Socioeconomic Background and Achievement

OTIS DUDLEY DUNCAN

DEPARTMENT OF SOCIOLOGY
UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN

DAVID L. FEATHERMAN

DEPARTMENT OF RURAL SOCIOLOGY
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN

BEVERLY DUNCAN

DEPARTMENT OF SOCIOLOGY
UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN

SP SEMINAR PRESS New York and London 1972

Chapter 5

Intervening Variables, I: Intelligence

Work with the intervening variable, intelligence, was one of the major preoccupations of the project, extending over much of its duration. This degree of commitment to the task seemed justified, for one main reason, among others. While there is a widely accepted assumption that mental ability is a primary source of variation in occupational achievement—being built into such bodies of practice as vocational counseling and job placement, for example—there is a surprising and distressing lack of information on how ability actually combines with other determinants of success in the world of work. It is difficult to resist offering the suggestion that the imbalance in the state of knowledge and in the collection of relevant information on the topic is due to the fact that it has been left largely to psychologists and practitioners trained in psychology, who have not been inclined to investigate and discover its social import.

Our work on this topic may be listed as follows: a reconsideration of the conceptual relationship between intelligence and occupational status; a review of sources of reliable evidence on the correlation of measured mental ability with variables implicated in the process of achievement; and the construction of a model to represent the role of intelligence in that process. The final results on the last item are presented in detail in Duncan (1968a) and will only be summarized here; a somewhat lengthier treatment is given of the other two topics and of certain preliminary investigations carried out before constructing the final model.

TABLE 5.1

Barr Rating, Occupation Title, and Description, with Matching Group Socioeconomic Status Score (SES) and NORC Prestige Rating for Occupations Included in Barr Scale (See text for sources)[a]

P.E. Value	Occupation	Description	Group	SES	NORC	Note[a]
0.00	Hobo	–	Omit	–	–	–
1.54	Odd jobs	–	2	–	–	a
2.11	Garbage collector	–	2	–	–	a
3.38	Circus roustabout	Does heavy, rough work about the circus	2	–	–	a
3.44	Hostler	Care of horses in livery, feed and sales stables	2	–	–	a
3.57	R. R. Sec. Hand	Replaces ties, etc., under supervision	2	03	22.20	a
3.62	Day laborer	On street, in shop or factory as roustabout	2	–	–	a
3.99	Track layer	Does heavy work under supervision	2	–	–	a
4.20	Waterworks man	A variety of odd jobs, all unskilled	5	21	–	–
4.29	Miner	Digger and shoveller, etc.	1	10	24.32	–
4.81	Longshoreman	Loads and unloads cargoes	1	11	26.86	–
4.91	Farm laborer	Unskilled and usually inefficient	3	06	21.36	b
4.98	Laundry worker	Various kinds of work in laundry (practically unskilled)	2	15	19.01	–
5.27	Bar tender	–	1	19	19.86	–
5.41	Teamster	–	2	–	–	c
5.44	Sawmill worker	Heavy work, little skill required	2	05	30.75	–
5.59	Dairy hand	Milking, care of stock under supervision	3	–	–	b
5.81	Drayman	–	2	–	–	c
5.87	Deliveryman	Delivers groceries, etc., with team or auto	2	32	–	–
6.14	Junkman	Collector of junk	5	59	–	–
6.42	Switchman	Tending switch in R.R. yards	1	44	32.78	–
6.66	Smelter worker	Metal pourers, casting collectors, etc.	2	18	–	–
6.27	Tire repairer	In general automobile repair shop	5	08	–	–
6.85	Cobbler & shoemaker	Repairman in shoe shop	2	12	–	–
6.86	Munition worker	Average	5	08	–	–
6.92	Barber	Not owner. Has charge of chair	1	17	37.93	–
6.93	Mov. picture operator	Operates machine which projects pictures	2	43	–	–

TABLE 5.1 continued

P.E. Value	Occupation	Description	Group	SES	NORC	Note[a]
7.02	Vulcanizer	Understands the process of hardening rubber	5	22	–	–
7.05	General repairman	Repairs broken articles. Uses wood-working tools	5	19	–	–
7.06	Ship rigger	Installing cordage system on sailing vessels, working under supervision	5	32	–	–
7.17	Telephone operator	–	1	45	40.36	–
7.19	Cook	In restaurant or small hotel	1	15	25.97	–
7.23	Streetcar conductor	–	2	30	–	–
7.24	Farm tenants	On small tracts of land	2	14	21.52	e
7.30	Brakeman	On freight or passenger trains	1	42	34.65	–
7.33	City fire fighter	Handles the ordinary fire-fighting apparatus	1	37	43.81	–
7.39	R.R. fireman	On freight or passenger train	2	45	–	–
7.54	Policeman	Average patrolman	1	40	47.77	–
7.71	Structural steel worker	Heavy work demanding some skill	2	34	–	–
7.73	Tel. & tel. lineman	–	2	49	–	–
7.77	Bricklayer	–	1	27	35.66	d
7.79	Butcher	Not shop owner. Able to make cuts properly	1	29	32.12	–
7.91	Baker	–	1	22	34.18	–
8.02	Metal finisher	Polishes and lacquers metal fixtures, etc.	2	22	–	–
8.04	Plasterer	Knowledge of materials used necessary	2	25	–	–
8.08	General painter	Paints houses, buildings and various structures	1	16	29.78	–
8.22	Harness maker	–	5	32	–	–
8.40	Tinsmith	Makes vessels, utensils, etc., from plated sheet metal	2	33	–	–
8.49	Letter carrier	–	1	53	44.66	–
8.50	Forest ranger	–	2	48	–	–
8.58	Stone mason	–	1	–	–	d
8.75	Plumber	Av. trained plumber employee	1	34	40.58	–
8.89	Gardening, truck farming	Owns and operates small plots	2	–	–	e
8.99	Electric repairman	Repairs elec. utensils, devices and machines	5	27	–	–
9.28	Bookbinder	Sets up and binds books of all sorts	2	39	–	–

TABLE 5.1 continued

P.E. Value	Occupation	Description	Group	SES	NORC	Note[a]
9.37	Carpenter	Knows wood-working tools. Can follow directions in various processes of wood construction work	1	19	37.33	–
9.37	Potter	Makes jars, jugs, crockery, earthenware, etc.	2	21	–	–
9.54	Tailor	Employee in tailoring shop	2	23	–	–
9.72	Salesman	In drygoods, hardware, grocery stores, etc.	2	39	27.13	–
10.11	Telegraph operator	In small town	3	47	–	–
10.21	Undertaker	In small town. Six mo.–yr. special schooling	1	59	53.40	–
10.26	Station agent	In small town. Acts as baggage man, freight agent, operator, etc.	3	60	–	–
10.26	Mechanical repairman	In shop or factory. Keeps machines in condition	2	27	–	–
10.29	Dairy owner and mgr.	Small dairy, 50–100 cows	2	–	–	e
10.53	Metal pattern maker	–	2	–	–	f
10.54	Wood pattern maker	–	2	44	–	f
10.54	Lithographer	Makes prints from designs which he puts on stone	2	64	–	–
10.76	Linotype operator	–	2	52	–	–
10.83	Photographer	City 1000–5000. A few months' training, experience in studio	3	50	–	–
10.86	Detective	Traces clues, etc. Employee of detective bureau	2	36	–	–
10.99	Electrotyper	Prepares wood cuts	2	55	–	–
11.17	Traveling salesman	Sells drugs, groceries, hardware, drygoods, etc.	2	47	41.53	–
11.34	Clerical work	Bookkeepers, recorders, abstractors, etc.	1	51	47.56	–
11.35	R.R. Pass conductor	–	1	58	40.86	–
11.51	Store keeper & owner	Small town retail dealer, general or special store	1	33	46.07	–
11.74	Foreman	Small factory, shop, etc.	1	53	45.05	–
11.78	Stenographer	Writes shorthand and uses typewriter	1	61	43.34	–

TABLE 5.1 continued

P.E. Value	Occupation	Description	Group	SES	NORC	Note[a]
12.02	Librarian	In small institution or public library	3	60	54.58	–
12.06	Nurse and masseur	Graduate	1	46	61.51	–
12.74	Chef	Employed in large first-class hotels	4	15	–	–
12.84	Editor	Small paper, considerable job work	4	82	–	–
12.89	Primary teacher	No college training, 2 yrs. special training	3	–	–	g
12.96	Landscape gardener	–	2	11	–	–
13.08	Grammar grade teacher	Normal graduate expects to make profession teaching	3	72	60.08	g
13.20	Osteopath	Training equal to college grad.	2	96	–	–
13.21	Pharmacist	In town of from 1000–5000 pop.	3	82	60.75	–
13.29	Master mechanic	Thorough knowledge in his field of mechanics	5	27	–	–
13.30	Music teacher	2–4 yrs. special training, not college graduate	4	–	–	h
13.31	Manufacturer	Employs from 10–50 men. Makes simple articles	2	61	65.16	–
13.54	Dentist	Graduate. 2–5 yrs. experience in small town	3	96	–	–
13.58	Art teacher	In high school. 3 or 4 years' special training	4	67	–	–
13.71	Surveyor	Transit man. City or county surveyor	1	48	53.27	–
13.31	Train dispatcher	Must be mentally alert	5	71	–	–
14.45	Land owner & operator	Very large farms or ranches	2	–	–	e
14.70	Musician	Successful player or singer in good company	4	52	14.70	h
15.05	Secretarial work	Private secretary to high state or national officials	4	61	–	–
15.14	High school teacher	Coll. or Normal grad. Not the most progressive	3	72	63.11	–
15.15	Preacher	Minister in town of 1000–5000. College graduate	3	52	68.99	–
15.42	Industrial chemist	Thorough knowledge of the chem. of mfg. processes	2	79	–	–
15.43	Mechanical engineer	Designs and constructs machines and machine tools	2	82	–	–
15.71	Teacher in college	Degree A.B. or A.M. Not the most progressive	3	84	78.26	–

