

This article was downloaded by: [University of Arizona]

On: 11 November 2012, At: 19:53

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954

Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Oxford Review of Education

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/core20>

Spearman's g and the Problem of Educational Equality

Arthur R. Jensen

Version of record first published: 03 Aug 2006.

To cite this article: Arthur R. Jensen (1991): Spearman's g and the Problem of Educational Equality, *Oxford Review of Education*, 17:2, 169-187

To link to this article: <http://dx.doi.org/10.1080/0305498910170205>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Spearman's *g* and the Problem of Educational Equality

ARTHUR R. JENSEN

In my essay in this journal in 1975, I quoted the British geneticist, J.B.S. Haldane:

The test of the devotion ... to science will, I think, come when the accumulation of the results of human genetics, demonstrating what I believe to be the fact of innate human inequality, becomes important. ... I believe that any satisfactory political and economic system must be based on the recognition of human inequality.

Today many scientists, like Haldane, are concerned with problems of human inequality. These scientists are neither elitists nor racists, nor are they promoting a privileged class or forging an underclass. In fact, many are troubled by the unfortunate consequences of human inequality. Accumulated evidence, however, forces them to accept the reality of human inequality in certain educationally, economically, and socially important characteristics and to recommend that it be openly recognised by the leaders of society and the general public. Human variation we know reflects the interaction of multiple causes, some originating in the course of human evolution, and others in both the ontogenetic development and life experience of individuals. Also, part of the variance, or individual differences, in certain behavioural traits, is a biological phenomenon, a fact of nature. Human variation should be grasped and dealt with constructively, by bringing to bear on it as much knowledge, humaneness, generosity, reason, wisdom and farsightedness as society can possibly muster. This, I believe, is Haldane's message.

In the latter half of this century, however, this message has been shunned and scorned in the public arena, for largely historical and political reasons. The public media have portrayed the science that deals with the interaction of environmental and genetic causes of individual differences in mental ability—the so-called IQ controversy—as an argument between the good guys and the bad guys. Among the more informed, however, this portrayal is increasingly seen as fatuous and passé. In fact, the old canards of the IQ controversy are now rejected by a majority of the well informed (e.g. those surveyed by Snyderman & Rothman, 1988). On this score, unfortunately, the education establishment is probably the most laggard, with the possible exception of the mass media. This is unfortunate because the one institution for which a realistic concern with human differences is most obviously required is, of course, public education.

In America, the past 20 years have witnessed official denial of individual and especially racial group differences in traits that are highly germane to the basic aims of education. This naive egalitarianism has spawned three main consequences: (1) increasing expression of national concern over the deterioration of public education;

(2) a perception of growing disparity in educational achievement between different segments of the population; and (3) official sanctions not only for equality of educational *opportunity*, but for equality of educational *outcome* as well.

Where equality of educational opportunity has become a reality, the visible manifestations of inequality in educational outcome are the symptoms of more basic sources of variance in educational performance. Before going into the particulars of some of the most salient and troubling symptoms, I should explain what research preponderantly has indicated is their root cause—namely, Spearman's *g*.

SPEARMAN'S *g*: ITS NATURE AND CORRELATES

Charles Spearman, arguably Britain's greatest psychologist, discovered *g* almost 90 years ago. The italicised lower-case letter *g* is the symbol he adopted for the *general factor* that exists in any and every large collection of diverse mental tests. Any test, task, or pursuit that makes demands on mental effort reflects some source of individual differences (variance) in common with every other such indicator of mental ability. This source of individual differences, or variance, that is common to all cognitive activity is known as Spearman's *g* (or psychometric *g*, or just *g*, for short). In other words, the *g* factor reflects whatever it is—presumably some attribute of the brain—that causes *individual differences* in performance on any indicator of mental ability to correlate positively with performance on every other such indicator, regardless of its specific information content, sensory modality, or form of response.

The degree to which various tests are correlated with *g* (or are *g*-loaded, in psychometric parlance) can be determined by mathematical methods known as *factor analysis*. The results show that tests differ markedly in their *g* loadings. These differences in *g* loading, however, are not predictable from the superficial features of ability tests, such as verbal, nonverbal, or performance tests, speeded or non-speeded, individual or group administered. Short of performing a proper factor analysis, the best clues to a test's *g* loading are the number, variety, and complexity of the test items' cognitive demands. Of course, the test items must fall within a psychometrically suitable range of difficulty for the persons tested.

Conventional IQ tests are always highly *g*-loaded, but IQ is not perfectly correlated with *g* and therefore cannot be exactly equated with *g*. IQ tests usually measure some admixture of other factors in addition to *g*, such as verbal, numerical, spatial visualisation, and memory abilities. Some part of the IQ variance may also reflect differences in certain classes of acquired information, often of a rather scholastic nature. These sources of variance may slightly improve the practical validity of the IQ for certain purposes, but they dilute the assessment of *g*. The specific content factors in certain IQ tests, independently of *g*, may reflect a small part of the variance in *past* scholastic achievement, but IQ scores with the *g* factor statistically removed have virtually no predictive validity for children's *future* scholastic performance. Clearly, the *g* factor is the chief active ingredient in IQ tests' predictive validity. The predictive validity of *g* applies not only to academic performance at all levels of education, but also to performance in nearly all types of jobs. Because occupations differ in their *g* demands as do mental tests, highly *g*-loaded indicators, such as IQ, are better able to predict probability of success in various occupations as they rank higher on the scale of complexity of their cognitive demands.

Hence the *g* factor, which cannot be adequately or accurately described in terms of any specific test or particular type of performance, can best be thought of as *general*

mental ability. This term, however, scarcely suggests all that we now know about the important properties of *g*. During the past decade there has been renewed interest in theory and empirical research on the nature of *g*, not only from a psychometric standpoint, but also in terms of basic information processes and their biological underpinnings. I can only mention briefly some of the main findings and conclusions of this research.

Non-intellectual Correlates of g

Although *g* is identified by factor analysing conventional psychometric tests, it also has connections with many variables that lie completely outside that realm. Therefore, it would be a mistake to think of *g* as merely some kind of psychometric artifact or hypothetical construct without meaning or reality beyond the scores obtained from typical mental tests or the mathematical procedures of factor analysis used to determine their *g* loadings (Jensen, 1986; 1987a).

For example, there is a direct relationship between the magnitude of tests' *g* loadings and the following non-psychometric variables: the heritability coefficients (i.e. proportion of genetic variance) of the tests, the degree to which inbreeding depression lowers test scores, the degree to which test scores are raised by heterosis (outbreeding), and features of the average evoked potential, that is, the latency and amplitude of the brain's electrochemical response to an external stimulus (all these variables reviewed by Jensen, 1987a).

