Besides the large number of trials involved; a potential role for chance in accounting for their good fit is further suggested by one peculiar fact that the four best-fitting cohorts of Table 9 have in common. They exhibited the highest percentages (8.5% to 11.4%) of persons (white males in this case) whose education placed them in the open-ended, uppermost category. The best-fitting cohort of all, 25-29, had the highest such percentage. This coincidence may be implicated in the good fit attained by these four Philadelphia cohorts, as the following analysis shows.

For the 1950 and 1960 age cohorts, respectively, correlations between percentages of white males in the open-end category and white standard deviations were -.93 and -.92 for years of schooling. The strong correlations indicate that white standard deviations were often reduced by ceiling effects as the amount of schooling acquired by white males increased. In absolute terms, the effect was greatest in the best-fitting Philadelphia cohorts, judgeing from the size of their open-end percentages.

Recall that reducing the white standard deviation, even if only through ceiling effects, increases the standardised black-white difference, $(M_b - M_w)^*$, in Equation (4a) and also increases the term (S_b/S_w), so that it becomes greater than 1.0 in some cases. As was mentioned in the discussion of log-earnings in Table 8, by IQ-model standards the term $(M_b - M_w)^*$ is always too small for SES variables. An artifact that enlarges that term will in general benefit the fit of SES, therefore, but also artifactually. Increasing the ratio of standard deviations, (S_b/S_w), can also help to compensate for a deficiency in $(M_b - M_w)^*$, as we saw in connection with log-earnings in Table 8. In the case of boys, where z_b was effectively nullified in the 1950 data (which reflect the court record criterion), the term (S_b/S_w) would have little impact on fit, but in the case of girls its impact would be greater. The fit for girls is not invariably better, however, because sometimes the combined effect of the two terms is too great, leading to erors of fit in the other direction.

Thus, the four very best cohort fits in Panel B of Table 9 are subject to an additional source of doubt concerning their validity, since they depend to some degree on a beneficial artifact affecting estimation of the white standard deviation. (However, results from the variant model, below, which does not involve standard deviations, maintain the relatively good fit for these cohorts.) Because men's education increased from 1950 to 1960, the poorer fit of the 1960 data seems to offer an exception to the described link between goodness of fit and high percentage of white males in the open-end category. However, the exception is only an apparent one; the 1960 data were less subject to the artifact, a higher education cutoff having been employed for the open-end category ("17 or more" versus "16 or more" years in 1950).

This overview of Table 9 shows that results that come close to rivalling those for IQ are confined to Panels A and B. But those panels were also the ones to which the IQ-surrogate hypothesis appplied most strongly. Hence, there was never an occasion on which years of schooling performed as well as IQ for the same prevalence rates, and those occasions on which years of schooling performed almost as well were heavily subject to the interpretation that it was IQ, after all, that enabled its surrogate variable to look good in the model. Although Panels A and B also contained the most suggestive evidence for meaningful interaction, such a major hypothesis demands dependable replication that was not forthcoming from Panels C and D. Indications that the meaningful interaction may have depended on artifacts due to ceiling effects also cast doubt upon the evidence for hypotheses requiring such an interaction.

Table 10 demonstrates the effect of applying the variant model to the data on which Table 9 was based. As before, only variance reductions based on medians need be considered, and interest continues to focus on the men's side.

The major patterns that were consistent with the IQ-surrogate hypothesis in Table 9 persist in the 1950 data of Table 10 (Panel A). Controlling age and sex both improved the fit, indicating that the IQ-surrogate hypothesis does not depend on the original model alone. The pattern of meaningful interaction also persists and thus seems to indicate that the interaction may not depend entirely on the artifact described above concerning the white standard deviation, because standard deviations do not figure in the variant model. However, to some extent the good fit of the variant model does benefit from interpolation within census categories assumed to have rectangular distributions, especially in the case of the lower prevalence rates of girls. In view of the wider separation of black and white prevalence rates for the court record criterion, the location of black and white critical values within the same census category of education cannot be considered an artifact when the fit is good. But the final stages of fitting the model depend on interpolations within such categories, and at that point the similarity between critical values for blacks and whites is often enhanced by the assumed rectangularity.

In panel A, the mean boy-girl fit with age controlled is not as good, on the average, as that achieved with the same medians by the original model in Table 9. The fit of the best-fitting cohort, 25-29, deteriorates slightly, but the average fit of the four best-fitting cohorts improves slightly when compared to the original model based on means, and somewhat more when compared to the original model based on medians. Consequently, the interaction involving younger and older cohorts is also somewhat accentuated in Table 10, although not exemplified as well by the 25-29 cohort. The relative successes of the variant model over the original are therefore few and spotty in Panel A. The variant model also tends to perform more poorly than the original on the prevalence rates of girls. In general, the 1950 data indicate that when the original model fits the data well, so does the variant model. Even when concern over the artifact in Table 9 is set aside, this is not an especially compelling argument in favour of the variant model, because the two are expected to be equivalent when normality holds, and to be more or less equivalent when normality is approximated. Approximating normality reasonably well in the lower half of the distribution, at their particular time and place, may account for the success of schooling in the case of the best-fitting cohorts in Tables 9 and 10. Like the original model, the variant model in Panel A never quite equals the fit of IQ to the same prevalence rates.

Being based on the low prevalence rates of the training school criterion, the 1960 data in Table 10 (Panel B) once again illustrate the full effect of the second artifact to which the variant model is subject, described above. In consequence, the fit is superficially so good everywhere that no meaningful patterns remain. The poorer fit in the youngest cohort merely reflects the eventual depopulation of the lowest education category by the historical trend toward increased education. As a result, interpolation in that low category finally produced critical values for blacks and whites that have some degree of separation. The consistent contrast in fit between Panels A and B testifies to the misleading nature of the artifact. So does the fact that when black prevalence rates are arbitrarily set equal to white prevalence rates in Panel B, 73.5% of the cases in which the fit was 99% or better decrease in fit by less than 1%.