TABLE 5.1 continued

P.E. Value	Occupation	Description	Group	SES	NORC	Note*
15.75	Lawyer	In town of moderate size. Income $1000–5000	1	93	75.66	–
15.86	Technical engineer	Thorough knowledge of the processes of an industry	2	–	–	i
16.18	Artist	High class painter of portraits, etc.	4	67	–	–
16.26	Mining engineer	Thorough knowledge of mining and extraction of metals	1	85	61.61	–
16.28	Architect	Training equal to college grad.	1	90	70.52	–
16.58	Great wholesale merchant	Business covering one or more states	5	70	–	–
16.59	Consulting engineer	In charge of corps of engineers	2	–	–	i
16.64	Educational administrator	Supt. city 2000–5000 Coll. or Normal graduate	5	72	67.40	–
16.71	Physician	6–8 yrs. prep. above H.S. Income $5000 and up	1	92	81.55	–
16.91	Journalist	High class writer or editor	4	82	58.83	–
17.50	Publisher	High class magazine and newspaper or periodical, etc.	5	79	–	–
16.81	University professor	Has A.M. or Ph.D., writes, teaches, and does research	1	84	78.26	–
18.06	Great merchant	Owns and operates a million dollar business	Omit	–	–	–
18.14	Musician	(Paderewski)	Omit	–	–	–
18.33	High National official	Cabinet officers, foreign ministers, etc.	Omit	–	–	–
18.85	Writer	(Van Dyke)	Omit	–	–	–
19.45	Research leader	Like Binet or Pasteur	Omit	–	–	–
19.73	Surgeon	(Mayo Bros.)	Omit	–	–	–
20.71	Inventive genius	(Edison type)	Omit	–	–	–

* P.E. values for combined titles:

a. 3.09, Odd jobs, garbage collector, circus roustabout, hostler, R.R. section hand, day laborer, track layer (group 2)

b. 5.25, Farm laborer, dairy hand (group 3)

c. 5.61, Teamster, drayman (group 2)

d. 8.18, Bricklayer, stonemason (group 1)

e. 10.22, Farm tenants (gardening, truck farming), dairy owner and manager, landowner and operator (group 2)

f. 10.54, Metal pattern maker, wood pattern maker (group 2)

g. 12.99, Primary teacher, grammar grade teacher (group 3)

h. 14.00, Music teacher, musician (group 4)

i. 16.23, Technical engineer, consulting engineer (group 2)

5.1 Observations on the Concept of Intelligence*

As an entry into the problem, let us summarize an exercise that may have more than an antiquarian interest. At one time there was rather wide use of a scale, purportedly measuring the standing of occupations, which was devised in the early 1920's. The Barr scale is briefly described in Volume I of *Genetic Studies of Genius* (Terman, 1925, p. 66): "Mr. F. E. Barr drew up a list of 100 representative occupations, each definitely and concretely described, and had 30 judges rate them on a scale of 0–100 according to the grade of intelligence which each was believed to demand. The ratings were then distributed and P.E. values were computed for all the occupations. The P.E. values express in the case of each occupation the number of units of intelligence which, according to the composite opinion of these 30 judges, the occupation demands for ordinary success." The listing of the occupations, with their descriptions and P.E. values, is reproduced in Table 5.1. (Note that there are actually 120 titles in the list.)

Socioeconomic status scores (Duncan, 1961a) are available for entries in the list of detailed occupations given in the *1960 Census of Population: Classified Index of Occupations and Industries*. Each occupation has a two-digit score ranging from 00 to 96 which was computed on the basis of 1950 Census data on income and education levels prevailing in the occupations. Prestige ratings are available for a group of occupations included in a 1964 study by Hodge, Siegel, and Rossi (Siegel, 1970) at the National Opinion Research Center (NORC). The NORC list indicates the matching detailed Census occupation title, with an indication as to the quality of the match.

The initial task was to match as many as possible of the Barr scale titles with the Census titles and with NORC titles. No attempt was made to match NORC titles with Census titles directly since NORC had already done this. Having arrived at two sets of titles that were assumed to match, the correlations between the Barr scores and the socioeconomic scores and between the Barr scores and the NORC scores were obtained.

The descriptions included with most of the occupation titles in the Barr scale were not always the most helpful in determining a match with one of the other two listings. In some cases the descriptions were such that no match was possible, for example, "Surgeon (Mayo Bros.)," and these titles were deleted. (One wonders what "ordinary success" as a Dr. Mayo might be.) Of the remaining 112 titles, some were combined using an arithmetic average as indicated in Table 5.1 to facilitate a match. In deciding upon matches, seven descriptions

* For a review of the concept of intelligence and of the development of its measurement vis-à-vis performance in modern society, and for estimates of its heritability, see Jensen (1969). An articulation of the role of intelligence in the stratification of contemporary American society and its projected role in the future appears in Herrnstein (1971).

were thought *probably* to bias anyone reading the description so that he would think of only a certain small segment of the workers included under that title, and that this small segment would not be typical. Such an example would be the title "Chef," with the description "Employed in large first-class hotels." A third group of eleven Barr scale titles contained descriptions that were thought *possibly* to bias a person reading a description. For instance, the title "Pharmacist," with a description of "In town from 1000–5000 population," in which case the size of the town was thought possibly to bias a person's judgment as to the amount of intelligence needed to perform the job. There were 64 Barr scale titles for which the description was thought not to influence a person's judgment in a biasing way. These were subdivided into two groups. One group contained 30 Barr scale titles with nonbiased descriptions and with a high quality NORC-Census match; the other group contained 34 Barr titles with nonbiased description but with only a Census match. A final group of Barr scale titles consisted of 14 titles with remote Census matches without regard to the bias effect of the description. Thus five groups of titles were obtained as follows:

Group	Number
1	Thirty titles with good Barr-Census-NORC matches and nonbias descriptions
2	Thirty-four titles with Barr-Census matches and nonbias descriptions but *no* NORC match
3	Eleven titles with Barr-Census matches having *possible* bias descriptions, and which may or may not have an NORC match
4	Seven titles with Barr-Census matches having *probable* bias descriptions, and which may or may not have an NORC match
5	Fourteen titles with *remote* Barr-Census matches which may or may not have NORC matches and/or biased descriptions
Total	Ninety-six titles

An analysis of covariance was made, with the five groups just named as the "treatment variable"; the socioeconomic status scores as the independent variable; and the Barr scale scores as the dependent variable. The five groups of titles were examined to determine if a common slope prevailed, whether the slope differed from zero, and whether one regression line would fit all groups.

The results, seen in Table 5.2, of the three appropriate F tests suggest that there is a common slope, that it is not equal to zero, and that one regression will fit all five groups. Since one regression line will fit all groups, it seems that our worries about bias descriptions and remote matches are without grounds. Thus the five groups can be combined, and the 96 matches obtained can be used in further analysis of Barr-Census comparisons.

TABLE 5.2

F-Tests for Covariance Analysis

Test	N_1 df	N_2 df	Observed F value	P	Table F value for given P	Result
Common slope	4	86	2.20	.95	2.47	Accept
Slope = 0	1	86	143.78	.99	6.94	Reject
One regression line fits all groups	4	90	2.37	.95	2.47	Accept

A similar analysis could have carried out for the Barr-NORC matches as well, but with the conclusions just presented and a look at the scatter plot for the Barr-NORC matches, it was concluded that there was no need for eliminating any of the Barr-NORC matches. In total there are 41 Barr-NORC matches, consisting of the 30 that also have "good" Census matches according to NORC, and 17 additional ones that are thought to have a lower quality match with the Census.

The principal results of the foregoing analysis are as follows: (1) The 96 Barr-Census title matches reveal a correlation between the Barr scale scores and socioeconomic scores of .81. (2) The 47 matches of titles of the Barr scale with NORC titles have a correlation of .91 between Barr scale scores and NORC prestige scores. For these 47 titles the correlation between Barr scores and socioeconomic scores is .90; between socioeconomic and prestige scores, likewise .90.

The purpose of this analysis, of course, was not to ascertain how the "intelligence" of individuals is actually related to the prestige or socioeconomic status of the occupations they pursue. (This topic will be discussed presently.) Instead, we wished to substantiate a point for future reference: 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. In short, we suggest here, with the intention of elaborating the idea later, that "intelligence" is a socially defined quality and this social definition is not essentially different from that of achievement or status in the occupational sphere. It is not mere coincidence, therefore, when psychologists find that "the kinds of occupational criteria which intelligence tests predict best are measures of the complex status characteristic we call *occupational levels*" (Tyler, 1964, p. 176).

None of these results, of course, resolves the ancient question of what intelligence "really is," or of the degree to which intelligence is actually required for the performance of occupations varying in social status or prestige.

Yet it is surely significant that the preconceptions of psychologists about occupational performance in relation to intelligence—preconceptions which, presumably, are built into conventional intelligence tests—so closely coincide with the public's view of the social worth or standing of occupations. If, as sociologists believe, the occupational role is a central element in the structure of a differentiated society, the abilities required for satisfactory performance of that role must be fairly directly involved in the achievement occupational status.

It is not utterly fanciful to reconstruct the history of intelligence testing in a way that it is seldom presented. As we usually think of the matter, psychologists analyzed mental functions and then abstracted a component, "intelligence," which they took to be a general factor in the relative efficiency of human organisms. They then devised tasks apparently requiring this factor in various degrees and incorporated them into standard sets called "intelligence tests." Once such tests were administered to population samples, it was discovered that they were predictive of the amount of success in school and work people would enjoy.

The reconstruction we wish to suggest is the following. Every society implicitly designates certain key roles in which performance is variable, with the quality of the performance being a basis for the assignment of status. (Other statuses, of course, may depend upon factors besides performance—the so-called ascribed statuses.) Where the society is one with a complex division of labor, many differentiated occupations are pursued, and these occupations are highly salient among the key roles whose pursuit is a basis for status achievement. Adequate performance in a high status occupation is taken by the social group as prima facie evidence of social capability. However, poor performance in a high status occupation leads to uncertain tenure of the status, and performance—whether good, bad, or indifferent—of a low status occupational role is not seen as providing any sizable increment to consensual estimates of a person's value to society. What we call "occupational prestige" corresponds to an unmistakable social fact. When psychologists came to propose operational counterparts to the notion of intelligence, or to devise measures thereof, they wittingly or unwittingly looked for indicators of capability to function in the system of key roles in the society. What they took to be mental performance might equally well have been described as role performance. Indeed, it was clear in the minds of the pioneers of mental testing that they wished to tap capacity to perform well in another social situation—that of the school. For their immediate purposes, it was unnecessary to expand upon the sociological observation that the school is itself (among other things) a primary mechanism for selecting incumbents of occupational roles.