Other variables found to be correlated with highly *g* loaded tests are the rate of glucose metabolism by the brain while the subject is engaged in mental activity (Haier *et al.*, 1988), the speed of neural and synaptic transmission in the brain (Reed & Jensen, in press), and a number of anatomical variables such as height, head size or cranial capacity and brain size, myopia, and other physical variables (reviewed in Jensen & Sinha, 1991). Although the causal pathways between these physical variables and *g* are not yet understood, the correlations are entirely real, and indicate that the population variance on standard mental tests reflects latent traits, primarily *g*, that are profoundly enmeshed with organismic variables. It is a reasonable hypothesis that most of these correlations between physical and behavioral traits came about through natural selection in the course of human evolution.

Also, there are a number of behavioral measures which show substantial correlations with psychometric *g*, although these measures are derived from tasks that are not ordinarily thought of as having intellectual content. Individual differences in performances on these kinds of tasks reflect information *processes* rather than any kind of acquired knowledge. These tasks are known in experimental cognitive psychology as elementary cognitive tasks (ECTs). They usually involve no verbal or symbolic content, but consist simply of the onset of a light or a sound, to which the subject responds by either releasing or pressing a button. Most ECTs are so simple that everyone can do them easily, and the only important source of variance is the subject's speed or reaction time (RT). In the various ECTs used in my laboratory, for example, the average RTs are less than 1 second. Yet these RT measures, especially when combined from several different ECTs, are quite substantially correlated with *g* (Jensen, 1987b; Vernon, 1987). Another class of ECTs that are correlated with *g* (about 0.50) is known as inspection time (IT). It is measured as the duration of a visual or auditory stimulus required to make a simple discrimination, such as which of

two lines is the longer, when one line is double the length of the other (Kranzler & Jensen, 1989). IT is typically less than 0.1 second.

All the research on these RT and IT tasks leads to the conclusion that psychometric *g* as measured by non-speeded IQ tests, is importantly related to individual differences in the speed of information processing. Seemingly small but reliable differences between individuals in their speed of information processing, when interacting with environmental experiences over a number of years, can result in very large differences in acquired knowledge and skills, as shown by comparisons of RT and scholastic knowledge and skills in high-IQ and average-IQ children (Cohn *et al.*, 1985; Jensen *et al.*, 1989). Also, the *g* of psychometric tests is one and the same general factor found in a wide variety of indices of individual differences in cognitive learning ability (Jensen, 1988a).

Spearman's Hypothesis

This particular hypothesis is especially pertinent to the problem of educational equality in America, because the most conspicuous feature of the problem is the large and seemingly intractable difference in the average level of educational performance between the black [1] population, on the one hand, and the white [2] and Asian [3] populations, on the other. Spearman (1927) was the first to note that the size of the black-white difference on a number of mental tests varied considerably from one test to another, and to conjecture that the size of the black-white differences on various tests is directly related to the tests' *g* loadings. In other words, he hypothesised that the black-white difference is essentially due to a difference in *g* rather than to differences in specific characteristics or in the information content of particular tests. This phenomenon, which I have termed Spearman's hypothesis, has now been tested extensively in at least 13 independent samples comprising 41,000 black and white subjects and 86 different mental tests. Every study, without exception, substantiates Spearman's hypothesis, which thus is fully borne out as an empirical fact (Jensen, 1985a, 1985b, 1987c). In the American population, the black-white difference on IQ tests, which are highly correlated with *g*, has remained virtually constant, at about 1.2 standard deviations (SD), at least since large-scale mental test results were first available, some 70 years ago.

Because of the central role of *g* in individual differences in scholastic aptitude and achievement, Spearman's hypothesis must be regarded as crucial in any scientific explanation of the black-white disparity in performance at all levels of education and, in fact, in all types of behaviour to the extent that they are correlated with *g*. The established relationship of *g* to educability in general, and the fact that the black-white difference on mental tests and their educational correlates is basically a difference in *g*, must be considered as the foremost fact in understanding the most visible problems of education.

Because educational performance is highly but not perfectly correlated with *g*, black-white differences in scholastic achievement tend to be slightly smaller than the *g* difference. The fact that the racial gap in scholastic performance is generally *smaller* than the *g* difference (which is about 1.2 SD) suggests that other, noncognitive factors, such as motivation or effort, do not systematically increase the scholastic achievement differences over that predicted from the *g* difference alone. No explanatory variables besides *g* are necessary.

Understanding the black-white difference in educational outcomes is tantamount to

understanding the causal nature of g . Sociological hypotheses of blacks' scholastic shortfall that invoke such notions as white racism, victimisation, discrimination, the culture of poverty, caste, linguistic differences, different learning styles, peer culture, poor self-esteem, alienation, and the like, simply prove unnecessary. This is not to say that these things do not exist; but there is no evidence that they contribute to explaining the black deficit in educational achievement, which can be fully accounted for in terms of Spearman's hypothesis and the fact that g is the main source of variance in educational achievement. There is nothing in the sociological hypotheses mentioned above that would predict the relative size of the black-white difference on various mental tests, or in various scholastic subjects, as well as the tests' g loadings.

The same point can be made about predicting the proportional representation of blacks in various occupations. Ranking occupations according to their g demands is a better predictor of the black/white ratios in those occupations than the sociological factors claimed to account for the increasing racial disparity at the more highly g -demanding levels of the occupational hierarchy (Gottfredson, 1986, 1988).

THE SHIFT FROM EQUALITY OF OPPORTUNITY TO EQUALITY OF OUTCOME

In the late 1960s and early 1970s, it became apparent that schools which provided both whites and blacks with equal educational facilities and equally qualified teachers produced far from equal levels of achievement. The main source of variance was found to be the one individual pupil characteristic over which schools can have little or no control, namely, g .

This fact is most evident in those school systems that have a record of being especially dedicated to equality of educational opportunity. A good example is the school system of Berkeley, California, with which I happen to be most familiar. Since 1968, every conceivable measure was taken to ensure equality of opportunity in every school, including the busing of pupils from different neighborhoods to achieve complete racial integration of all classrooms. The school superintendent said, "You name it, we've tried it." Yet in scholastic achievement scores the black-white differences, averaging at least one standard deviation, have persisted for more than 20 years since these conditions of equal opportunity were instituted. The case of this school system, and many others like it, demonstrates that equality of educational opportunity does not result in equality of educational outcome. Hence it would be hard to see how the educational system itself, in such cases, could be held responsible for the inequality of educational outcome. Unequal outcomes, of course, are no surprise with respect to individual differences *within* any particular racial or ethnic group. But many were surprised and dismayed to see the persistence, year after year, of a near-perfect correlation between the *mean* IQs of each of the schools' various racial/ethnic groups and their *mean* levels of achievement, with the mean differences between racial/ethnic groups ranging more than one standard deviation. Factor analysis of the test results showed clearly that the *between*-groups difference consisted of the same sources of variance as the *within*-group differences, with one and the same g factor accounting for most of the variance both *between* and *within* racial groups (Jensen, 1980). Racial disparities in scholastic performance do not appear to be due to a distinct set of causes from those that determine individual differences *within* racial groups; the same g factor is implied in both cases.