Table 10 presents the final opportunity for examining the variant model. Nothing there seems to warrant the conclusion that SES variables are placed at a special disadvantage by the original model, and that (meaningful) good fits would emerge consistently for those variables if only the assumption of normality was relaxed. Although the variant model did perform better in Panel A of Table 10 than in Table 7, that improvement fails to elevate the model above the original one. Moreover, the improved performance occurs within a context that supports interpreting the

Table 10: Goodness of Fit Statistics from Tests of Variant IQ-type Models Using Years of School Completed in 1950 and 1960 by Men and Women Within Age Cohorts to Account for Race Differences in Prevalence of Philadelphia Court Record and Nationwide Training School Delinquents

Age	Education of men		Education of women	
	Boys	Girls	Boys	Girls
A. Philadelphia, 1950: Between-group variance reduction (percent)				
25-74 (entire range)	52.4	+178.1	+155.4	+1,536.8
25-29	90.0	99.9	77.0	99.9
30-34	96.2	99.6	85.3	97.5
35-39	98.1	98.2	54.8	80.1
40-44	99.9	94.9	70.6	23.5
45-54	79.0	65.2	+4.6	+383.3
55-64	+7.5	+99.6	56.2	+177.0
65-74	64.0	79.7	70.2	46.6
Algebraic mean of 7 cohorts:	74.2	62.6	42.5	+30.4
B. Nationwide, 1960: Between-group variance reduction (percent)				
25-74 (entire range)	99.9	100.0	97.8	99.9
25-29	81.5	99.2	40.6	95.7
30-34	98.7	99.9	92.9	99.3
35-39	99.8	100.0	99.9	99.9
40-44	96.5	99.9	96.3	99.9
45-49	97.2	99.9	98.9	99.9
50-54	99.2	99.9	99.9	100.0
55-59	99.7	99.9	99.9	100.0
60-64	99.9	100.0	99.7	99.9
65-69	99.9	100.0	99.8	99.9
70-74	99.9	100.0	99.9	100.0
Algebraic mean of 10 cohorts:	97.2	99.9	92.8	99.4

Note: The variant model uses medians, but not means. For sources and other details, see Table 9. In this table, entries of 100.0% represent fits with very small residuals. Unless the residual was very small, entries of 99.9% were not rounded upwards.

variable as a surrogate for IQ at least as strongly as it supports viewing some narrow specification of SES (i.e., education of younger adult males) as the key factor accounting for black-white differences in delinquency of boys and girls. This concludes the discussion of education in relation to the IQ-surrogate hypothesis.

Prior to the analyses concerning years of schooling in Panel B of Table 9, the single best fit of all SES variables in the original model had been achieved by Duncan's SEI. Recall that in Table 7 the mean boy-girl reduction of variance attained by the SEI had been 96.4% when men with farm backgrounds during childhood were excluded. This high percentage was eventually surpassed or tied by years of schooling in the case of the four younger cohorts in Panel B of Table 9, but there the concern was with the less severe, and hence easier to fit, criterion of court record delinquency. Furthermore, the SEI has been far more consistent than even the refined versions of schooling in Table 9 in producing large reductions in the between-group variance (Tables 7 and 8). Even when handicapped by the inclusion of men with farm backgrounds, the SEI was more consistently successful in reducing variance than all other SES variables, with the limited exception of 1972 log-earnings in Panel B of Table 8. Clearly, we need to pursue the reasons for the SEI's superior overall performance in the model and for the rather substantial impact that excluding men with farm backgrounds had on the goodness of fit, just as we did in the case of years of schooling for men.

Fortunately, we have Duncan himself to thank for having provided invaluable clues on both counts. As the result of certain analyses, Duncan et al. (1972: 77) concluded:

> The psychologist's concept of the "intelligence demands" of an occupation is very much like the general public's concept of the prestige or "social standing" of an occupation. Both are closely related to independent measures of the aggregate social and economic status of the persons pursuing an occupation.

Specifically, Duncan et al. found that correlations between the SEI scores of occupations and psychologists' ratings of those occupations according to their usual intelligence demands (on the Barr scale) ranged from .81 to .90, depending on which of two overlapping lists of occupational titles was employed. The above correlations were based on 96 and 47 job titles, respectively. For the shorter list, it was possible to intercorrelate the Barr scale, the SEI, and prestige ratings of the occupations and those correlations ranged from .90 to .91. Elsewhere, in the course of developing her own theory concerning the functional basis of the occupational hierarchy, Gottfredson (1985: 141) has reported a correlation of .82 between occupational prestige and a dimension of intellectual difficulty derived by factor analysis from a variety of job analysis ratings. But perhaps most important for present purposes is the simple empirical fact that the mean IQs of occupational groups often correlate from .90 to .95 with mean ratings of the SES of the groups' occupations (Jensen, 1980: 340). (Naturally, such correlations are highest when the occupations are grouped into relatively few categories.)