Our argument tends to imply that a correlation between IQ and occupational achievement was more or less built into IQ tests, by virtue of the psychologists' implicit acceptance of the social standards of the general populace. Had the first IQ tests been devised in a hunting culture, "general intelligence" might well have turned out to involve visual acuity and running speed rather than vocabulary and symbol manipulation. As it was, the concept of intelligence arose in a society where high status accrued to occupations involving the latter in large measure so that what we now *mean* by intelligence is something like the probability of acceptable performance (given the opportunity) in occupations varying in social status.

This argument, however, does not imply that the correlation of IQ with occupational status—assuming the latter to be measured on a scale of prestige or (what is nearly equivalent) socioeconomic rank—will be perfect. First, there are many social contingencies (just alluded to by the term "opportunity") which may militate against a matching of capacity to perform occupational roles and actual performance. Second, any test is a small sample of the almost unlimited sorts of personal assessments that could be made; it is thus a fallible basis of inference.

It is an empirically contingent question, therefore, as to how well occupational achievement can actually be predicted from test scores. If our argument were entirely cogent, we might suppose that if all the "social contingencies" bearing upon occupational achievement were properly taken into account, residual variation would be solely due to "intelligence." To accomplish a demonstration of this hypothesis, however, we should require a model that correctly locates intelligence itself in a causal complex and correctly specifies its role in status achievement vis-à-vis the many other contingent factors. One way of stating our purpose in this research is to indicate that we are trying to make progress in this direction. Naturally, we do not expect any such decisive result as that suggested by the statement of the ultimate objective of research.

5.2 Correlates of Intelligence

In a search for psychological data on sizable propulations which could be roughly matched with our demographic data on occupational achievement, we were pleasantly surprised to learn of a very substantial body of information summarized by Byrns and Henmon (1936). They report scores of some 100,000 Wisconsin high school seniors given selected tests of "scholastic aptitude" during 1929–1933. The summary is in terms of ten broad groups of parental occupation and 77 specific occupation titles. Three different tests had been used, so the authors aggregated the results only after making a percentile transformation. Their Table V shows for each of the 77 parental occupations the number of students tested, and the first-, second-, and third-quartile scores of students identified with that parental occupation.

Our interest was in the correlation of students' scores with the *status* of the parental occupation as measured by Duncan's (1961a) socioeconomic index of occupational status. To study this correlation, we had to match the occupation titles given by Byrns and Henmon with the Census titles for which the index is defined. This led us to make certain omissions, such as students whose parents were "retired" or classified in a "miscellaneous" category. We also omitted a few occupations, containing only a small number of students, where the title strongly suggested exclusive application to the female parent, for example, "nurses" and "dressmakers." Altogether, we retained the data pertaining to 88,883 of the original 100,820 students. In some instances we had to make a combination of two titles given by Byrns and Henmon to achieve approximate comparability with a Census category, and sometimes Census titles had to be combined. We ended up with 64 occupation titles for use in our restudy of Byrns and Henmon's data. The process of matching occupation titles is inevitably somewhat arbitrary and subjective, but in the light of our experience it is difficult to believe that the results would have been greatly different in the hands of any knowledgeable investigator. For future use, however, we offer to psychologists the suggestion that their data could be more generally useful if some care were taken to render occupation and other social categories consistent with those employed in official statistical sources.

One further manipulation was required before we could compute the statistics of interest to us. Since Byrns and Henmon used percentiles as their score value, they obtained a roughly rectangular distribution of scores—approximately 10 percent being scored 0–9, 10 percent 10–19, and so on, to 10 percent in the top interval, 90–100 (see their Table I). We assumed that the underlying score distribution was normal, and converted the median percentile scores in their Table V to normal deviates—or, actually, to probits, making use of Table IX in Fisher & Yates (1948). The probit values were then transformed to standard scores with mean 100 and standard deviation 20. For example, Byrns and Henmon report that children of Druggists had a median percentile score of 67.6. In the normal distribution 67.6 percent of the population falls below a score corresponding to .4565 standard deviation units above the mean or a probit value of 5.4565. This probit value multiplied by 20 is 109.13, which (rounded to 109) we took to be the mean standard score of children of Druggists.

Let X_j be the occupational status score of the jth occupation on Duncan's scale, $\bar{Y}_j$ be the mean standard score of children of parents classified in the jth occupation, and n_j the number of such children in the data of Byrns and Henmon, so that $\Sigma_j n_j = N = 88{,}883$. If Y_{ij} is the standard score of the ith child in the jth occupation category, we have

$$\bar{Y} = \sum_j \sum_i Y_{ij}/N = \sum_j n_j \bar{Y}_j/N = 100.607,$$

differing slightly from 100, presumably because of the omissions noted above and/or errors of rounding. By assumption,

$$\text{Var}(Y) = \sigma_Y^2 = 20^2 = 400,$$

an assumption we cannot check numerically because of the way in which the data are tabulated.

We find that

$$\bar{X} = \sum_j n_j X_j/N = 32.41,$$

a value whose representativeness we shall assess presently; and

$$\text{Var}(X) = \sum_j n_j(X_j - \bar{X})^2/N = 500.99,$$

whence $\sigma_X = 22.38$, a value likewise subject to an external check.

Finally, we require

$$(\text{Cov}\, Y, X) = (1/N)\sum_j n_j \bar{Y}_j X_j - \bar{Y}\bar{X} = 90.47.$$

From the foregoing, we obtain immediately the regression coefficient $b_{YX} = .1806$, the intercept $a_{YX} = 94.75$, and the correlation $r_{YX} = .2021$, so that $r_{YX}^2 = .041$. We may also compute the squared correlation ratio, $eta_{YX}^2 = .052$, taking each of the 64 occupation titles as a distinct category.

With this large a sample, the difference, eta_{YX}^2 minus $r_{YX}^2 = .011$, is no doubt too large to attribute to sampling error. The scatter diagram (Fig. 5.1), however, gives only the slightest suggestion of curvilinearity: Most of the variation of occupation-specific means from the regression line is simply scatter of particular occupations at comparable status levels.

In view of the uncertainty about the status score for the occupation, Farmer, and the very substantial number of farmers' children in the Wisconsin data, we are pleased to have observed the near coincidence of the actual mean and the regression estimate. The mean standard score for farmers' children is 96. With an occupational status score of 14, the regression estimate for this group comes out at 97.3.

Our correlation $r_{YX} = .20$ may be compared with the result stated by Byrns & Henmon (1936, p. 287): "the correlation between mental ability of the student and the rank of the parental occupation, here discovered, for the entire group of students is only +.18." Although we are not quite sure how the latter value was computed, we are reassured by the fact that our manipulations have not resulted in any pronounced distortion of the conclusion originally reached.

In using the foregoing estimates from the data of Byrns and Henmon, one must bear in mind the selectivity involved in their definition of the study population. From the 1930 Census we learn that 53.7 percent of 17-year-old

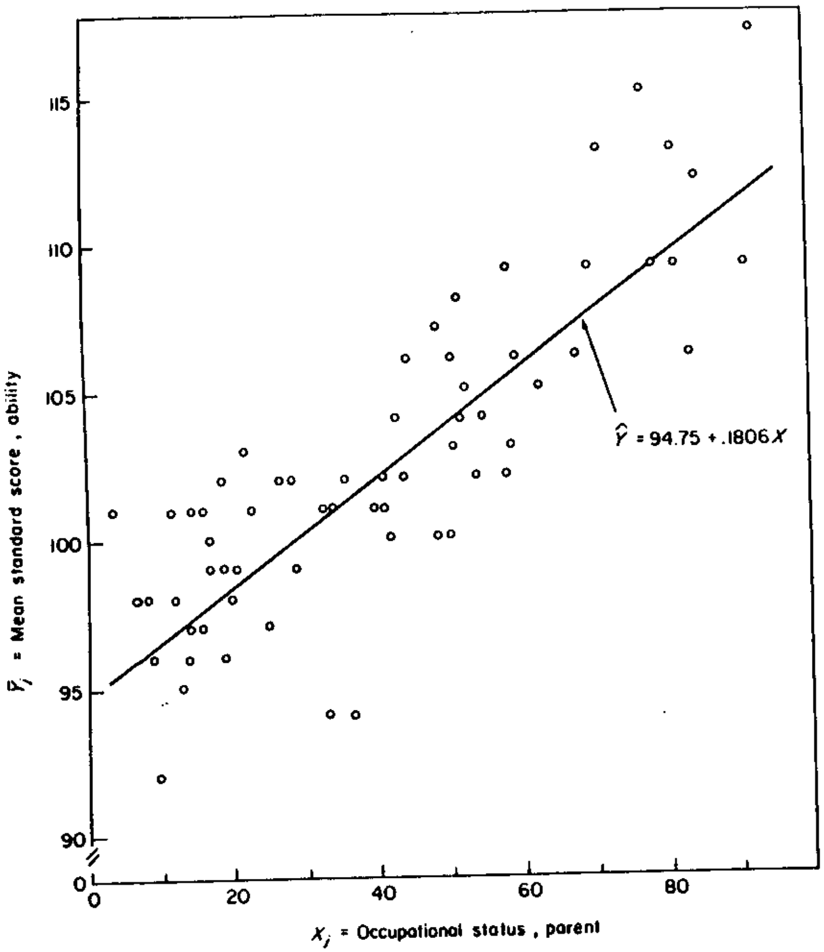

Fig. 5.1. Scatter plot, mean standard scores of ability on occupational status of parent, Wisconsin high school seniors, 1929–1933 (after Byrns & Henmon, 1936).

boys in Wisconsin were enrolled in school. Not all of them, however, were high school seniors, the group covered by the testing program. We probably can secure a better estimate of the coverage of the testing program by considering 1940 Census data on educational attainment of Wisconsin men 25–29 years old, who were, of course, of high school age around 1930. Of these men, 38.0 percent are reported as high school graduates and an additional 4.6 percent as having completed 3 years of high school. In round numbers, therefore, the testing program from which Byrns and Henmon secured their data must have covered about 40 percent of the Wisconsin boys reaching age 18 in the period

1929–1933. The authors give no indication of how far the testing program may have fallen short of covering the target population on account of absences from school and the like.

To evaluate the occupation statistics derived from Byrns and Henmon, we consider national OCG data on native white men 47–51 years old in March 1962 (who were, therefore, 15–19 years old in 1930). For comparability with the population studied by Byrns and Henmon, we exclude those who completed less than 4 years of high school. For this select group of high school graduates we find the mean of father's occupational status is 35.01 (vs. 32.41 derived from the Wisconsin data), with a standard deviation of 23.74 (vs. 22.38). The agreement seems satisfactory inasmuch as we have no reason to assume strict equivalence of the two populations.