This fact and its obvious consequences have resulted, from the early 1970s to the

present, in a radical shift in educational goals—from equality of *opportunity* to equality of *outcome*. Indeed, if equality of outcome is deemed the only acceptable result of equality of opportunity, it might be said that the reality of *g* differences literally *forces* a shift of emphasis from equality of opportunity to equality of outcome. In this shift that has taken place, two distinct phases are discernable: (1) efforts made to genuinely reduce inequality of outcome, and (2) measures taken to create the illusion of equal outcome.

EFFORTS AT GENUINELY REDUCING INEQUALITY OF OUTCOME

A number of interventions were implemented in the hope that they would actually reduce inequalities in outcome by improving the basic abilities, or skills as they are usually called, involved in school learning. The theory on which many of these efforts were based takes no account of *g* differences between the targeted groups and the rest of the school population. Rather, it was based on what has been termed the “specificity doctrine” (Jensen, 1984), which holds that intelligence is *learned* and consists of a myriad of particular items of information and a repertoire of specific cognitive skills that can be taught. Economic, social, and cultural disadvantage was hypothesised as the chief hindrance to children’s acquisition of educationally crucial kinds of knowledge and skill. Lagging scholastic progress, typical of many children from an educationally disadvantaged background, was attributed to deficiencies in the specific knowledge and skills assumed to be prerequisites for learning various scholastic subjects. According to the specificity doctrine, the essential difference between fast and slow learners is a difference in the amount of previously acquired bits of knowledge and skills they possess that are deemed essential for coping with specific scholastic tasks. Interventions such as special pre-school and compensatory education programmes, therefore, sought to inculcate in disadvantaged children the kinds of prerequisite knowledge, skills, and attitudes regarding education presumably possessed by pupils who thrive academically.

Unfortunately, both short-term pre-school interventions and compensatory programmes designed for school-age children with a poor prognosis for success in school had surprisingly little effect on IQ or achievement level. What little effect there was on these specifically targeted variables dwindled to naught within a year or so after the treatment groups entered regular classes (see review by Clarke & Clarke, 1989).

It is especially instructive to examine the outcome of the most extensive and intensive pre-school enrichment programme ever undertaken—the famous Milwaukee Project, which was glowingly publicised by the mass media while in progress but has been conspicuously shunned by the media since the publication of the study’s final results (Garber, 1987).

The Milwaukee Project was a multimillion dollar long-term experiment in which scholastically at risk black children from the inner city of Milwaukee were given an unprecedented degree of cognitive stimulation. The director of the programme, the late Rick Heber, even proclaimed that the early environment of such famous geniuses as Sir Francis Galton and John Stuart Mill (who were also famous child prodigies) would appear grossly deprived compared with the intellectually super-stimulating environment created for the infants in the study’s experimental group. Beginning shortly after birth and continuing until they were 6 years of age, 17 infants were taken from their homes to an infant stimulation centre. There the infants spent the whole day, 5 days a week, being mentally stimulated and trained in order to inculcate them

with the specific knowledge and cognitive skills believed to constitute intelligence. Virtually every environmental factor in the child psychology literature thought to influence the development of intelligence was applied to these children throughout their 6 years in the programme.

At the beginning of their first year in regular public schools, the six year-old first-graders in the experimental group averaged about 30 IQ points higher than a control group selected by the same criteria. By the end of first grade, however, the experimental and control groups did not differ significantly in scholastic achievement, and by the fourth grade (9 to 10 years of age) both groups were at approximately the 10th percentile in achievement scores—a level of academic performance commensurate with an IQ of 80, which was approximately the control group's mean IQ. Normally, IQ (as an approximate index of g) is highly correlated with scholastic performance, and if the intensive training of cognitive skills had actually raised g rather than just taught the experimental group the knowledge and skill content of the IQ test (i.e. its specificity), the experimental and control groups would have differed widely in scholastic performance. The fact that they did not indicates that the specificity of the IQ test had been trained up, while g remained unaffected. Hence the IQ was hollow with respect to g and hence non-predictive for the treated group, but was g -loaded and normally predictive for the control group. (For further details of this analysis and its interpretation, see Jensen, 1989b.) The evidence consistently shows that, for any particular method of teaching, individual differences in the amount of transfer of training from one scholastic skill to other, related or more advanced skills, is purely a function of g . Almost any child can be trained to perform any particular skill, given sufficient time. But the transfer value, or generalisability, of that skill to other tasks or problems depends largely on g .

Scores on specific tests can often be dramatically raised by teaching to the test (often done intentionally), while not affecting the level of g . Yet it is g , of course, that accounts for the tests' practical predictive validity for a wide variety of criteria. The Milwaukee Project and many other intensive efforts to raise significantly not just IQ, but g itself, have so consistently failed to achieve this aim (see Spitz, 1986) that, until good evidence to the contrary is forthcoming, we must conclude that g is not susceptible to manipulation by strictly psychological or educational techniques. This conclusion is consistent with the theory that g is essentially a function of biological factors, both genetic and constitutional, and if it is to be appreciably altered, it would have to be altered by biological means. Hence good or bad physical health (both pre- and post-natally), presence or absence of perinatal trauma and childhood diseases, and poor or optimum nutrition throughout the developmental years probably contribute more of the environmental variance in g than do any psychological or educational factors, assuming a normally humane social environment.

Since it appears unlikely that variance in g itself can be significantly reduced by behavioral means, a number of psychological-educational methods have been tried that take the reality of g differences for granted, but then aim to get around the educational and employment problems associated with the wide range of individual differences in g . The three main classes of these techniques are known as (1) *Aptitude X Treatment Interaction (ATI)*; (2) *Mastery Learning*; and (3) *Thinking Skills Training*.

(1) *ATI* consists of discovering the distinct optimum type of instruction for each level of ability. Evidence on *ATI* indicates modest success in improving the scholastic achievement of below-average pupils by means of *direct instruction* in the specific content and skills of the curriculum. Highly structured lessons are designed to

minimise the burden of complex information processing for the learner, thereby lessening the importance of g . The effects of ATI are generally not dramatic, but ATI seems preferable to subjecting pupils at every level of ability to the same instruction, which typically results in high failure rates and discouragement among less able pupils.