These various correlations and their theoretical interpretations by Duncan et al. and by Gottfredson establish that the SEI would function as an excellent surrogate for IQ on the aggregate level and that that relation is likely to endure. This means that although the SEI would not discriminate among individuals nearly as well as an IQ test would, that index describes occupationally defined aggregates of individuals in much the same way as occupationally defined aggregates of IQ scores would whenever entire populations, such as blacks and whites, are the focus of interest. Consequently, although the parameters of a distribution of SEI scores for individuals would have no easily specified relation to the IQ parameters of that distribution, those SEI parameters would very likely maintain *roughly* the same relation to the SEI parameters of a second distribution of individuals that the IQ parameters of the first distribution had to the IQ parameters of the second distribution, particularly if both distributions consisted of men. Such a rough correspondence would suffice for producing substantial reductions in the between-group variance of IQ-type models more or less consistently when the SEI is substituted for I2.

Persons currently in farm occupations usually have a history of farming in their families. Therefore, excluding men with farm backgrounds in their own childhoods excludes also virtually all men who themselves have farm occupations currently; this two-stage process of exclusion can be understood as a refinement that permits the SEI to fulfil its role as IQ-surrogate more effectively. Blau and Duncan (1967: 286), for example, commented that they were not wholly confident of the SEI as an index of the socio-economic position of farm occupations, and that the prestige ratings given farming by the general public tend to be higher than the SEI would predict. These difficulties were apparent when the SEI was created. Duncan (1961a: 131, 1961b: 155) noted that the real income component of the SEI was underestimated for farm occupations, where some production is often for home use, and that the index could not reflect "the considerable variation among farmers in size of farm enterprise, type of farm tenure," and so on (1961b: 156). These concerns indicate that Duncan's index would function relatively poorly as a surrogate for IQ where farming is concerned.

The extension of the IQ-surrogate hypothesis to the SEI is further supported by the impression one gets from scanning the data that the configuration of the parameters that enter into Equation (4a) often corresponds to the actual configuration of IQ parameters better in the case of the SEI than in the case of other SES variables. A favourable configuration, as represented by the nationwide IQ parameters in Table 2, is one in which $(M_b - M_w)^*$ and (S_b/S_w) equal about -1.1 and .80, respectively. Note that the term $(M_b - M_w)^*$ summarises the configuration for three of the four required parameters in one stroke.

For the 1962 OCG-I data in Table 8, $(M_b - M_w)^*$ approximated -1.1 rather well in the case of the SEI, where values ranged from -.87 to -.91. And in the case of the best SEI fit of all, accomplished by excluding men with farm backgrounds in Table 6, that key value equalled -.97, which is even closer to -1.1. However, following the Civil Rights Act of 1964, the values of the standardised difference in population means for the 1973 OCG-II data decreased in absolute magnitude to range between -.55 and -.80. But the SEI retained its generally good fit nevertheless as the result of a largely offsetting change in (S_b/S_w). For example, in 1962 the values of that term ranged from .52 to .66, but by 1973 they increased to range from .65 to .87 and thus they became closer to their target value of .80 as defined by IQ.

It is illuminating to consider how these offsetting changes came about. Since the SEI standard deviations of whites changed hardly at all in the interim, the change in (S_b/S_w) was due to changes in the standard deviations of blacks. In 1962, the black standard deviation was too small, probably because of floor effects in the range of occupational status. But as the black SEI mean rose, the black distribution acquired room to spread out, with the result that the value of (S_b/S_w) in the 25-64 age group was .81 in 1973 and thus almost exactly in agreement with the ratio of IQ standard deviations of .80.

Aside from the evidence on goodness of fit, it is difficult to compare these configurations of parameters in any convenient way with those of other SES variables, and so the impression based on scanning the data must suffice. These examples suggest that the black-white configuration of parameters in the case of the SEI maintains a rough correspondence to the black-white configuration of IQ parameters even when some of the SEI parameters have been disturbed, because compensating changes seem to occur elsewhere in the configuration. Such robustness testifies to the fundamental relation between the distribution of intelligence and the distribution of occupational status that Gottfredson (1985: 141) had in mind when she suggested that "the hierarchy itself probably *arises* from and is shaped by the variance in intelligence among workers in a society." Recall also the parallelism noted above between IQ and dollar earnings with respect to the relative sizes of the black and white standard deviations.

Summary of evidence concerning the IQ-surrogate hypothesis. The IQ-surrogate hypothesis seems to account for the better of the results obtained from inserting SES variables into the model. Each of a series of independent refinements of the data that would clearly enable a promising SES variable to function even better as a surrogate for IQ for the purpose of estimating parameters has resulted in net improvements in fit. Those refinements consisted of disaggregating years of schooling by sex and then by age, and also of excluding men with farm backgrounds from analyses using Duncan's index of occupational status.

Not the least significant piece of evidence by any means was the unequivocal identification of the SEI itself by Duncan et al. as an excellent surrogate for intelligence on the aggregate level. Numerous correlations were cited in support of that conclusion.

On the individual level, of course, years of schooling is already well-known as the SES variable having the highest correlation with IQ (i.e., .5 to .6), and so its role as an occasionally successful surrogate for IQ is not altogether surprising either. What may seem anomalous to some readers is that Duncan's index of occupational status performed more consistently than years of schooling in the model despite its having a lower correlation with IQ at the individual level (e.g., about .4). However, that difference has no necessary relation to their relative performances on the aggregate level, and as Gottfredson's (1985) theory makes clear, there are good reasons why the occupational world stratifies workers as to mean IQ quite successfully, even though it relies heavily on years of schooling instead of IQ in doing so. It is even conceivable that occupational status represents a more universal and stable surrogate for IQ and for group differences in IQ than years of schooling does, in view of its greater responsiveness to economic realities. For example, the difference between blacks and whites in median schooling has narrowed to less than half a year in recent cohorts (U.S. Bureau of the Census, 1979: Table 71), which amounts only to about .12 white standard deviation, but the black-white difference in crime and delinquency has remained as great as ever.