Having considered the correlation of mental ability with one important item of socioeconomic background, we turn to the problem of estimating its correlation with measures of achievement. A search of the literature suggests that the best historical data for a general population relating IQ measured at an early age to subsequent educational attainment are those compiled by Benson (1942). She followed up 1989 pupils in the sixth grade of 64 elementary schools in Minneapolis who had been given the Haggerty Intelligence Examination: Delta 2 in April 1923. Records of subsequent achievement (highest grade completed) were obtained for 1680 cases.

Benson reports, "A product-moment coefficient of correlation of $.57 \pm .01$ was obtained between IQ and grade level attained" (p. 164). Her Table I is a cross-tabulation of IQ (10-point intervals) by six levels of attainment. We scored the latter as follows, to conform with our practice in anlyzing OCG data:

3: "Did not enter high school" (but presumably finished at least sixth grade and, for the most part, eighth grade)
4: "Entered high school but did not graduate"
5: "Graduated from high school but did not enter college"
6: "Entered college but did not receive any degree"
7: "Received bachelor's degree"
8: "Took graduate work or received advanced degree"

Using these scores and the midpoints of IQ intervals, we found a correlation of .542 ($r^2 = .294$). The regression of education on IQ was .0321, with an intercept of .99. Mean IQ was 112.4 with a standard deviation of 19.38. (We have ignored the "Stanford-Binet equivalents" also given by Benson; these have a somewhat smaller standard deviation.)

As a rough check on the plausibility of Benson's follow-up data, we looked at 1940 Census data on educational attainment of persons 25–34 years old in Minneapolis. The comparison with Benson's distribution in Table 5.3 is

TABLE 5.3

Percent Distribution by Educational Attainment, for Persons 25–34 Years Old Living in Minneapolis in 1940 and for Sample Studied by Benson

Years of school completed	1940 Census	Benson
Elementary		
None to 5	1	–
6 to 8	20	16
High School		
1–3	23	35
4	36	31
College		
1–3	11	10
4	6	6
5 or more	3	2
Total*	100	100

Source: *1940 Census of Population*, Vol. IV, Part 3; Benson (1942, Table I).
* Excludes attainment not reported.

moderately reassuring. She, of course, missed the 1 percent of children failing to reach sixth grade. The 309 cases not located in the follow-up were known to be negatively selected on IQ. A median of 108 is reported for the 1989 cases originally tested as against 112 for the 1680 cases followed up. We infer that the mean IQ of the 309 lost cases was around 86. Disproportionate numbers of them probably were early dropouts. This fact may help to account for the underrepresentation of persons failing to enter high school in Benson's sample, but it leaves us puzzled at the overrepresentation of those completing 1–3 years of high school. An alternative explanation, of course, is response error in the Census data or lack of comparability between the two sources. In illustration of the latter, it seems likely that many of Benson's respondents who "entered high school but did not graduate" actually dropped out before finishing the ninth grade. In that event, the Census type of question would classify them as Elementary, 8 years, rather than High School, 1–3 years.

Altogether, one can feel considerable confidence when taking a value of .5 or .6 as the correlation between IQ and educational attainment in cohorts completing their schooling during the 1930s. Interestingly enough, this seems to be about the value obtained in correlating IQ scores obtained on adults with their past history of schooling—a point we can check more carefully with the CPS-NORC data and other sources.

We present next some calculations on data summarized by Harrell & Harrell (1945), whose paper shows summary statistics of the AGCT (Army General Classification Test) scores of 18,782 white enlisted men in the Army Air Forces

Air Service Command during World War II. The statistics are classified into 74 previous civilian occupations of these men. The Harrell report contains no information on age, educational attainment, geographic origin, or other social characteristics of the sample. Apparently, occupations infrequently represented in this population were simply omitted from the tabulations. To match Census occupation titles (approximately), it was necessary to combine certain of the Harrell categories. Hence, the present analysis concerns 69 occupation groups.

The AGCT was designed to have a mean of 100 and standard deviation of 20. The Harrell sample as a whole yields a mean of 106.6 with standard deviation 19.1. Evidently, selection into the Air Force enlisted man population involved some screening for intelligence.

When the 69 occupations are scored on Duncan's (1961a) status scale, we obtain a mean of 31.8 and standard deviation of 19.2. These figures suggest that the sample is not highly unrepresentative of civilian occupations of young men. Duncan & Hodge (1963), for example, report a mean of 35.5 with standard deviation 22.1 as of 1940 for a Chicago sample of white men 25–34 years old in that year. The Harrell sample may, therefore, underrepresent men at the extremes of the occupational status distribution.

In Table 1 of the Harrell paper we have the mean and standard deviation of the AGCT scores of men in each occupation. It is, therefore, easy to compute the within-occupation and between-occupation sums of squares; the total sum of squares follows at once. We find eta^2 of AGCT on occupation is .2288. The correlation coefficient of AGCT with occupational status is .4241; hence r^2 is .18. The regression coefficients are .4264 for occupational status on AGCT and .4218 for AGCT on occupational status. In causal models we would probably wish to think of occupational status as a function of intelligence. Hence the former regression is perhaps the more relevant one. The Harrell table, however, shows mean AGCT for given occupation. It is only the latter regression, therefore, that we can inspect for evidence of curvilinearity. The scatter diagram (Fig. 5.2) shows little evidence of systematic departure from a linear relationship.

A second set of Army data for white enlisted men is available in a report by Stewart (1947). A similar collection of data for civilian samples tested with the GATB (General Aptitude Test Battery) will also be studied here (United States Bureau of Employment Security, 1962).

Stewart's data pertain to 81,553 white enlisted men in 227 different occupations. Occupations infrequently represented in her original sample were omitted from the published report. The occupational categories for Stewart's data are, therefore, considerably more detailed than those used in Harrell and Harrell's data described previously.

Some occupations on Stewart's list were discarded: specifically, all titles

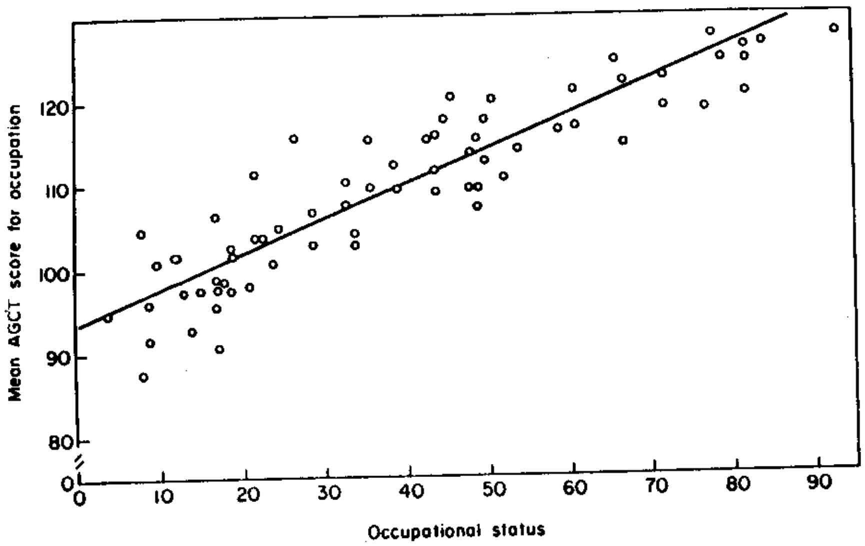

Fig. 5.2. Regression of AGCT score on previous civilian occupational status, as measured by scores on Duncan (1961a) scale.

with a "student" prefix, such as "Student, Medicine," and a few which could not be given a Census code. Stewart reports values of percentiles 10, 25, 50, 75, and 90 for each occupation. We used only the median (P_{50}) values and treated them as occupation-specific mean scores. Since Stewart does not report a standard deviation for her whole sample, we took it to be 20 and used this figure in calculating the variance and sum of squares of AGCT scores. The mean AGCT score for the 62,233 cases included in our calculations is 101.6. The statistics on occupational status in this sample are mean 25.9 and standard deviation 18.6. Thus Stewart's sample has rather lower means on both AGCT and civilian occupational status, but the standard deviation of the latter is quite comparable with the value observed in the Harrell and Harrell material.

The following tabulation compares the regression statistics obtained from the two sources:

	Stewart	Harrell
Correlation, AGCT and occupational status	.446	.424
Regression, AGCT on occupational status	.481	.422
Regression, occupational status on AGCT	.414	.426
Eta^2, AGCT on occupation	.253	.229

In view of the differences in population coverage and the detail of the occupational classification, the similarity between the two sets of results is remarkable. It is difficult to foresee any use for these results where the differences will be of material consequence.

Turning to the civilian data, we consider GATB scores on Aptitude G (Intelligence) collected by the United States Employment Service. Like the AGCT, this score is designed to have mean 100 and standard deviation 20 in the general population. The sample providing data for specific occupations, however, is not a cross-section sample but a collecion of samples of specific occupations obtained in what appears to have been an ad hoc and expedient fashion. While the occupation titles are extremely specific, they do not cover the total occupation structure to the degree that the military data do.

The source publication gives sample size, means, and standard deviation for each of the specific occupations. We deleted a considerable number of occupations the samples for which were predominantly female. The mean for all of the 17,173 cases covered in the source was 100.36; for the 7858 deleted cases it was 90.68; for the 9315 cases studied here it was 108.53. (These figures, of course, are not relevant to the question of general sex differences in intelligence.) Despite the upward bias of this sample's mean, the standard deviation remained 19.99, or effectively 20.

There is likewise an upward bias in the distribution of status scores for the occupations included. The mean for the group studied here is 43.85 with a standard deviation of 21.86.

If the sample for each specific occupation were representative, there would be no bias in the regression coefficient of intelligence (G score) on occupational status. This regression is .504, which may be compared with the regression of AGCT on occupational status of .481 from the Stewart data and .422 from the Harrell data.

The GATB analysis provides a correlation coefficient of .551 between intelligence and occupational status, which is somewhat higher than those obtained from the two AGCT series, .446 and .424, respectively. In view of the (probable) nonrepresentativeness of the occupations covered in the GATB data, one would not accept this as an estimate for the general population.

An interesting feature of the GATB data is the high value of eta^2 for the correlation ratio of G scores on occupation. A value of .490 is obtained, in contrast to .253 and .229 for the two AGCT studies. Yet we have seen that the linear regression of intelligence on occupational status score is not markedly higher in the GATB data than in at least one of the AGCT sets. Evidently, the detailed occupational coding and/or the sampling technique of the GATB study produced a good deal of interoccupation variation in intelligence not captured in the military data. However, this additional variation is not particularly related to occupational *status*.