(2) *Mastery Learning* requires pupils to attain a high level of mastery (say, 90%) of the knowledge or skill content of a given lesson before being permitted to go on to the next lesson in a specially well-planned sequence. Because time-to-learn is highly correlated with g , and time for study is limited, there is necessarily a trade-off between degree of mastery and amount learned. There is no way that mastery learning can get around this trade-off. The result is that low- g pupils master a relatively small amount of scholastic material in the same amount of time that high- g pupils master much more. For slow learners, this method more or less insures that they will master more than they would if mastery was not a prerequisite for beginning the next lesson. This is advantageous for slow learners because their educational attainment extends to a somewhat higher level than would be likely under non-mastery instruction. Mastery also circumvents a cumulative achievement deficit, which occurs when deficits in early stages of learning hinder learning at more advanced levels, thus resulting in a *cumulative deficit* as the pupil is confronted by increasingly advanced material. The final result is a lower asymptote of achievement than would be the case under mastery learning. But the range of individual differences in g is so great in the typical school that, under mastery learning, some high school pupils will still be struggling to master fractions and decimals, while others will be breezing through calculus and Boolean algebra.

(3) *Training Thinking Skills* consists of teaching pupils fairly general strategies for problem solving, critical thinking, analysis and interpretation of written material, self-guidance of learning, and monitoring one's own progress. Though such training has positive effects on educational achievement, the chief beneficiaries have been the abler pupils. This should not be too surprising, because the efficiency of acquiring and applying these rather complex thinking tools is strongly related to g , as is their transfer to a wide range of intellectual pursuits.

In summary, each of these methods has certain educationally desirable effects, yet the evidence leaves little doubt that they do not overcome the problem of the wide range of g in schools or the work force. Consequently, faced with the considerable g demands of a technological society, some sizable proportion of our population, perhaps even more than 20%, unfortunately suffers severe limitations in academic educability, economic productivity, or even self-sufficiency. Psychological and educational research has not yet discovered, and may never discover, any means that would trivialise this disadvantage. Certainly in the foreseeable future we will have to face the reality of the personal and social problems associated with large differences in g .

THE PRESENT CRISIS IN EDUCATION

Some 30 years of federally funded psycho-educational research on how to raise the scholastic performance of pupils in the educationally at risk segment of our population and implementation of innumerable large-scale programmes have not delivered what the public had been led to expect. Official pronouncements have, in fact, warned of worsening educational outcomes, not just in terms of the groups targeted for special concern, but for the overall level of academic performance in the nation's schools. The National Commission of Excellence in Education published its now famous report, *A*

Nation at Risk (1983), warning that American students were doing poorly compared to their counterparts in other developed countries with considerably smaller per capita educational budgets. The Commission also warned that this fact augurs dire consequences for America's future position in the world economy. It has been noted that America ranks surprisingly below many other developed countries in literacy rate and that America's black and Hispanic [4] populations may be on a par in this respect with many Third World countries. Some observers have suggested that the nation is increasingly becoming socially and economically divided between the educational haves and have nots in terms of educational outcome. They fear that, in literacy, general educational level, and ability to assume productive roles in a technological society, a troubling segment of the population is becoming a kind of Third World within the USA. This is most conspicuous in the inner cities of our largest urban centres.

More recently, the National Research Council issued a detailed report (Jaynes & Williams, 1989) on the current status of blacks in the USA, which includes a chapter loaded with tables and graphs of statistics on the educational scene. Here are some of its conclusions:

... there remain persistent and large gaps in the schooling quality and achievement outcomes of education for blacks and whites. Black high school dropout rates remain higher than those for whites, black performance on tests of achievement lags behind that of whites, and blacks remain less likely to attend college and to complete a college degree. After the mid-1970s, the college-going chances of black high school graduates have declined, and the proportion of advanced degrees awarded to blacks has decreased... blacks's status in higher education, as undergraduates, graduates, and faculty, has worsened or stalled since the mid-1970s. Several indicators, in particular college attendance rates, show signs of slippage, with blacks' status deteriorating relative to that of whites and of other minorities. (pp. 377-378)

Although the NRC's report notes that "substantial progress has been made toward the provision of educational resources to blacks" (p. 378), there are "persistent and large gaps in schooling quality" (p. 379). This is most evident in predominantly black inner-city schools, where a kind of critical mass phenomenon is often observed. That is, when the proportion of pupils with low scholastic aptitude exceeds some critical threshold, teacher morale, classroom discipline, and the quality of instruction all markedly deteriorate. Of course, under these conditions, the few pupils with above-average academic potential are inevitably short-changed. The effects of the critical mass phenomenon are compounded by the intrusion of the social pathology of the surrounding neighborhood. The whole problem is extremely complex and has multiple causes, of course. But certain major features of this distressing condition seem clearly related to *g*. Concerning educators' attempts to improve these conditions in inner-city schools, for example, Maeroff (1988) observes

... there is a tendency to revel in delusions of improvement. Order may be restored, but oppression reigns. Test scores may rise, but concepts remain ungrasped. Facts may be memorised, but students cannot apply them in solving problems. Dropouts may be kept in school, but the diplomas they receive are not backed by skills and knowledge. (p. 638)

THE MEAN AND THE VARIANCE OF EDUCATIONAL ACHIEVEMENT

Studies of instructional methods that actually improve the level of achievement, provided they do not impose artificial restraints on the progress of learning by the higher-*g* students, have found it impossible to raise the overall average level of achievement and also to reduce its variance. So consistently has been found that it could almost be called The First Law of Individual Differences, to wit: In achievements that do not have a low performance ceiling, instruction that succeeds in raising the group mean also increases the variance among individuals. As differences between groups simply reflect the average of the individuals that compose them, they of course show the same phenomenon. Hence educational conditions that maximise achievement will magnify both individual differences and group differences. Although this militates against approaching *equality* of educational achievement, either for individuals or for different groups, improving the *mean* level of achievement would have the decidedly beneficial effect of raising a larger proportion of every group above the critical level of knowledge and skills required for productive employment and economic self-sufficiency. Let it not be overlooked that this is an exceedingly desirable outcome for any society whose welfare depends largely on the educated intelligence of its population.

Nevertheless, the overriding problem of education in America is still viewed as the necessity of equalising the educational outcomes of different racial groups. That seems the primary goal. Raising the overall level of educational achievement in all segments of the population presently appears to be of secondary concern to educators and politicians. With the primary goal still seemingly far out of reach after a generation of educational innovations to achieve it, there are increasing signs of frustration and despair.