Other evidence in support of the IQ-surrogate hypothesis consisted of the absence of a consistent meaningful interaction in Tables 8 and 9 between age of cohort and goodness of fit. That lack constitutes a critical gap in any argument that would attempt to compensate for the poor performance of the general population's SES in IQ-type models by stressing that the SES of some age cohorts may be more relevant to delinquency than that of others, either because younger cohorts contain the parents of children in the delinquency age-range or because young adults provide adolescents with examples of the socio-economic position they can expect to attain. The singular appearance of such an interaction involving the four younger cohorts in Panel B of Table 9 was not replicated in Panel D. Moreover, the presence of improvements in fit among even the older cohorts, despite the interaction, indicated that special causal relevance through proximity to adolescence was certainly not the only possible source of the improvements.

Further evidence consisted of the universal differential between men's and women's standard deviations for years of schooling in Table 9, the simultaneous involvement of that differential in improvements in fit based on men's education, and the fact that those improvements themselves involved prevalence rates whose structure was known to bear a strong relation to black and white IQ parameters. A consistently greater black-white difference in absolute years of schooling for men than for women was also found to be in accord with the Spearman hypothesis. Such a complex set of relations clearly supports the IQ-surrogate hypothesis.

Finally, although a few of the better examples of fit obtained by using years of schooling and the SEI in the model could be considered excellent, none of them ever equalled or surpassed the mean boy-girl fit achieved by IQ for the same set of prevalence data. This in itself represents a strong indication of the relevance of IQ to black-white differences in delinquency.

Final remarks. Thus concludes what can be regarded as an extended exercise in convergent and discriminant validation (Campbell and Fiske, 1959). The developments cited at this article's beginning represent typical examples of convergent construct validation, through empirical replication and extension of mutually consistent hypotheses. The less typical principle of discriminant validation was then employed to distinguish between IQ and SES as explanations of the black-white difference in delinquency. For the original and variant models employed, differences in delinquency prevalence rates were found to be more consistent with black-white differences in IQ than with black-white differences in SES. The argument eventually returned full circle to convergent validation when the IQ-surrogate hypothesis was found to account for those few instances in which SES performed better than at other times. IQ is thus the better candidate, and indeed the only one so far, for explaining why the black-white difference assumes the particular magnitudes it does, and for implementing the important principle when deciding causality that effects should be proportional to their causes (Einhorn and Hogarth, 1986).

Conclusions

1a. Recent developments concerning progress with the race-IQ-delinquency model were described. Laub extended four key propositions, which Gordon had based on official juvenile delinquency data for whites, to blacks and to adults using victimisation data. Gordon reported evidence for the normality of *g*, and numerous studies continued to validate IQ tests as measures of *g* for blacks and whites. Large black/white ratios in crime statistics were validated by Hindelang, Laub, and Langan. Hirschi and Hindelang showed the association between delinquency and IQ to be at least as strong as that between delinquency and SES. Modest negative correlations between IQ and delinquency were demonstrated consistently and exceptions were found to be special cases. Differences between blacks and whites in the IQ-delinquency correlation were accounted for by known differences between them in location on the IQ continuum and by hypothesised differences in contextual effects relating to their IQ distributions. Estimates of black and white IQ parameters were found to be stable and consistent over time. Additional replications of the fit of the IQ model to lifetime rates of adult imprisonment that had been reported by Langan and Greenfeld were noted.

1b. Simulations ruled out equivalent IQ-specific delinquency rates as an explanation for IQ-commensurability of prevalence rates, and revealed that IQ-commensurability required IQ-specific rates to be higher for blacks than whites. This requirement accounted for the missing "surplus" of black delinquents anticipated on sociological grounds and thus opened the door to contextual effects of IQ distributions on delinquency. Such effects were hypothesised to depend entirely on the IQ distributions and certain empirical generalisations, such as regression meanwards of parental IQ. In turn, this hypothesis led to a new rationale for IQ-commensurability (the empirical consistency observed between black-white differences in delinquency, on the one hand, and black-white differences in IQ parameters, on the other hand). Because IQ parameters specify exactly both normal IQ distributions as well as the proportions of those distributions located beyond given cutoffs, it is reasonable, according to the new rationale, that individual and contextual effects determined by the IQ distributions would cumulate to produce delinquency prevalence rates that continue to reflect the differences in IQ parameters by corresponding to IQ cutoffs that are similar for blacks and whites.

2. One vague criticism that probably represented a reaction commonly heard from sociologists attributed the success of the model to SES rather than to IQ. Accordingly, the performance of SES variables in the original race-IQ-delinquency model has been extensively reviewed. All known data sets containing SES measurements for relevant populations collected at times reasonably contiguous to the avail-

able prevalence rates of delinquency were examined. These included examples of all of the more widely recognised and important SES variables, such as male income, family income, educational attainment, and occupational status. In all, more than 125 boy-girl pairs of trials were conducted for just two sets of sex- and race-specific delinquency rates.

3. On the average, the measures of sociology's master variable did not perform well when substituted for IQ in the original model. Raw income for males performed worst of all, which is particularly interesting since group differences in delinquency are often attributed to poverty. Family income did better, but was still below average among variables that produced any boy-girl reduction. For the 125 boy-girl trials involving SES displayed in the tables, the mean reduction in variance amounted to 18.6%. If we consider only the 95 trials in which the boy-girl mean represented a net reduction in variance, the average reduction was 61.9%. Even the best few examples of fit may owe their success to chance when they have been cast up by such a large number of trials. In contrast, IQ averaged a 98.4% reduction in only four trials.