In consequence, the GATB occupation-specific means show a good deal more scatter around the regression of intelligence on occupation than do the

AGCT means, even though the regression itself is much the same (see Fig. 5.3).

To summarize: Two sets of military and one set of civilian data give essentially consistent indications of the degree of relationship between tested intelligence and occupational achievement. If anything, the civilian data suggest a slightly stronger relationship. The difference, if not due to technicalities solely, could be due to the fact that the military data refer to former civilian occupations of very young men, many of whom had doubtlessly not yet established their occupational career lines at the time of induction.

5.3 Preliminary Models

The information in hand, to this point in the discussion, is summarized in Table 5.4. We wish now to indicate how these data may be used in securing an extension of the basic model of the process of occupational achievement. Figure 5.4 recapitulates one set of estimates on this model from Blau & Duncan (1967, p. 174). For the moment we are taking as the terminal occupational status the "first job" reported in the OCG data, assuming it is the nearest possible approximation to the data on previous civilian occupation given in the

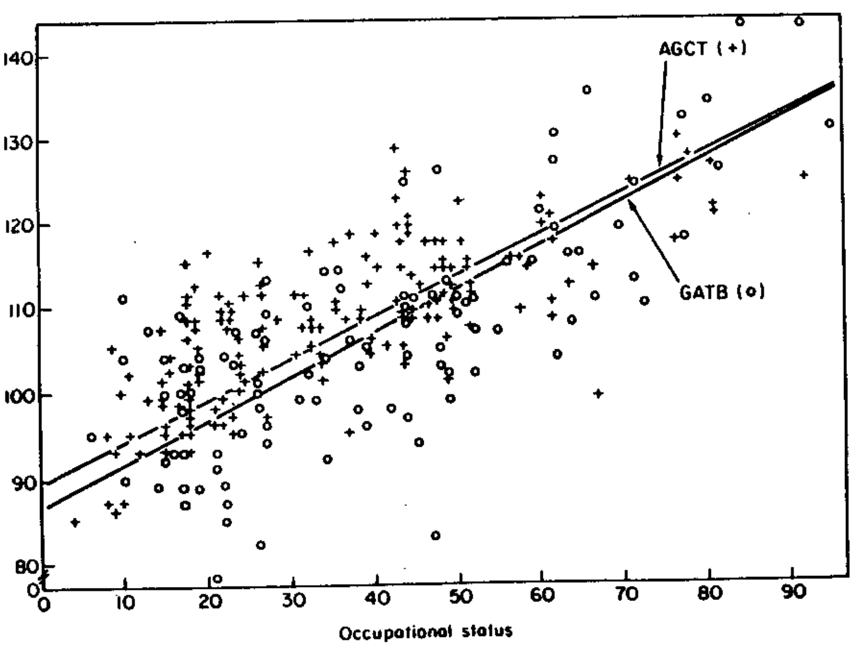

Fig. 5.3. Regression of mental ability on occupational status (Duncan score). AGCT data from Stewart (1947); GATB data from United States Bureau of Employment Security (1962).

TABLE 5.4

Correlations Used in Estimating Path Coefficients in Fig. 5.5[a]

Variable	Q IQ	X Father's occupation	V Father's education	U Respondent's education	W First job
Q, Respondent's intelligence	–	.20	.25	.54	.43
X, Father's occupational status	1	–	.52	.44	.42
V, Father's education	2	5	–	.45	.33
U, Respondent's education	3	5	5	–	.54
W, Status of first job	4	5	5	5	–

[a] Key to sources:
1. Byrns and Henmon (1936)
2. Unpublished data of W. H. Sewell (WISC data set)
3. Benson (1942)
4. Average of two sets of AGCT data (Harrell & Harrell, 1945; Stewart, 1947)
5. OCG study, all men 20–64 years old

studies of military mental ability tests. The status of "first job" is represented as being dependent upon educational attainment and status level of father's occupation. Educational attainment, in turn, is taken to depend upon the father's occupational status and his educational attainment—although one might well have used alternative measures of socioeconomic background. Both occupational statuses—the father's (as of respondent's age 16) and the respondent's first job—are scaled on Duncan's socioeconomic status index for detailed occupations. Educational attainment is the number of years of regular schooling completed.

The models are linear causal systems which are hypothesized to account for the observed associations among measured variables. The path coefficients shown for Fig. 5.4 were estimated from data for men 20–64 years of age.

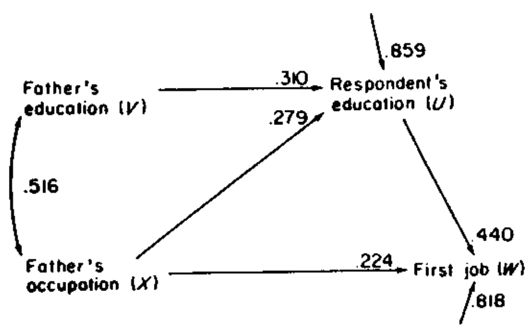

Fig. 5.4. Basic model, with estimates based on OCG data for all men 20–64 years old.

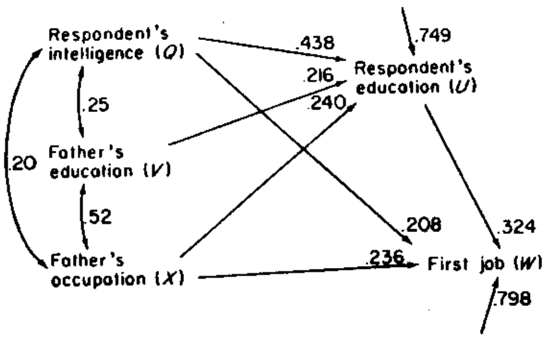

Fig. 5.5. Extended model based on data in Table 5.4.

Slightly different results are obtained in making the estimates for different populations, such as men with nonfarm background, white men, or men in a more restricted age range. Such differences are not at issue here, however, as we have introduced Fig. 5.4 only for illustrative purposes.

The extension we wish to entertain involves considering measured intelligence as a background factor, along with the socioeconomic measures. Thus Fig. 5.5 represents the status of the first job as depending directly upon educational attainment, both directly and indirectly upon father's occupation and respondent's intelligence, and only indirectly upon father's education. Respondent's education, in turn, depends upon intelligence and the two socioeconomic background items. It should be noted that the model requires no assumption concerning the nature of the linkage between socioeconomic background factors and intelligence. Such a correlation could arise on the basis of either genetic or social mechanisms or, more likely, a combination of both. For the purposes of the present model, it suffices to recognize that the correlation exists. A quite different model would be required to represent hypotheses about how the correlation is produced.

The data used in estimating path coefficients for Fig. 5.5 are shown in Table 5.4. At this juncture we are venturing to combine into one model estimates of correlations obtained for several different populations. This tactic will also be followed in our subsequent work. Ideally, we would resort to this procedure only if we had equally representative and reliable samples of the very same population. This condition will seldom be met, and we shall have to assume comparability of data from different sources when there is every reason to believe such comparability does not strictly hold. *From this standpoint, all results obtained on this procedure had best be regarded as hypotheses for ultimate verification upon a single population for which all relevant measures can be obtained.* Even though our procedure is hazardous, it is not actually different from the informal practice of investigators who draw conclusions by com-

paring information developed in two or more studies. Or, rather, the difference is that we are here undertaking formally what is common in informal practice. Presumably, any liability of the procedure should be more apparent when it is controlled by the formalism of an explicit model than it would be in the absence of such a control.

The importance of taking intelligence explicitly into account is suggested by a comparison of Fig. 5.4 with Fig. 5.5. As far as occupational achievement is concerned, inclusion of intelligence as a background factor does not markedly increase the proportion of "explained" variance in status level of the first job. The residual factor for W (first job status) is .818 in Fig. 5.4 as compared with .798 in Fig. 5.5. Translating these into proportions of variance not accounted for (the square of the residual path), we have 67 percent of the variance of W not accounted for in Fig. 5.4 as against 64 percent in Fig. 5.5. This result dashes any hope that availability of intelligence test scores will enable the investigator to improve markedly the prediction of early occupational achievement as compared with what he can do with education and socioeconomic background factors alone.

However, Fig. 5.5 gives us a rather different interpretation of the nature of the process of status achievement from that implied by Fig. 5.4. A substantial direct path from intelligence to first job must be entered into the system. When this is done, the apparent direct effect of education is diminished, for in Fig. 5.5 education is represented as affecting the first job only insofar as it operates independently of intelligence, as well as socioeconomic background. Phrased otherwise, the apparent effect of education in Fig. 5.4 includes some variation in first job status that is actually due to intelligence, given that intelligence affects educational attainment.

This result assumes a certain importance in view of the current interest in estimating "returns from education" in an economic sense. Conventional Census data reveal that amount of income earned rises with increments to years of schooling. Economists studying the rate of return to education have noted, however, that number of years of schooling partly reflects differences in ability. It has been observed, moreover, that the effect of education on income is transmitted, in considerable measure, via occupational level. At this point we cannot yet include income as a further output of Fig. 5.5, although a model with this feature is presented subsequently. However, we are in a position to look at the respective roles of education and intelligence as determinants of occupational status.

If we look at the gross association of education with occupational status, as measured by the simple (zero-order) correlation between the two, we find the substantial value of .54 (see Table 5.4). Figure 5.4 suggests, however, that education is operating in part to transmit the effect of socioeconomic background so that its direct effect in a model incorporating such background items is reduced to .44. Even this figure is seen to be an overestimate in the light of Fig. 5.5, where with both socioeconomic background and intelligence included in the system, the direct effect of education shrinks to .32. In this model, education is estimated, then, to have a direct effect of .32 and an indirect effect, due to its correlation with antecedent determinants of first job, of .22 (the original simple correlation, .54, less the direct effect, .32). To be sure, the remaining direct effect is still substantial, and its significance is enhanced by the fact that it is measuring the impact of education on occupational achievement independently of some of the obvious determinants of both education and occupation level.