EDUCATIONAL FADS AND ILLUSIONS AS SYMPTOMS OF FRUSTRATION

A great deal of so-called educational research is a parade of fads, each new one as ephemeral as previous ones. These consist of variations of innovative classroom procedures, new seating arrangements, instructional methods involving different pupil-teacher and pupil-pupil interactions, and, in recent years an increased emphasis on making instruction more interesting in the sense of entertainment. The intellectual demands of education are eclipsed by programmes to socialise pupils, enhance their self-esteem, heighten their aspirations, unleash their self-expression, and generally make them feel good about themselves. However laudable these aims are in their own right, evidence is lacking that they have any significant effect on what is ordinarily thought of as scholastic achievement. In general, there is a dearth of good research on the theoretical basis or practical efficacy of most of these innovations; in fact, many do not even last long enough to allow replicated studies. But if the sum total effect of two decades of educational innovations were actually superior to the more traditional or conventional classrooms, the achievement gap between the more and the less advantaged segments of the school population should have decreased over the past decade, because the most innovative programmes have predominated in the schools that serve those neighbourhoods with the highest percentage of pupils at risk for poor scholastic achievement. One sees these innovative programmes extolled in newspapers, popular magazines, and TV programmes. Everyone is exuberant about their promise, but rarely is there a *bona fide* evaluation of their outcomes. If they really work, why do they not persist and why do they not become more widespread? Insufficient funding is not a

convincing explanation, at least in the USA, which has the world's highest per capita expenditure for education.

Charismatic Teachers

As examples of what is possible in raising the achievement level of at risk pupils, enormous publicity has been given to a handful of what can be called charismatic teachers. The implication is that if only the majority of teachers were as dedicated as these charismatic teachers and acquired their successful techniques, the problem of group differences in educational outcomes would virtually disappear. In the few instances that the claims for extraordinary gains in achievement have been looked into, they typically turn out to be either grossly exaggerated or completely non-existent. Or, when the claims are actually valid, they can usually be explained in terms of special selection of only the ablest students (see Spitz, 1986, pp. 81–82). It is noteworthy that the purported outcomes are usually one-shot successes that can seldom be repeated in subsequent classes even by the same charismatic teacher whose original feat had been lionised by the mass media. But what is most remarkable about these anecdotal reports of charismatic teachers is how unquestioningly they are received by the public and even by professional educators. Educational researchers have been curiously derelict in applying their observational techniques and edumetric methodology to bear on verifying the charismatic teacher's effect on pupils' achievement and discovering the factors responsible for it.

Abolition of Standardised Tests

In the State of California and in many other school systems in the USA school psychologists are now legally prohibited from using standardised tests of intelligence and academic aptitude with minority black and Hispanic children who are referred for scholastic problems. The official claim is that the tests are biased for these minorities, despite overwhelming evidence that the tests have the same practical validity for minorities as for the majority. Standardised achievement tests are still in general use, but there are organisations that strive to see them abolished also.

With the banning of IQ tests for minority pupils, school psychologists have substituted diagnostic procedures that claim to assess something importantly different from IQ—learning potential, or LP. Its proponents claim it is superior to conventional IQ tests for assessing the capabilities of children from poor or culturally different backgrounds. Essentially, the child is tested, then is given brief coaching and practice on the test material, then is retested; the gain score (i.e. the difference between the pre- and post-test scores) indicates the child's LP. Connected with some of these instruments are prescribed intervention techniques for developing the cognitive skills of slow learners. LP measures and their associated interventions are now the vogue in schools systems that have banned the traditional mental tests. Research on LP, however, shows that when children are grouped by LP scores, 80 to 90% would have been placed in the same groups on the basis of ordinary IQ tests. Both kinds of measures predominantly reflect one and the same g factor, but LP, being based on a difference score, has lower reliability than the IQ. A comprehensive review of the research on LP concludes that "LP assessments provide little to no diagnostic information beyond that supplied by traditional IQ tests" (Glutting & McDermott, 1990, p. 306). When this fact becomes more widely known, LP also will probably be

banned along with IQ tests. What cannot be banned, of course, is the connection between *g* and educational attainment, whether or not either variable is formally measured.

FORCED APPEARANCE OF EQUAL OUTCOMES

Massive evidence has established the fact that scores on highly *g*-loaded tests are correlated with performance in school, college, armed forces training programmes, and in many jobs to the extent that they are intellectually demanding. Such tests therefore have become widely used in educational and personnel selection. Without question, the tests have practical utility to educational institutions and employers, in terms of efficiency and costs, especially when there is a fairly high ratio of the number of applicants to the number of openings. Rank ordering all applicants' test scores and top-down selection of the number of applicants needed to fill the openings ensure a more capable student body or work force than any other means of selection.

In a racially diverse population, a problem that arises from such top-down selection is that there are rather large group differences in the means and standard deviations of the distribution of scores on the selection test. This creates the condition now known as *adverse impact*, which means that use of the test, with top-down selection, results in disproportionate selection of various groups compared to their proportions in the pool of applicants or in the general population. This phenomenon, which occurs whenever any tests are used with top-down selection, has given rise to a variety of protests and suggested solutions.

Biased Tests

This is still the popular claim, but it is rapidly disappearing in informed circles. Researchers have found that tests do not under-predict the educational achievement or job performance of minorities. Tests measure *individual differences*, and are entirely colour blind in doing so. Their predictive validity is the same for blacks and whites alike; the average difference in their tests scores is reflected in their average performance difference in the predicted criterion to the extent that the test scores are correlated with the criterion *within* each group. But the *claim* of test bias dies hard. When it was found, for example, that the average failure rate of blacks who take the National Teachers Examination was about four times that of whites (67% vs 16%), test bias was alleged as the cause. Although a panel of experts could not find evidence of bias, it was still asserted that the bias in the test items is "too subtle and too elusive to be detected" (*New York Times* 16.3.83). Hence the drive by some groups to abolish selection tests entirely. Other selection criteria are suggested instead, such as interviews, diplomas, and educational credentials, all of which have less validity than tests and are much more liable to bias.

Reducing Tests' Adverse Impact

Attempts to modify existing tests or to construct new tests in ways that reduce racial disparities in the score distributions have two main effects: (1) they necessarily reduce the *g* loading of the tests; and (2) they markedly decrease the tests' predictive validity *within* each of the groups. In brief, it has not been found possible to increase test validity for individuals *within* any group without also increasing the mean difference

between groups. The reason is that the main factor in individual differences is the same as the main factor in group differences and is also the main factor that accounts for tests' practical validity. This presents an inescapable dilemma if the improvement of test validity is a primary concern.