4. A variant model, not depending on normality for SES variables, was explored. In general, the variant model performed worse than the original model when artifactual results were excluded. When the variant model did perform well, improvement over the original model was minimal. Thus, the variant model performed well when the original model performed well, but offered no cure for poor performances of SES in the original model.

5. In view of the large number of trials considered, the few occasions on which SES variables yielded their better fits in the models could be due to chance influences. This chance hypothesis must, of course, be tempered by the fact that the better results were not distributed randomly over the SES variables considered. Instead, best results were concentrated mainly in two variables, years of schooling for men and the Duncan index. Furthermore, the few examples of excellent fit that were found also appeared in conjunction with those variables, although only under circumstances in which the variables had been refined so as to enhance their suitability as surrogates for IQ. The Duncan index performed well in its single trial that excluded men with farm backgrounds, and so in that form the index is not as subject to the criticism that the good result was due to chance as years of schooling is in view of its many trials in the refined form.

6. No theory of delinquency has singled out parental years of schooling or occupational status (Duncan's SEI), let alone their specially enchanced manifestations, as the ones most critical to black-white differences in delinquency. Consequently, any claim for their special relevance to delinquency theory that emerges following the findings reported here must inevitably have a conspicuously *ad hoc* status.

7. In no case did an excellent fit for any SES variable apply to more than one of the two criteria of delinquency, which were juvenile court appearance and commitment to training school. There was not always an opportunity to test for replication across delinquency criteria, but where there was (years of schooling in Table 9), replication did not appear. *Not even the best outcomes involving SES ever equalled those for IQ for the same prevalence rates.* From several standpoints, therefore, complete interchangeability with IQ was never demonstrated.

8. The two SES variables (education and Duncan's SEI) responsible for the best outcomes were also the two best-qualified to act as surrogates for IQ in the adult population, on a variety of grounds. First, they happen to be the SES variables that usually rank first and second in size of correlation with IQ at the individual level. Second, we have the testimony of Duncan et al. concerning the SEI, which equated the public's concept of occupational prestige with the intelligence demands of jobs, as well as Gottfredson's revision of functional theory reinstating the importance of intelligence to occupational stratification. Third, major variations in the goodness of fit of education and the SEI were associated with manipulations of their suitability

as surrogates for IQ.

9. For years of schooling especially, those manipulations of suitability consisted of demographic analyses that were essentially passive, in that instead of interfering with the variable itself by statistical means (e.g., partialling) or by excluding certain observations altogether (e.g., farm backgrounds for the Duncan index), they simply permitted the IQ-surrogate relationship to unfold further if indeed it was present (i.e., through disaggregation). By analogy, the effect is as though the manipulations permitted SES to rotate toward IQ in factor space, if SES was so inclined, and it did. Since SES joined IQ in factor space (rather than vice versa), SES must be interpreted in terms of IQ (rather than vice versa). In view of the logic underlying the ensuing improvements in fit-not to mention the chronological and causal priority of IQ over adult occupational status recognised in the status-attainment literature (e.g., Duncan et al., 1972)-it is evident that SES functions in the model as a spurious surrogate for IQ rather than the other way around. This proposition was termed *the IQ-surrogate hypothesis*. The fact that IQ does a much better job than SES of fitting the prevalence rates points to the same conclusion: IQ is the potent variable, not SES.

10. By providing indirect or roughly proportional estimates of black and white IQ parameters in both adult populations, the best-fitting SES variables simultaneously provided analogous estimates for both juvenile populations, because within each race the two generations share the same IQ parameters. Since the juvenile generation is very likely the one whose IQ parameters matter most for delinquency (Gordon, 1986a), this generational linkage enabled SES to function as a surrogate for IQ in the race-IQ-delinquency model whenever the proportionality relationship between IQ and SES parameters was sufficiently protected. Consequently, even when SES variables were most successful in the model, that success itself constituted evidence for the importance of IQ to black-white differences in delinquency.

11. One can always find fault with the conditions under which the SES hypothesis, or any other, has been tested. Lack of normality and reliance on parameters based to some extent on open-end categories could be mentioned as examples (although for Table 9 it was shown that open-end categories might benefit the fit of SES). Before voicing such complaints, supporters of explanations based on SES must state how they were able to overcome such difficulties in reaching their own conclusion (that SES is the relevant variable). Those who would brush aside the IQ results by exclaiming "Of course!" have an obligation to back up their implied criticism with evidence, and at the least, to specify the exact variable they have in mind.

12. The ways of studying between-group differences employed in this article have led to results that reinforce the earlier conclusions concerning SES reached by Savitz (1973: 508), who noted, "Socio-economic status was found to underpredict black delinquency and the factors related to socio-economic status did not fully account for the racial differences in delinquency, but they did reduce, to a small extent, the differential." He was referring to linear or categoric models.

13. A major by-product of this article has been the demonstration that the extraordinarily good fit to race differences in prevalence of delinquency that IQ achieves is not easily obtained using just any set of correlates of SES.

14. If there is any variable other than IQ or SES that critics of the proposed model have in mind, they should name it. The only conceivable candidate remaining would seem to be school achievement, but that variable would probably produce a good fit in the model only when measured by standardised achievement tests that have good reliability and high *g* loadings (e.g., Jensen, 1980: 97), as distinct from school grades. Thus, school achievement would also qualify as simply another IQ-surrogate. Any attempt to shift the heuristics of the underlying variable from those possessed by IQ to those suggested by schooling must confront the fact that the black-white IQ difference exists even before pupils enter school, that it holds for tests whose content is related to the school curriculum only remotely at best, that it is not changed in the course of schooling, that it has not decreased over time despite

the substantial reduction of black-white differences in the amount of schooling attained, and that the difference has been demonstrated not to depend on differences between blacks and whites in mastering specific information, but rather on differences in the same underlying *g* that accounts for individual differences within each race (Jensen, 1985; Gordon, 1985; 1987).