The reduction in the apparent role of education as between Figs. 5.4 and 5.5 contrasts with the lack of change in the direct effect of father's occupation in the two models. Since the correlation between father's occupation and respondent's intelligence is only .2 (according to the estimate used here), inclusion of intelligence in the model hardly affects the estimate of the direct effect of father's occupation; we find it to be .22 in Fig. 5.4 and actually a little higher at .24 in Fig. 5.5. Again, it should be remembered that this is an estimate of the *net* impact of father's occupation on respondent's first job, taking into account its correlation with respondent's intelligence and the fact that it works partly via its influence on education—that is, net of these indirect paths of influence. The net or direct effect of father's occupation is, of course, less than its gross association with first job, which comes to .42 (see Table 5.4).

It will have been noted that no direct effect of father's education on first job is shown in the model. This is the case because such direct effect is very nearly zero and not statistically significant. A version of Fig. 5.4 which included a direct path from V to W yielded a coefficient of .014, rather less than the standard error of the coefficient (Blau & Duncan, 1967, p. 174).

The inclusion of intelligence in the model not only puts a new—and presumably more realistic—interpretation on the roles of education and socioeconomic background factors in occupational achievement; it also leads to a more adequate accounting for the variation in education. In Fig. 5.4, where only socioeconomic background was considered, the residual factor for U is .859, implying that 74 percent of the variance in educational attainment is unexplained. In Fig. 5.5, with the residual path of .749, this figure is reduced to 56 percent. The model, even so, far from exhausts the variance in education.

For advocates of equal opportunity, it may be reassuring that intelligence is clearly more important than socioeconomic background as a determinant of educational attainment. Its direct effect is .438, its indirect effect, due to correlation with socioeconomic background, .102 (the sum of these, .54, being the simple correlation between education and intelligence). However, it is clear

that socioeconomic background influences how far boys go in school, quite apart from differences in measured ability. This is apparent from the path coefficients for V and X, father's education and occupation.

To analyze the matter in a slightly different way, we can note that, by itself and including its role in mediating effects of socioeconomic background, intelligence accounts for 29 percent of the variance in schooling (the square of .54, the zero-order correlation between the two variables). In combination, intelligence and the socioeconomic background items account for 44 percent of the variance in education. The increment of 15 percentage points (44–29) is the net contribution of socioeconomic background, as measured by the two characteristics of the father, quite apart from any indirect effects of socioeconomic origins operating via intelligence. Whether this net influence amounts to an inequality of "opportunity" or represents inequality in some kinds of social, economic, or psychological resources which a family may bestow upon the child remains to be estimated.* Experiences with models of this kind suggests that inclusion of explicit measures of economic resources, such as family income, would not alter greatly the estimate of the combined impact of all socioeconomic factors on educational attainment. What other kinds of "resources" should be postulated is a separate question. In subsequent discussion, we bring one other suggestive item of information to bear upon this difficult question.

The next complication of the basic model to be considered arises from the introduction into it of a variable reflecting family structure, to wit, number of siblings. For the remainder of this discussion of preliminary models we shall omit consideration of occupational status so that the model presented next has only one output variable, educational attainment. Fig. 5.6 is taken from the work of B. Duncan (1965). Her work was based on data for a subgroup of the OCG sample, consisting of native white males 27–61 years old in 1962. The model happens to include a variable, labeled "intact family," which we shall subsequently ignore; but its inclusion probably has little effect on the path

* Estimates of talent loss (under various definitions) for men and women in the WISC data set appear in Sewell (1971) and Sewell, Hauser, and Shah (forthcoming). For young men of equivalent mental ability, there are fewer from lower socioeconomic origins going on to higher education, completing college, and entering postgraduate study than from higher status origins. In the highest IQ quartile, more than 1.5 times as many students from the highest quartile on socioeconomic background as in the lowest socioeconomic quartile go on to post-high-school education. Somewhat greater talent loss is experienced by lowest status, lowest ability Wisconsin students as compared to highest status, lowest ability boys and girls. Sewell and his colleagues interpret their data as demonstrating educational discrimination by social origins (and sex) at all levels of higher education, even in a state with an outstanding record of providing public and private scholarships and low tuition rates in its diverse system of public higher education.

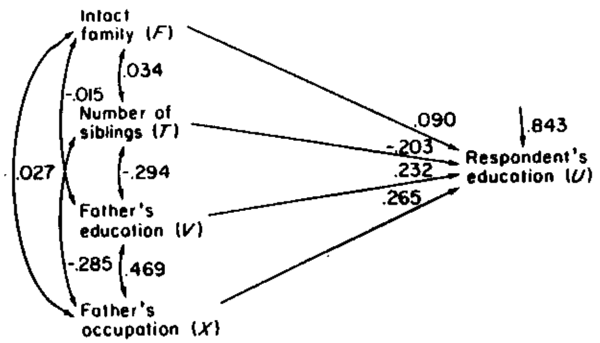

Fig. 5.6. Basic model of educational attainment with estimates for native white men 27–61 years old (B. Duncan, 1965b, Chapter 3).

coefficients for the other determinants of educational attainment since its correlations with them are so low. The intercorrelations among the background variables are shown in the diagram. Their respective correlations with the dependent variable are as follows, $r_{UF} = .087$; $r_{UT} = -.344$; $r_{UV} = .441$; $r_{UX} = .434$. When the population is limited to native white males, father's education diminishes in relative importance (as compared to father's occupation) as a factor in educational attainment of respondents. This is apparent both in the simple correlations and in the path coefficients for Fig. 5.6. An appreciable negative path for number of siblings is estimated, and this factor is itself negatively correlated with the other background factors.

The new interpretation of these data which is required when intelligence is taken into account appears as Fig. 5.7. The crucial item of information is the correlation of intelligence with number of siblings. This correlation is not available in any of the sources from which we have obtained other correlations

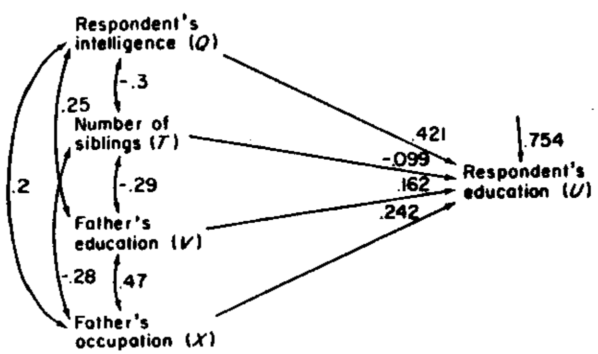

Fig. 5.7. Educational attainment as a function of family background and intelligence.

for intelligence. However, the correlation of IQ with number of siblings has been studied rather extensively (Anastasi, 1956), and there is a sizable and controversial literature on the interpretation of this relationship. The most representative figure for the observed correlation between number of siblings and standard intelligence tests in unselected populations seems to be about $-.3$, and this value has been selected for the purpose of our illustrative calculations. In Fig. 5.7 we are unable to include the "intact family" variable, but as already indicated, it seems unlikely that its insertion into the model would alter the other paths appreciably.

Close comparison between Figs. 5.5 and 5.6 are not warranted since the latter pertains to a somewhat different population from the former. However, there is general resemblance between the two except, of course, for the additional path to education in Fig. 5.6 and the omission from it of the first job. With intelligence and socioeconomic background held constant, number of siblings retains a significant direct effect on schooling. It is interesting, however, to see the extent to which the sibling variable operates via other variables in the system. Its zero-order correlation with educational attainment, noted previously, is $-.344$; with socioeconomic background items held constant, the direct effect shrinks to $-.203$ (in Fig. 5.6); and with intelligence as well included among the background factors there is a further shrinkage to $-.099$. A full interpretation of this outcome would require a considerable elaboration of Fig. 5.7, to "explain" the intercorrelations of the background variables taken as given in that model. Such an explanation would raise complicated issues of "heredity vs. environment," for there is no agreement on the extent to which the inverse correlation of number of siblings and intelligence represents environmental effects on intellectual development as over against dysgenic fertility patterns (Burt, 1947; Nisbet, 1953).

Since Fig. 5.7 includes both intelligence and socioeconomic background, the path coefficient of approximately $-.1$ for number of siblings must represent other influences than these. The most obvious observation is that children in large families enjoy lesser economic resources per head than children in small families, given that the families are at the same socioeconomic level. The same may be true of other resources as well. In a large family, parental aspirations may not be as sharply focused on any one child, designated at random, as in the small family.

An elaboration of Fig. 5.7 on a quite conjectural basis appears in Fig. 5.8. Here we raise the question of how similar the educational outcomes would be for two brothers in the same family. We assume that since the two brothers have the same father and the same number of siblings, values of T, V, and X are the same for them. This is actually a simplification, because for example, father's occupation is specified as of the date the respondent was 16 years old, and two brothers would not have attained that age in the same year, apart from the case

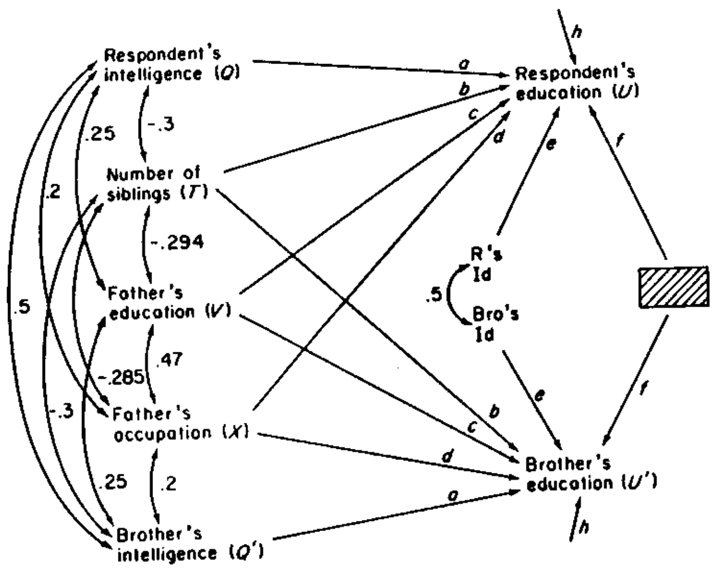

Fig. 5.8. Correlation between siblings in regard to educational attainment, interpreted in the light of Fig. 5.7.

of twins. We assume, moreover, that the socioeconomic background factors act in precisely the same way for the two brothers so that there is only one set of path coefficients applying either to respondent's education or to brother's education. A further assumption is that intelligence of either brother is intercorrelated with background items in the same way as for the other brother and acts in the same fashion on schooling.