Abandoning Colourblind Top-down Selection

A procedure for eliminating adverse impact completely is known as within-group scoring. It has been officially sanctioned for use by the US Employment Service, whose General Aptitude Test Battery (GATB) is probably the most widely used of all tests for personnel selection (Hartigan & Wigdor, 1989). Like all other g loaded tests, the GATB has shown adverse impact on blacks and (to a lesser degree) Hispanics. To eliminate adverse impact, the raw scores (i.e. number of correct answers) are converted to percentile ranks *separately* within each racial/ethnic group. The applicants for a particular job then are ordered on these within-group percentile scores and they are selected from the top down until the required number of positions is filled or the supply of applicants who score above some minimum level deemed necessary for probable success on the job is exhausted. The effect, of course, is that the person's racial classification, as well as the person's actual level of performance on the test, determines his or her percentile score, on which selection is based. For example, a raw score that places a white applicant at the 50th percentile based on the white norms might place a black applicant with the same raw score at the 86th percentile of the black norms. The racially adjusted scores thus favour members of the group with the lower mean raw score. It is tantamount to adding or subtracting some value from every applicant's raw score according to the applicant's racial/ethnic classification, with the values determined so as perfectly to equalise the adjusted score distributions of all of the racial/ethnic groups. The practical effect of this procedure is drastically to reduce the predictive validity of the test. I have even seen one instance in which top-down selection based on racially adjusted scores resulted in a *negative* validity coefficient; that is, in actual job performance the majority of the lower-scoring (i.e. on racially adjusted scores) employees (mostly whites) out-performed the higher-scoring employees (mostly blacks). If, for the same persons, the racially adjusted scores were converted back to raw scores, the raw scores would, of course, show a *positive* predictive validity coefficient.

Affirmative Action

Originally, affirmative action was a national effort to increase the representation of minorities in college enrolments and in higher status jobs by actively encouraging and recruiting qualified minority persons—measures widely accepted as fair and beneficial to society as a whole. There is a high degree of national consensus on the moral imperative of non-discrimination. When it became evident, however, that these measures by themselves fall far short of achieving the desired racial parity, affirmative action policy sanctioned another tactic: relaxed qualifications, racial preferences, and other special dispensations for under-represented minorities. Thus affirmative action became highly controversial, being equated by many with the injustice of racial quotas. Some critics even renamed it affirmative discrimination and reverse discrimination. The idea of selection based on any criteria other than pertinent qualifications was strongly opposed by a vast majority.

Yet a lowering of selection standards for certain groups is inescapable if the primary goal is racial parity in college admissions and in hiring. (Parity is defined as occurring when the proportions of different population groups selected correspond to their proportions in the general population or in the pool of available applicants.)

Take the case of admissions into high-prestige selective colleges. The lower-bound IQ of students who succeed in such colleges is about 110; some 25% to 30% of the white and Asian populations in the US falls above that level, while less than 5% of the black population does so. To achieve racial parity in enrolments, therefore, the cutoff score for blacks, in IQ terms, would have to be lowered to 95. Under that admissions policy for blacks, only about 20% of them would be above the level of ability typically needed to succeed academically. Since the selection test (commonly the Scholastic Aptitude Test, or SAT), predicts academic performance with the same degree of accuracy (or inaccuracy) for white, black, and Asian students, a much higher proportion of black than of white and Asian students would be expected to fail, assuming the same grading standard for all students.

An additional problem arises if the institution tries to achieve racial parity by setting different admission standards for various groups while at the same time it imposes a limit on the total number of admissions. If there is top-down selection within each group, admission standards have to be raised for some groups to limit their admissions to make room for groups for whom lowered standards are necessary to bring their number up to parity. This has the inevitable effects of narrowing the range of ability in the group with the higher admission standard and widening the competitive gap between it and the groups admitted with lower academic qualifications.

A striking example of these effects can be seen in the University of California, Berkeley, which for several years now has had an energetic affirmative action policy that comes close to achieving parity for blacks and Hispanics, at least in freshman admissions. The results are dismayingly (Sarich, 1990a,b).

First, white students are now markedly under-represented in admissions; although whites constitute about two-thirds of all the high school graduates who are eligible by the traditional standards (the academically top 12.5%), they constituted only about one-third of those admitted as freshmen in 1989. The 12.5% of black students in the freshman class, however, is almost five times greater than the percent of blacks (2.5%) in the total pool of those eligible by traditional standards. The corresponding percentages for Hispanics are 23% and 6.9%. Asians are as yet virtually unaffected one way or the other by the Affirmative Action policy; they constitute 20% of the academically eligible and 23% of the freshman class, but Asians are only about 8% of all California high school graduates. In summary, at present the ratio *number admitted/number eligible* (by normal standards) for the different groups is: white 0.5, Asian 1.1, Hispanic 3.3, black 4.8. True parity would equal 1.0.

Second, in academic aptitude and performance the undergraduate student body has become a two-tiered student population comprising 57% white and Asian in the upper tier and 35% black and Hispanic in the lower, with only 10% overlap between the SAT score distributions of the two tiers. That is, 90% of the combined black and Hispanic group have lower SAT scores than 90% of the white and Asian group, a difference in average levels of academic aptitude that is reflected in the bimodality of the distribution of examination grades in large freshman courses.

Third, the most deplorable consequence for the special admission students is their extraordinarily high rate of attrition, which takes its greatest toll in the freshman and sophomore years. As many as 75% to 80% of the black students admitted in recent

years either flunk out or drop out before graduation. An added misfortune is that many motivated minority students who are thus lost to Berkeley, and perhaps leave with bad feelings about themselves or about academe, might have thrived in a more suitable college.

Legislating Equality of Educational Outcome

Since racial parity in college admissions does not persist through to graduation, still other measures would obviously be necessary to produce equality of educational outcome, as least as formalised by grades and graduation. A step in this direction was taken recently by the California State Legislature, in which bills were proposed setting a deadline by which time the racial/ethnic composition of the graduating classes of every state-supported college and university would be required to mirror the racial/ethnic composition of the population of all recent public high school graduates in California. The bills also made similar stipulations about the racial/ethnic composition of college faculties. Both the hiring and promotion of minority faculty would have to reflect their proportions in the state's population. These remarkable bills were vetoed by the Governor of California. He did not run for re-election and would not be held politically accountable. Similar bills will most likely be introduced in the legislature again, and a governor with hopes of re-election will probably think twice about exercising the veto.

Nothing at all has been officially stated about *how* the aims of such legislation can be realised. Perhaps it will be by changing course requirements—for example, cutting the number of required courses in basic sciences in the pre-med and medical school curricula. Other steps, already a reality in some colleges, include relaxed grading standards and removal of incompletes and failing grades from the college transcripts of special admission students.