It is time to consider the black-white IQ difference seriously when confronting the problem of crime in American society.

References

Alexander, Karl L. and Bruce K. Eckland
1974 "Sex differences in the educational attainment process." *American Sociological Review*, 39: 668-82.

Arbeiter, Solomon
1984 *Profiles, College-Bound Seniors, 1984*. New York: College Entrance Examination Board.

Blau, Peter M. and Otis Dudley Duncan
1967 *The American Occupational Structure*. New York: Wiley.

Bogue, Donald J.
1969 *Principles of Demography*. New York: Wiley.

Campbell, Donald T. and Donald W. Fiske
1959 "Convergent and discriminant validation by the multitrait-multimethod matrix." *Psychological Bulletin*, 56: 81-105.

Cattell, Raymond B.
1938 "Some changes in social life in a community with a falling intelligence quotient," *British Journal of Psychology*, 28: 430-450.

Coleman, Richard P. and Lee Rainwater
1978 *Social Standing in America: New Dimensions of Class*. New York: Basic Books.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld and Robert L. York
1966 *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office.

Duncan, Otis Dudley
1961a "A socioeconomic index for all occupations." Pp. 109-38 in Albert J. Reiss, Jr., *Occupations and Social Status*. New York: Free Press.
1961b "Properties and characteristics of the socioeconomic index." Pp. 139-61 in Albert J. Reiss, Jr., *Occupations and Social Status*. New York: Free Press.
1968 "Inheritance of poverty or inheritance of race?" Pp. 85-110 in Daniel P. Moynihan (ed.), *On Understanding Poverty: Perspectives from the Social Sciences*. New York: Basic Books.

Duncan, Otis Dudley, David L. Featherman and Beverly Duncan
1972 *Socioeconomic Background and Achievement*. New York: Seminar Press.

Department of Defense (DOD)
1982 *Profile of American Youth: 1980 Nationwide Administration of the Armed Forces Vocational Aptitude Battery*. Washington, DC: Office of the Assistant Secretary of Defense.

Department of Health, Education, and Welfare (DHEW)
1976 *Intellectual Development and School Achievement of Youths 12-17 Years: Demographic and Socioeconomic Factors*. (Vital and Health Statistics: Series 11, Data from the National Health Survey; No. 158). (DHEW Publication No. HRA 77-1640.) Washington, DC: U.S. Government Printing Office.

Einhorn, Hillel J. and Robin M. Hogarth
1986 "Judging probable cause." *Psychological Bulletin*, 99: 3-19.

Farrington, David P.
1986 "Age and Crime." In Michael Tonry and Norval Morris (eds.), *Crime and Justice: An Annual Review of Research*, Vol. 7. Chicago: University of Chicago Press.

Featherman, David L. and Robert M. Hauser
1978 *Opportunity and Change*. New York: Academic Press.

Flynn, James R.
1984 "The mean IQ of Americans: Massive gains 1932 to 1978." *Psychological Bulletin*, 95: 29-51.

Gordon, Robert A.
1967 "Issues in the ecological study of delinquency." *American Sociological Review*, 32: 927-44.
1973 "An explicit estimation of the prevalence of commitment to a training school, to age 18, by race and by sex." *Journal of the American Statistical Association*, 68: 547-53.
1975 "Crime and cognition: An evolutionary perspective." Proceedings of the II International Symposium on Criminology. Sao Paulo, Brazil: Oscar Freire Institute.
1976 "Prevalence: The rare datum in delinquency measurement and its implications for the theory of delinquency." Pp. 201-84 in Malcolm W. Klein (ed.), *The Juvenile Justice System*. Beverly Hills: Sage.
1980a "Research on IQ, race, and delinquency: Taboo or not taboo?" Pp. 37-66 in Edward Sagarin (ed.),

Taboos in Criminology. Beverly Hills: Sage.
1980b "Implications of valid (and stubborn) IQ differences: An unstatesmanlike view." *Behavioral and Brain Sciences*, 3: 343-44.
1984 "Digits backward and the Mercer-Kamin law: An empirical response to Mercer's treatment of internal validity of IQ tests." Pp. 357-506 in Cecil R. Reynolds and Robert T. Brown (eds.), *Perspectives on Bias in Mental Testing. New York: Plenum.*
1985 "The black-white factor is g." *Behavioral and Brain Sciences*, 8: 229-31.
1986a "Scientific justification and the race-IQ-delinquency model." Pp. 91-131, In Timothy F. Hartnagel and Robert A. Silverman (eds.), *Critique and Explanation: Essays in Honor of Gwynne Nettler.* New Brunswick, NJ: Transaction Books.
1986b Review of James Q. Wilson and Richard J. Herrnstein, *Crime and Human Nature. Business Horizons*, 29: 78-79.
1987 "Jensen's contributions concerning test bias: A contextual view." Pp. 77-154 in Sohan Modgil and Celia Modgil (eds.), *Arthur Jensen: Consensus and Controversy* (Falmer International Master-Minds Challenged). Barcombe, Lewes, Sussex, England: Falmer Press.

Gordon, Robert A. and Leon J. Gleser
1974 "The estimation of the prevalence of delinquency: Two approaches and a correction of the literature." *Journal of Mathematical Sociology*, 3: 275-91.