Finally, the crucial assumption concerns the correlation between the intelligence scores of the two brothers. This particular correlation has been studied rather extensively. A recent review article instances no less than 35 inquiries into the correlation between siblings in intelligence (Erlenmeyer-Kimling & Jarvik, 1963). Rather widely varying figures, from less than .3 to nearly .8, have been obtained. The median of the 35 correlations is, however, .49. Whether or not this is a mere coincidence, the empirical correlation is very close to the theoretical genetic correlation of .5 between siblings which follows from highly simplified assumptions. Without commenting on the implications of this coincidence, we shall take .5 as the correlation between brothers' intelligence scores.

In Fig. 5.8 it should be noted that educational attainment is not assumed to be directly affected by brother's intelligence so that there is no path from Q' to U or from Q to U'. Such a direct effect would be theoretically anomalous. As we shall see, however, this assumption does not require that the simple correlation between Q' and U or Q and U' be zero.

When either brother is considered separately, therefore, Fig. 5.8 merely repeats Fig. 5.7, and we shall therefore transfer the respective path coefficients from the latter to the former. Hence for the direct effect of intelligence on schooling (for either brother) we have $a = .421$; for the effect of number of siblings, $b = -.099$; for the effect of father's education, $c = .162$; and for father's occupation, $d = .242$. For the moment, let us disregard both the paths labeled e and f. Then the residual path, h, is the same as in Fig. 5.7, or .754, implying that 57 percent of the variance in educational attainment is not accounted for by the model (whichever brother is in question).

If Fig. 5.8, omitting paths e and f, were literally correct, we could derive the correlation between educational attainments of the two brothers, making use of the appropriate theorem from the theory of path analysis:

$$r_{UU'} = ar_{QU'} + br_{TU'} + cr_{VU'} + dr_{XU'}.$$

This expression includes a correlation ($r_{QU'}$) which is not among our empirically given coefficients. But it, too, is readily obtained from the model, assuming the model to be correct:

$$r_{QU'} = ar_{QQ'} + br_{QX} + cr_{QV} + dr_{QT}.$$

We first compute $r_{QU'}$ as .329; inserting this value into the earlier formula (along with the path coefficients and other designated correlations), we secure the implied value $r_{UU'} = .341$.

We are now in possession of a commodity that is all too rare in sociological analysis: a precise quantitative "prediction" from an explicit model. The prediction, of course, does not concern some future event in the real world but the result of an inquiry that might be undertaken to ascertain whether the implied relationship is correct. In this case, however, we shall not have to wait long to test the prediction. The OCG data include readings on the educational attainments of both the respondents and their oldest brothers (for the roughly half of the sample having an older brother and able to report his number of years of schooling). The correlation between respondent and *oldest* brother is not exactly what is called for by Fig. 5.8, which treats the two brothers symmetrically. But it is at least worth considering how well the OCG result for brother's education conforms to the outcome deduced from Fig. 5.8. In fact, the OCG data for all native non-Negro men 25–64 years of age indicate a correlation of .573 between respondent's education and education of oldest brother, which is considerably higher than .341, the value implied by the model. Evidently, the model is incorrect, or else the correlation between respondent and oldest brother is materially greater than the correlation between respondent and a randomly chosen brother (both propositions could hold, of course).

If, for the sake of argument, we take the true value of $r_{UU'}$ to be .573, we shall have to modify the original model. Such a modification—or, rather, two alternative modifications, among many possible ones—are shown in Fig. 5.8.

Let us consider first Fig. 5.8a, a version incorporating additional paths labeled *e* (omitting paths *f*). Here were have postulated two mystery variables denoted, respectively, as "Respondent's Id" and "Brother's Id" ("Id" being merely a label for something that behaves in the way to be described). The variable is assumed to have a direct effect, *e*, on educational attainment, and to be independent of intelligence, number of siblings, and socioeconomic background. Moreover, the correlation between Respondent's Id and Brother's Id is taken as .5. Id, therefore, might be a trait determined by a simple genetic mechanism, independently of any genetic determination of intelligence and unaffected by socioeconomic environment. If the reader cares to think of some unconscious motivational factor arising in such a fashion, he may find some help in the imagery. With this purely illustrative postulate, the model is rendered consistent with our information on the correlation between educational attainments of brothers by inserting an appropriate value for the paths, *e*. We have already computed the correlation $r_{UU'}$ produced by paths $a, \ldots, d$ as .341. Since Id is assumed to be uncorrelated with the other background variables, Fig. 5.8a implies that $r_{UU'} = .341 + .5e^2$. Taking $r_{UU'}$ as .573, we can solve for $e = .681$. Moreover, the increment to explained variance in U (or U') amounts to $e^2 = .464$, so that in Fig. 5.8a, the residual, *h*, is reduced to .323, implying that 10 percent of the variance in educational attainment is unexplained in the model. Evidently, Id is quite a powerful variable, as dynamic psychologists have long suspected!

This "fun with numbers" is not advanced as a serious theory of the determinants of educational attainment. The purpose of the exercise is to illustrate one line of argument and the consequences thereof. We are trying, in effect, to imagine the response of a behavioral scientist to Fig. 5.7 as he seeks to muffle his disappointment with the large unexplained residual. Many such scientists react initially by speculating about variables left out of the model. Here, we have seized upon the remark that a behavioral scientist might have uttered in a seminar discussing Fig. 5.7, and have followed it to its logical conclusion. We imagine him contending that the model omits "motivation," and observing that high motivation and low motivation are found in both lower and middle class youth for reasons that are difficult to apprehend. To translate such a remark into some definite implication we have to specify the formal properties of a model embodying the speculative hypothesis being advanced. Other translations than the one just considered could, of course, be entertained; and the consequences would then be somewhat different. What one would like to see in discussions of empirical results, when they take a speculative direction, is an attempt to make the speculations specific enough that their consequences can actually be confronted.

The consequence in Fig. 5.8a may (or may not) strike the reader as far-fetched. "Id," whatever it may be, turns out to have a greater net effect on schooling than any of the other determinants with which we are familiar. If this is just a way of stating that the region of our ignorance exceeds the area of our knowledge, no harm is done. If it is, on the other hand, a programmatic dictum, then we know we have a hard job ahead in seeking to measure and identify a powerful factor whose source and nature are at the moment entirely mysterious.

In Fig. 5.8b, which includes paths labeled *f* while omitting those designated *e*, we consider a slightly different mystery variable to account for the previously unaccounted for correlation between brothers' educational attainments. The mystery variable is now no longer a trait that might be observed in each brother individually but a characteristic of the family or environment which is common to the two brothers. It is designated as a "gray box," whose content will remain unspecified. Here, in contrast to the previous conjecture, we are assuming that the gray box has exactly the same content for both brothers, whereas in Fig. 5.8a the two Id scores were only correlated to the extent of .5. Again, the model permits easy calculation of the unknown path, for $r_{UU'} = .573 = .341 + f^2$, yielding the value of .482 for *f*. In turn, the residual paths, *h*, take on the value .580, implying that 34 percent of the variance of educational attainment for each brother is unexplained.

Figure 5.8b, like its alternate, is a highly specialized modification of Fig. 5.7. The gray box is assumed to be utterly uncorrelated with either intelligence or socioeconomic background. This property immediately rules out such candidates for its content as income, "cultural level" of the home, or even such practices as age at weaning and toilet training (which are thought to vary by social class). To be sure, if the critic can suggest a variable like one of these, and if he is willing to specify not only how it affects schooling but also in what degree it is related to background items, we can entertain still another version of Fig. 5.8 to represent this hypothesis. He will, in any event, have to think of a variable with quite a sizeble influence on schooling, for *f* is the largest path in the diagram; and, if the gray box variable were allowed to be positively correlated with background factors, its correlation with schooling would have to exceed *f*.

5.4 Ability and Achievement: Final Model

The final model developed in the project is represented in Fig. 5.9a and reduced forms thereof are shown in Fig. 5.9b and in Fig. 5.10. The estimates of path coefficients pertain as nearly as possible to the population of United States white men 25–34 years old in 1964. About half the correlations among the

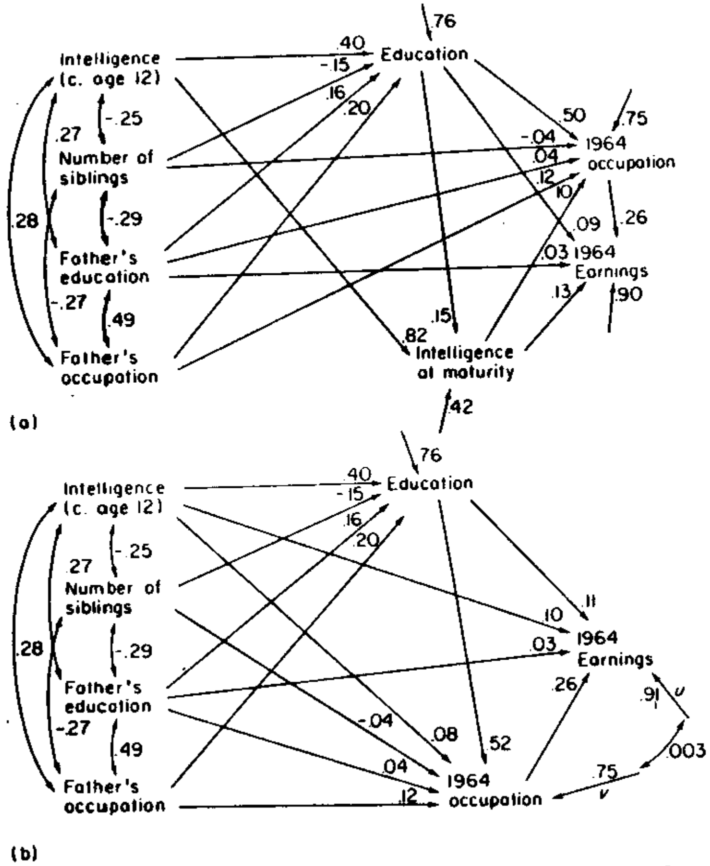

Fig. 5.9. (a) Final model of ability and achievement and (b) semireduced form omitting intelligence at maturity. Source: Duncan (1968a).

eight variables in the model are taken from the CPS-NORC data set. The remainder are either taken from other published sources, such as those discussed earlier, or are derived from the model itself. The details of the estimates, together with some evaluation of them, are given in full in Duncan (1968a).