There is talk on college campuses about the importance of cultural diversity. Amid the demands for a curriculum which pays more attention to the experience and values of America's underprivileged minorities, it is often added that black and Hispanic students, to succeed in academe, need professors of their own race or ethnic group as role models. Putting aside the question of the specific merits of this argument (and the obvious counter-example of Asian students, who are quite successful academically, despite the very low number of Asian teachers and professors they typically encounter), the practical constraints on hiring significant numbers of minority faculty are far greater than in the case of student admissions. To understand why, one must recall that certain minorities, particularly blacks, are extremely under-represented on college faculties, after more than two decades of affirmative action effort in minority recruitment. There is a widespread implicit assumption that if only much more of the academic talent pool of under-represented racial/ethnic minority populations were drawn into university graduate programmes, the supply of minority faculty in the nation's colleges and universities could eventually approach parity. This assumption, however, fails to take realistic account of the probable size of the talent pool, especially in those disciplines with the poorest minority representation. At present, for example, about half of all doctoral degrees (PhDs and EdDs) earned by blacks are in the field of education, while blacks and Hispanics combined earn fewer than 2% of all doctoral degrees in the sciences and engineering, although these groups constitute about 30% of the college-age population of the USA. In science and technology, then, the parity gap is extreme.

What is the realistic likelihood of appreciably increasing the percentage of minority university faculty, assuming the usual qualifications? University faculty are generally hired with top-down selection from the pool of new PhDs. The average PhD has been found to have an IQ of about 130, so let us assume that the majority of those hired as assistant professors in major universities are above that IQ level, as are most probably those eventually promoted to associate and full professorships. Given these fair assumptions, we can then make a rough statistical estimate. The calculation is best based on the black population, about which we have more certain knowledge of the IQ distribution and their population frequency than we have for other minority groups.

In the entire country, then, how many blacks turning 18 years of age in any one year would have a level of aptitude equivalent to an IQ of 130 or above—the level predictive of success in an academic career? Now the US black population has a mean IQ of about 85, with a standard deviation (SD) of 15. IQ 130 is 3 SDs above the black mean IQ of 85. As IQ has an approximately normal distribution, we can infer that only about 13.5 persons in 10 000 would fall above 3SD (i.e. IQ 130 in this case). According to recent US Census figures, there are about 420 000 black youths in any one-year cohort, so the total number above IQ 130 would be about 600. This, to begin with, is indeed a small pool of high-level academic talent, considering that it has to be divided among the demands of the many institutions competing each year for black students with the potential successfully to pursue advanced degrees and enter the professions, including college and university teaching positions. Not quite all of these 600 or so youths would be expected to graduate from high school or opt for college. Of those who go on to college, most will enter the world of business and various other professions. Some small fraction will elect to pursue a doctoral degree. A wide variety of attractive opportunities besides an academic career await those who earn advanced degrees. Hence there would be far too few black candidates in each year's crop of graduates who are normally qualified for faculty positions to come anywhere near meeting the desired racial parity in the colleges and universities of the entire nation.

The most common response to this dilemma is to disparage top-down selection for minorities. Instead of seeking the best available regardless of race or ethnicity, there are those who advocate preferential hiring of minority candidates who fall within the range of adequacy. This probably amounts to minimal qualification for the job and that is apt to be a downward sliding scale to meet increasing demands for parity. Non-traditional criteria for hiring or promotion would replace the usual, most relevant qualifications for performance in a particular position.

The drawbacks to such policies, which should be obvious, remain hidden far in the background of public concern. Consider their damaging effect on race relations. Affirmative action, when it involves racially preferential selection, creates a college or workplace where the association between race/ethnicity and performance becomes conspicuous, to the disadvantage of the less competitive groups. Also, discrimination on the basis of race instead of selection based strictly on relevant criteria engenders resentment and racial tensions. Many people still believe that racial discrimination is best eliminated by strictly enforced *non*-discrimination, rather than by any form of preferential treatment based on race. Racial preferences and privileges in educational selection, hiring, and promotion seem especially stigmatising to those minority persons who meet the traditional qualifications for pursuing their ambitions and can succeed without any special dispensations. Their real ability and achievement are devalued when they are identified as belonging to a group for which selection criteria are based in part on race or ethnicity. Those who would qualify by colourblind criteria and

succeed without special dispensations are the ones who should protest race-conscious selection most loudly.

CONCLUSION

I have argued that both individual and average group differences in Spearman's g are presently the primary factor in virtually all of the manifestations of inequality of educational outcomes, from kindergarten through graduate school, as well as their correlation with employment. Proposed solutions to the social and economic problems associated with educational inequality, including the troubling issue of increased racial tensions, will, I believe, fail in the short run and prove harmful to the whole society in the long run, to the extent that the central role of g in this picture is not fully recognised. Yet it is virtually a social taboo in this day to bring g into public discussions of the current problems of education. Popular cause-and-effect explanations eschew it, and theories based exclusively on broad sociological factors without recognising the concept of *ability* have brought forth unworkable solutions. The very terminology of explanations promulgated by many professionals of the race problem in America consists of the rhetoric of *victimisation* and *blame*: white racism, caste thinking, oppression, discrimination, white supremacy, and appeals to white guilt (e.g. see the collection of essays edited by Thomas, 1990). The language of such commentators seems to reflect and nurture an incipient racial paranoia that augurs ill for future race relations. Thus one author complains that "the gatekeepers of power have devised new and more complex strategies for circumventing the attainment of equity and that institutions now have seized the 'quality' criterion as a mechanism for denying equity in higher education" (quoted from Thomas, 1990, p. 58).

The popular sociological, historical, and cultural theories of educational disparity and its social consequences are weakest in considering the direction of cause-and-effect relationships. They are much too broad and indirect for zeroing in on precisely what happens at the cognitive level when children achieve, or fail to achieve, in school. Understanding the causes of educational inequality means in large part understanding the nature and causes of differences in g . A quite different order of research than has prevailed heretofore will be needed to discover the psychological factors and brain processes through which broad cultural and social factors become correlated with educational outcomes. Proposed solutions to what we perceive today as the problems of educational inequality, if they are to be realistic and potentially effective, cannot ignore the findings of psychometrics and differential psychology. These will have to be coherently integrated with the data from sociology and anthropology. Moreover—if science is to be allowed to run its course—even biological, evolutionary, and genetic facts and theories will need to be considered as well. Indeed, to shun them in the mistakenly fatalistic belief that the human species is genetically fixed would be a tragedy.

NOTES

- [1] The designation 'black' here refers to Americans of African ancestry, formerly termed Negro, but also called Afro-American or African American. Nearly all American blacks have some genes (averaging about 20–25%) from non-African ancestors (usually European Caucasoid).
- [2] 'White' here refers to Caucasoid persons of predominantly European descent.