Gottfredson, Linda S.
1984 *The Role of Intelligence and Education in the Division of Labor.* Report No. 355. Baltimore: Johns Hopkins University, Center for Social Organization of Schools.
1985 "Education as a valid but fallible signal of worker quality: Reorienting an old debate about the functional basis of the occupational hierarchy." Pp. 123-69 in Alan C. Kerckhoff (ed.), *Research in Sociology of Education and Socialization*, Vol. 5. Greenwich, CT: JAI Press.
1986 "Societal consequences of the *g* factor in employment." *Journal of Vocational Behavior, 29: 379-410.*

Gove, Walter R.
1985 "The effect of age and gender on deviant behavior: A biopsychosocial perspective." Pp. 117-44 in Alice S. Rossi (ed.), *Gender and the Life Course.* Hawthorne, NY: Aldine.

Harlem Youth Opportunities Unlimited, Inc. (HARYOU)
1964 *Youth in the Ghetto.* New York: Harlem Youth Opportunities Unlimited, Inc.

Hindelang, Michael J.
1978 "Race and involvement in common law personal crimes." *American Sociological Review*, 43: 93-109.
1981 "Variations in sex-race-age-specific incidence rates of offending." *American Sociological Review*, 46: 461-74.

Hirschi, Travis
1969 *Causes of Delinquency.* Berkeley: University of California Press.

Hirschi, Travis and Michael Gottfredson
1983 "Age and the explanation of crime." *American Journal of Sociology*, 89: 552-84.

Hirschi, Travis and Michael J. Hindelang
1977 "Intelligence and delinquency: A revisionist review." *American Sociological Review*, 42: 571-87.

Humphreys, Lloyd G., Charles K. Parsons and Randolph K. Park
1979 "Dimensions involved in differences among school means of cognitive measures." *Journal of Educational Measurement*, 16: 63-76.

Jensen, Arthur R.
1980 *Bias in Mental Testing.* New York: Free Press.
1985 "The nature of the black-white difference on various psychometric tests: Spearman's hypothesis." *Behavioral and Brain Sciences*, 8: 193-263.

Jensen, Arthur R. and Cecil R. Reynolds
1982 "Race, social class and ability patterns on the WISC-R." *Personality and Individual Differences*, 3: 423-38.

Kaufman, Alan S. and Jerome E. Doppelt
1976 "Analysis of WISC-R standardization data in terms of the stratification process." *Child Development*, 47: 165-71.

Kennedy, Wallace A., Vernon Van de Rief and James C. White, Jr.
1963 "A normative sample of intelligence and achievement of Negro elementary school children in the Southeastern United States." Monographs of the Society for Research in Child Development 28 (6, Serial No. 90).

Langan, Patrick A.
1985 "Racism on trial: New evidence to explain the racial composition of prisons in the United States." *Journal of Criminal Law and Criminology,* 76: 666-83.

Langan, Patrick A. and Lawrence A. Greenfeld
1985 *The Prevalence of Imprisonment.* Bureau of Justice Statistics Special Report. Washington, DC: U.S. Government Printing Office.

Laub, John H.
1983 "Urbanism, race, and crime," *Journal of Research in Crime and Delinquency,* 20: 183-98.

Loeber, R. and T. Dishion
1983 "Early predictors of male delinquency: A review." *Psychological Bulletin,* 94: 68-99.

McGarvey, Bill, William F. Gabrielli, Jr., P.M. Bentler and Sarnoff A. Mednick
1981 "Rearing social class, education, and criminality: A multiple indicator model." *Journal of Abnormal Psychology,* 90: 354-64.

Mercy, James A. and Lala Carr Steelman
1982 "Familial influence on the intellectual attainment of children." *American Sociological Review,* 47: 532-42.

National Center for Educational Statistics (NCES)
1966 *Supplemental Appendix to the Survey on Equality of Educational Opportunity.* Washington, DC: U.S. Government Printing Office.

Plomin, Robert and J.C. DeFries
1980 "Genetics and intelligence: Recent data." *Intelligence,* 4: 15-24.

Reiss, Albert J., Jr.
1961 *Occupations and Social Status.* New York: Free Press.

Reynolds, H.T.
1977 *The Analysis of Cross-Classifications.* New York: Free Press.

Rowe, Alan R. and Charles R. Tittle
1977 "Life cycle changes and criminal propensity." *Sociological Quarterly,* 18: 223-36.

Savitz, Leonard D.
1973 "Black crime." Pp. 467-516 in Kent S. Miller and Ralph Mason Dreger (eds.), *Comparative Studies of Blacks and Whites in the United States.* New York: Seminar Press.

Scarr, Sandra and Richard A. Weinberg
1978 "The influence of 'family background' on intellectual attainment." *American Sociological Review,* 43: 674-92.

Sewell, William H., Robert M. Hauser and Wendy C. Wolf
1977 "Sex, schooling, and occupational careers." CDE Working Paper 77-31, Center for Demography and Ecology, University of Wisconsin-Madison.

Shuey, Audrey M.
1966 *The Testing of Negro Intelligence,* Second Edition. New York: Social Science Press.

Silberman, Charles E.
1980 "Fear." Pp. 367-86 in Egon Bittner and Sheldon L. Messinger (eds.), *Criminology Review Yearbook,* Volume 2. Beverly Hills: Sage.

Terman, Lewis M. and Maud A. Merrill
1960 *Stanford-Binet Intelligence Scale: Manual for the Third Revision, Form L-M.* Boston: Houghton Mifflin.

Thornberry, Terence P. and Margaret Farnworth
1982 "Social correlates of criminal involvement: Further evidence on the relationship between social status and criminal behavior." *American Sociological Review,* 47: 505-18.

Tuddenham, Read D.
1948 "Soldier intelligence in World Wars I and II." *American Psychologist,* 3: 54-56.