One important feature of the model is the incorporation of two measures of "intelligence": ability as measured at about age 12 and as measured at maturity. The important work of Bloom (1964) on stability of intellectual traits over the life cycle was consulted in selecting an estimate of .9 as the coefficient of intertemporal stability for intelligence for this segment of the life cycle. The reason why this feature is important is that prior research has left ambiguous the question of the degree to which intelligence measures are contaminated by educational attainment. Thus, in commenting on a Swedish study which

showed that IQ is positively associated with occupational mobility, Lipset & Bendix (1959, p. 234–235) remarked:

Instructive as these data are, they are vitiated in part by the high correlation between I.Q. and educational achievement (.82) and between educational achievement and mobility. Since the intelligence tests were made after the completion of education—in the course of the process of registering for the military draft—and since we know that education itself may result in some improvement of a person's I.Q., the problem of causal imputation is not resolved. Nevertheless, I.Q. tests do measure (even if they do not isolate) native ability, and to this extent [the Swedish] data give clear-cut evidence for the considerable effect of intelligence on social mobility.

In constructing the final model, we explicitly took account of the possibility "that education itself may result in some improvement of a person's I.Q." The

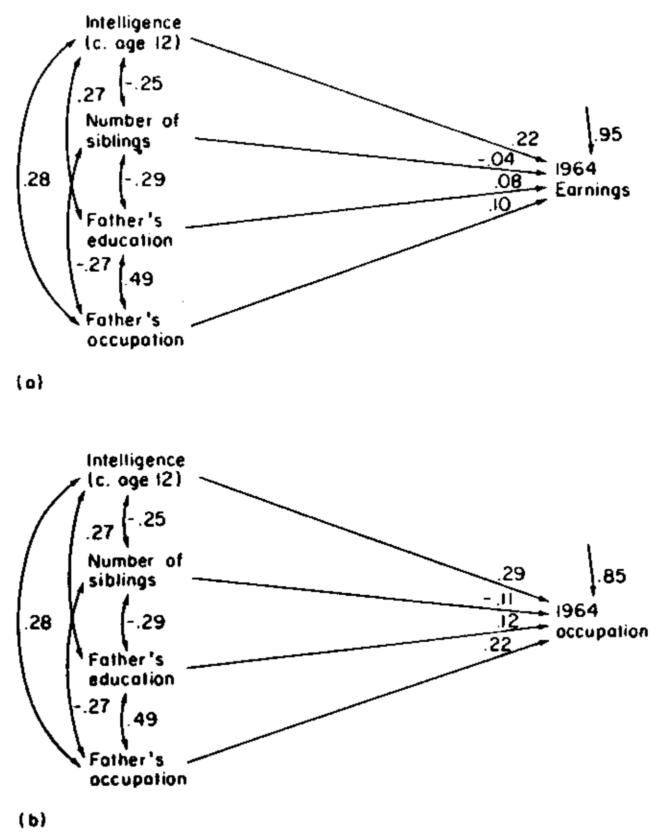

Fig. 5.10. Reduced forms of final model of ability and achievement.

estimate of the magnitude of this effect is a function of (1) the correlation be- tween education and mental ability of mature men; (2) the correlation between mental ability of sixth-grade children and subsequent educational attainment, estimated, as previously described, from the data of Benson (1942); and (3) the assumption as to the stability of mental ability over time, as estimated by Bloom.

Once the model is constructed, we may consider that the data on intelligence at maturity have served their purpose in allowing us to estimate all the co- efficients. We may then proceed to eliminate that variable from the model, deriving the semireduced from shown in Fig. 5.9b. If the original model is correct, we can be sure that the semireduced form does not suffer from the ambiguity to which Lipset and Bendix called attention; the variables can be temporally ordered with fairly little error.*

Perhaps the most interesting substantive result is that the bulk of the in- fluence of intelligence on occupation is indirect, via education. The direct path from intelligence to occupation (Fig. 5.9b) is only .08, whereas the indirect path via education is (.40)(.52) = .21, or more than twice as large. The sum of the two, .08 + .21 = .29, is shown in Fig. 5.10b as the entire effect of intelligence on occupation, apart from joint effects with the other three background variables.

The situation is somewhat different in regard to intelligence as a cause of differential earnings. In Fig. 5.9, it is clear that the effect of intelligence on earnings, net of the effects of education and occupation, is appreciable. Thus men with the same schooling and in the same line of work are differentially rewarded in terms of mental ability. In Fig. 5.9b the direct effect of intelligence on earnings, at .10, is almost as large as the sum of indirect effects via education and occupation, which comes to (.40)(.11)+(.40)(.52)(.26)+(.08)(.26) = .12. The combination of direct plus indirect influence, .10 + .12 = .22, is shown in the upper diagram of Fig. 5.10 as the entire effect of intelligence on earnings, net of the other three background factors.

5.5 Ability and Achievement: A Replication

As was indicated in the previous section, estimates of coefficients in the "final" model of ability and achievement were derived from data for the population of United States white men aged 25–34 in 1964; but a considerable part of the information used in making the estimates pertains to other popula-

* Note that estimates for Fig. 5.9b are computed according to the discussion of semireduced form models with correlated errors (cf. Chapter 2, formulas [2], [6], [7], and [8]). In the case of Fig. 5.9b, eliminating "intelligence at maturity" as an intervening variable in the semi- reduced form results in a small correlation between the residuals to 1964 occupation and earnings (r_{uv} = .003).

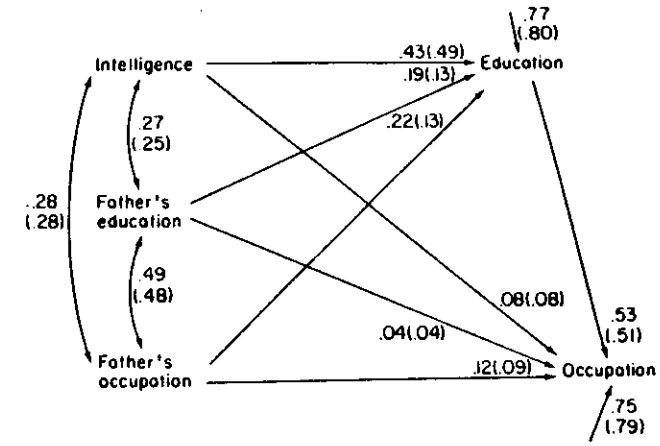

Fig. 5.11. Abridged version of the final model of ability and achievement with path coef- ficients estimated for two populations. (Figures not in parentheses are based on data sum- marized in Section 5.4; figures in parentheses are based on DAS data set.)

tions. Access to the DAS data set permits a completely independent replication, although one that does not contain quite enough information to estimate the complete model and one for which the issue of temporal stability of measured intelligence is left unresolved. Fig. 5.11 presents a comparison between estimates secured from the same data used in the previous section and those secured from the DAS data. In the DAS data, the measure of "intelligence" is the "Similarities" subscale of the Wechsler Adult Intelligence Scale. As we have used this scale here, it is interpreted to refer to intelligence as it would have been manifested at some point well before the termination of schooling. That this is a distortion of the probable facts is the main message of the manipulations involved in the work reported in Section 5.4. Yet the results of that work also suggest that the distortion is comparatively minor.

Indeed, the main difference between the two sets of path coefficients in Fig. 5.11 is that the path from intelligence to education is rather larger in the DAS data than in the set of estimates derived from CPS-NORC data and other sources. An exaggeration of this path is precisely what we would expect if a measure of adult intelligence is used as a proxy for childhood intelligence. Apart from this difference, and the corollary reduction in the paths from socioeconomic background factors, the two sets of results exhibit a very nice replication indeed. It appears that the final model of ability and achievement describes features of the process of achievement that are pervasive in American society.

5.6 Summary

Although it is not commonly defined in that way, there is a good argument for conceiving of intelligence as "ability to perform occupational roles." That the pioneers in the measurement of intelligence implicitly proceeded on some such notion is suggested by the very high correlation between the "intelligence demands" of various occupations, as estimated in the Barr scale, and the "prestige" of those occupations as reflected in ratings by the general public. Despite this high correlation, it remains empirically contingent whether particular individuals will find their way into occupations of varying status to a greater or lesser degree on the basis of the kinds of abilities reflected in measurements of intelligence.

Some significant bodies of published data provide estimates of the correlation of parental occupational status with measured intelligence, the correlation of intelligence measured in childhood with subsequent educational attainment, and the correlation of mental test scores of young men with occupations held at an early stage of the career. With appropriate caution, such estimates may be juxtaposed with other data available to the project in the construction of models explicating the role of ability in achievement.

Preliminary versions of this kind of model make it clear that, while intelligence has a substantial influence on amount of schooling, it does not fully explain the correlation of family-background factors with schooling. Moreover, a very substantial correlation between the levels of schooling of brothers can only partly be explained by common family-background factors and sibling resemblance in intelligence.

The work leading to a "final model" of ability and achievement takes explicit account of the possibility that intelligence measured at maturity may be partly a result of amount of schooling. What appears to be an appropriate adjustment for this effect, however, leaves intact the proposition that intelligence has a substantial influence on occupational achievement, apart from its correlation with family-background factors. Much of that influence, however, is mediated by educational attainment. Somewhat similarly, intelligence bears an important relationship to income (net of social origins) for men of equal schooling and occupational statuses. Unlike its relationship to occupational achievement, intelligence affects income directly to nearly the same extent as it influences differential earnings indirectly through schooling and occupational attainment. While the inclusion of intelligence test scores in a model of the process of achievement, therefore, increases appreciably the proportion of variation in occupational status "explained," there remains a very substantial amount of variation still "unexplained."

In Chapter 4 racial differentials in socioeconomic achievements were examined and found to contain a component for which differences in socio-

economic circumstances and structure of the family of origin, and prior attainments, could not account. This residual was interpreted as a handicap, or cost, to some men for being black. Recent commentary on the comparative achievement of whites and Negroes in the United States has suggested that the residual might be explained (in part) as an outcome of racial differences in the distributions of intelligence (Jensen, 1969). We do not wish to comment here on biogenetic, sociogenetic, and other mechanisms by which the documented difference in black–white IQ may have arisen. However, the separation in means by about one standard deviation on a normal IQ scale (about 20 points) does *not* nearly account for the occupational and economic handicaps of Negroes. As reported elsewhere (Duncan, 1968b), little racial variation in schooling remains after equating men for family of origin characteristics and intelligence. Nevertheless, substantial occupational and economic disadvantages accrue to black men, despite their parity with whites in social origins and education, and for constant IQ scores. These supplementary comments are offered to place in perspective the magnitude of the effects of intelligence on achievement and to summarize results from explorations with a more elaborated model not included in this volume.