- [3] 'Asian' (also 'Asian American') refers here to Mongoloid persons mainly of Chinese and Japanese origin, but including persons of Korean, Taiwanese, Vietnamese ancestry or other Mongoloid persons of Asian and Southeast Asian origin. East Indians, Pakistanis, and Filipinos may also be classified as 'Asian', but they are a very small proportion of all Asians in the US.
- [4] 'Hispanic' is a racial and ethnically unsatisfactory term, although it has gained wide currency in the USA. It refers to those whose ancestry or background is a country in which Spanish (or Portuguese) is the predominant language and whose surnames generally are Spanish. The group includes Americans of Puerto Rican and Cuban background, and a distinction is often made between Hispanics of Central and South American origin (called 'Latinos') and those of Mexican origin ('Mexican-American' or 'Chicano'). 'Hispanic' in California and generally in the Southwestern USA includes mostly Chicanos. Chicanos are also a quite heterogeneous group, with varying degrees of European (mostly Spanish) and Mexican-Indian ancestry.

REFERENCES

- CLARKE, A.M. & CLARKE, A.D.B. (1989) The later cognitive effects of early intervention, *Intelligence*, 13, pp. 289-297.
- COHN, S.J., CARLSON, J.S. & JENSEN, A.R. (1985) Speed of information processing in academically gifted youths, *Personality and Individual Differences*, 6, pp. 621-629.
- GARBER, H.L. (1987) *The Milwaukee Project: preventing mental retardation in children at risk* (Washington, DC, American Association on Mental Retardation).
- GLUTTING, J.J. & McDERMOTT, P.A. (1990) Principles and problems in learning potential, in: C. R. REYNOLDS & R. W. KAMPHAUS (Eds) *Handbook of Psychological and Educational Assessment of Children: Intelligence and Achievement* (New York & London, The Guilford Press).
- GOTTFREDSON, L.S. (Ed.) (1986) The *g* factor in employment, *Journal of Vocational Behavior*, 29(3), pp. 293-450.
- GOTTFREDSON, L.S. (1988) Reconsidering fairness: a matter of social and ethical priorities, *Journal of Vocational Behavior*, 33, pp. 293-319.
- HAIER, R.J., SIEGEL, B.V., NUECHTERLEIN, K.H., HAZLETT, E., WU, J.C., PAK, J., BROWNING, J.L. & BUCHSBAUM, M.S. (1988) Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography, *Intelligence*, 12, pp. 199-217.
- HARTIGAN, J.A. & WIGDOR, A.K. (Eds) (1989) *Fairness in Employment Testing* (Washington, DC, National Academy Press).
- JAYNES, G.D. & WILLIAMS, R.M. Jr. (Eds) (1989) *A Common Destiny: Blacks and American Society* (Washington, DC, National Academy Press).
- JENSEN, A.R. (1980) Uses of sibling data in educational and psychological research *American Educational Research Journal*, 17, pp. 153-170.
- JENSEN, A.R. (1984) Test validity: *g* versus the specificity doctrine, *Journal of Social and Biological Structures*, 7, pp. 93-118.
- JENSEN, A.R. (1985a) The nature of the black-white difference on various psychometric tests: Spearman's hypothesis, *Behavioral and Brain Sciences*, 8, pp. 193-219.

- JENSEN, A.R. (1985b) The black-white difference in *g*: a phenomenon in search of a theory, *Behavioral and Brain Sciences*, 8, pp. 246–263.
- JENSEN, A.R. (1986) *g*: artifact or reality?, *Journal of Vocational Behavior*, 29, pp. 301–331.
- JENSEN A.R. (1987a) The *g* beyond factor analysis, in: J. C. CONOLEY, J. A. GLOVER & R. R. RONNING (Eds) *The Influence of Cognitive Psychology on Testing and Measurement* (Hillsdale, NJ, Lawrence Erlbaum Associates).
- JENSEN A.R. (1987b) Individual differences in the Hick reaction time paradigm, in: P. A. VERNON (Ed.) *Intelligence and Speed of Information Processing* (Norwood, NJ, Ablex).
- JENSEN, A.R. (1987c) Further evidence for Spearman's hypothesis concerning black-white differences on psychometric tests, *Behavioral and Brain Sciences*, 10, pp. 512–519.
- JENSEN, A.R. (1989a) The relationship between learning and intelligence, *Learning and Individual Differences*, 1, pp. 37–62.
- JENSEN, A.R. (1989b) Raising IQ without increasing *g*? A review of "The Milwaukee Project: preventing mental retardation in children at risk," *Developmental Review*, 9, pp. 234–258.
- JENSEN, A.R., COHN, S.J. & COHN, C.M.G. (1989) Speed of information processing in academically gifted youths and their siblings, *Personality and Individual Differences*, 10, pp. 29–34.
- JENSEN, A.R. & REED, T.E. (in press) Conduction velocity in a brain nerve pathway of normal adults correlates with intelligence level, *Intelligence* (in press).
- JENSEN, A.R. & SINHA, S.N. (1991) Physical correlates of intelligence, in: P. A. VERNON (Ed.) *Biological Approaches to the Study of Human Intelligence* (Norwood, NJ, Ablex).
- KRANZLER, J.H. & JENSEN, A.R. (1989) Inspection time and intelligence: a meta-analysis, *Intelligence*, 13, pp. 329–347.
- MAEROFF, G.I. (1988) Withered hopes stillborn dreams: the dismal panorama of urban schools, *Phi Delta Kappan*, 69, pp. 633–638.
- NATIONAL COMMISSION ON EXCELLENCE IN EDUCATION (1983) *A Nation at Risk* (Washington, DC, US Department of Education).
- SARICH, V. (1990a) The institutionalization of racism at Berkeley. *Academic Questions*, 4, pp. 72–81.
- SARICH, V. (1990b) Making racism official at Cal., *California Monthly* [magazine of the California Alumni Association], 101, pp. 17–18.
- SNYDERMAN, M. & ROTHMAN, S. (1988) *The IQ controversy: the media and public policy* (New Brunswick, NJ, Transaction).
- SPEARMAN, C. (1927) *The Abilities of Man* (London, Macmillan).
- SPITZ, H.H. (1986) *The Raising of Intelligence: a selected history of attempts to raise retarded intelligence* (Hillside, NJ, Lawrence Erlbaum Associates).
- THOMAS, G.E. (Ed.) (1990) *US Race Relations in the 1980s and 1990s: Challenges and Alternatives* (New York, Hemisphere Publishing Corporation).
- VERNON, P.A. (Ed.) (1987) *Speed of Information Processing and Intelligence* (Norwood, NJ, Ablex).

Correspondence: Professor Arthur R. Jensen, School of Education, University of California, Berkeley, CA 94720, USA.