U.S. Bureau of the Census
1952 *U.S. Census of Population: 1950. Vol. II, Characteristics of the Population, Part 38, Pennsylvania.* Washington, DC: U.S. Government Printing Office.

1963 *U.S. Census of Population: 1960. Subject Reports. Nonwhite Population by Race.* Final Report PC (2)-1C. Washington, DC: U.S. Government Printing Office.
1964 *U.S. Census of Population: 1960. Vol. I, Characteristics of the Population, Part 1, United States Summary.* Washington, DC: U.S. Government Printing Office.
1965 *Consumer Income: Income in 1964 of Families and Persons in the United States.* (Current Population Reports: Series P-60; No. 47). Washington, DC: U.S. Government Printing Office.
1979 *The Social and Economic Status of the Black Population in the United States: An Historical View, 1790-1978.* (Current Population Reports: Special Studies: Series P-23; No. 80). Washington, DC: U.S. Government Printing Office.
1982 *Population Profile of the United States: 1981.* (Current Population Reports: Series P-20; No. 374). Washington, DC: U.S. Government Printing Office.
1983 *Statistical Abstract of the United States: 1984.* Washington, DC: U.S. Government Printing Office.

Wechsler, David
1974 *Manual for the Wechsler Intelligence Scale for Children-Revised.* New York: Psychological Corporation.

West, D.J. and D.P. Farrington
1973 *Who Becomes Delinquent?* London: Heinemann.

White, Karl R.
1982 "The relation between socioeconomic status and academic achievement." *Psychological Bulletin,* 91: 461-81.

Wigdor, Alexandra K. and Wendell R. Garner
1982 *Ability Testing: Uses, Consequences, and Controversies.* Washington, DC: National Academy Press.

Wilson, James Q. and Richard J. Herrnstein
1985 *Crime and Human Nature.* New York: Simon and Schuster.

Wilson, William J.
1978 *The Declining Significance of Race.* Chicago: University of Chicago Press.

Wolfgang, Marvin E., Robert M. Figlio and Thorsten Sellin
1972 *Delinquency in a Birth Cohort.* Chicago: University of Chicago Press.

This article has been cited by:

1. E.A. Cocodia. 2014. Cultural Perceptions of Human Intelligence. *Journal of Intelligence* **2**:4, 180-196. [CrossRef]
2. Ivan K. Ash, Ann L. Edwards, Bryan E. Porter. 2014. An investigation of state population characteristics that moderate the relationship of state seat belt law and use in the United States. *Accident Analysis & Prevention* **71**, 129-136. [CrossRef]
3. Stephen J. Thoma, Pitt Derryberry, Darcia Narvaez. 2009. The distinction between moral judgment development and verbal ability: some relevant data using socio-political outcome variables. *High Ability Studies* **20**:2, 173-185. [CrossRef]
4. Michael A. McDaniel. 2006. Estimating state IQ: Measurement challenges and preliminary correlates. *Intelligence* **34**:6, 607-619. [CrossRef]
5. Christopher Hensley, Phyllis Gray-Ray, Yuk-Ying Tung, Xiaohe Xu, Melvin Ray. 1999. A racial comparison of Mississippi's juvenile violent and property crimes: A test of self-control and subculture of violence theories. *Journal of African American Men* **4**:3, 21-44. [CrossRef]
6. HARRY F. WEYHER. 1998. CONTRIBUTIONS TO THE HISTORY OF PSYCHOLOGY: CXH. INTELLIGENCE, BEHAVIOR GENETICS, AND THE PIONEER FUND. *Psychological Reports* **82**:3c, 1347-1374. [CrossRef]
7. HARRY F. WEYHER. 1998. CONTRIBUTIONS TO THE HISTORY OF PSYCHOLOGY: CXH. INTELLIGENCE, BEHAVIOR GENETICS, AND THE PIONEER FUND. *Psychological Reports* **82**:3, 1347. [CrossRef]
8. David Lubinski, Lloyd G. Humphreys. 1997. Incorporating general intelligence into epidemiology and the social sciences. *Intelligence* **24**:1, 159-201. [CrossRef]
9. Robert A. Gordon. 1997. Everyday life as an intelligence test: Effects of intelligence and intelligence context. *Intelligence* **24**:1, 203-320. [CrossRef]
10. Robert A. Gordon. 1996. Parental licensure and its sanction. *Society* **34**:1, 65-69. [CrossRef]
11. C. R. Brand. 1996. The importance of intelligence in Western societies. *Journal of Biosocial Science* **28**:04. . [CrossRef]
12. Lee Ellis. 1996. A discipline in peril: Sociology's future hinges on curing its biophobia. *The American Sociologist* **27**:2, 21-41. [CrossRef]
13. Troy Duster. 1995. What's New in the IQ Debate. *The Black Scholar* **25**:1, 25-31. [CrossRef]

14. David A. Ward, Charles R. Tittle. 1994. IQ and delinquency: A test of two competing explanations. *Journal of Quantitative Criminology* **10**:3, 189-212. [CrossRef]

15. Edward M. Miller. 1994. Paternal provisioning versus mate seeking in human populations. *Personality and Individual Differences* **17**:2, 227-255. [CrossRef]

16. Richard Lynn. 1990. The causes and cures of criminality. *Personality and Individual Differences* **11**:3, 319-320. [CrossRef]

17. Arthur R. Jensen, Michael E. Faulstich. 1988. Difference between prisoners and the general population in psychometric g. *Personality and Individual Differences* **9**:5, 925-928. [CrossRef]

18. Linda S Gottfredson. 1986. Societal consequences of the g factor in employment. *Journal of Vocational Behavior* **29**:3, 379-410. [CrossRef]