INTELLIGENCE **11**, 193–198 (1987)

EDITORIAL: Psychometric *g* as a Focus of Concerted Research Effort

Arthur R. Jensen
University of California, Berkeley

"Intelligence" has been tried and apparently it has failed.

Anyone who reads the recent book *What is Intelligence? Contemporary Viewpoints on its Nature and Definition,* edited by Sternberg and Detterman (1986), in which 25 experts responded to the question posed in the title, could easily conclude that there are about as many different conceptions of "intelligence" as the number of experts.

This was also true back in 1921 when the same question was asked of an earlier group of experts. Comparing the two symposia, separated by 65 years, we find scarcely more consensus among the experts today than in 1921.

Shouldn't this be cause for dismay to us who work in this field? Is there something wrong with our science, that there is so little agreement, even about what are the key questions, much less the answers?

One of the intellectually appealing features of progress in the more advanced empirical sciences is the increasing consensus among scientists regarding the definition and nature of the main phenomena within their purview. Investigation of some coherent body of phenomena leads to more agreement about the crucial questions and problems. True, there is always a frontier of unsettled issues and controversies in science; but with consolidation of the gains of research, there is also a growing consensus. The history of the physical and biological sciences indicates that generally, as a topic develops, an increasing majority of researchers acknowledge the key phenomena that call for more intensive investigation. Further investigation widens recognition of converging lines of evidence, sharpens focus on pivotal questions, and stimulates more concerted effort, both cooperative and competitive, in the search for answers—the kind of answers based on enough empirical support to compel acceptance by a majority of experts. A well known example is Watson and Crick's discovery of the double helix structure of DNA. A broader example is the continuing study of evolution, from Darwin to the present.

If these desiderata prevailed in the study of human abilities during the more than 100 years since Sir Francis Galton introduced this subject into the domain of science—almost as long ago as the study of evolution—shouldn't we expect by

Correspondence and requests for reprints should be addressed to Arthur R. Jensen, School of Education, University of California, Berkeley, CA 94720.

now something more satisfying, scientifically, than the welter of diverse and contradictory opinions exemplified by the recent symposium? Where are indications of cumulative gains of research, converging lines of evidence, and generally accepted definitions, concepts, and formulations? In philosophy, religion, and politics, of course, we never expect such progress. We might justifiably despair of ever seeing it on the subject of "intelligence."

Yet we believe that our undoubtedly important observation of individual differences in cognitive abilities should be treated in the usual manner of empirical science. As we work to advance on this path, it would help if we could cut through the semantic impasse and break the perpetual kaleidoscope of concepts, notions, and philosophies concerning the definition of "intelligence."

There is no value, scientifically, in arguing about definitions per se. They are merely conventions, but two or more distinctly different concepts should not be represented by one and the same term. A good rule in science is *one word-one concept.*

However, "intelligence" has mainly three quite different meanings, each exclusively favored by different psychologists. Some regard "intelligence" as the sum total of *all mental abilities,* or as the entire repertoire of a person's knowledge and skills at a given time. Others equate "intelligence" with *g,* or the most *general factor* common to all mental abilities. The "contextualists" acknowledge a wide variety of abilities and skills but claim that what any particular culture means by "intelligence" is only a *selection* of certain abilities from the entire cognitive domain that are deemed most important in the culture. In this view, the particular selection of abilities subsumed in "intelligence" is a cultural artifact, a composite of whatever abilities are especially valued in the dominant culture, or "intelligence" may consist of just those abilities that some test constructor happens to select, either to achieve some specific practical validity or to accord with his particular theoretical notions. (Note how test designers promote each new "intelligence" test as measuring a better selection of abilities than any previous test.) The broadest contextualism views "intelligence" as a composite of not only certain culturally valued abilities, but also of certain personality traits, motives, values, interests, and attitudes that may be correlated with various criteria of "real life" achievements recognized in a given culture. It has even been suggested that "intelligence" should include deftness and grace in body movements.

Unlike the contextualist's "intelligence," which is *invented,* the constituent mental abilities and their basic process components are *discovered.* Psychologists, to be sure, invent tests that measure certain abilities, but they don't invent the ability factors and cognitive processes that are the sources of variance on the tests. Various item-homogeneous tests distinguishably measure different mental abilities when the single test's reliability coefficients are higher than the correlations between the various tests. (By "mental" I mean simply that variance in test scores doesn't consist mainly of variance due to sensory acuity or motor skill. By "ability" I mean performance that can be graded on speed, accuracy, or quality according to an objective standard.) Because we can invent new tests, thereby making it possible that we may discover other abilities not as yet identified psychometrically, "mental ability" is necessarily an open-ended concept. Yet each discovered ability will rarely, if ever, be uncorrelated with all other abilities. As far as we can now determine, virtually all measurable abilities are correlated positively with one another to varying degrees. The structure of their overlapping variances can be represented by factor analysis. All factor models agree in finding that the total variance in any large and diverse battery of ability tests can be resolved into a number of uncorrelated components. Hence, those who equate "intelligence" with the whole of mental abilities obviously have it easy in demonstrating that "intelligence" is not unitary, but multifactorial. Many orthogonal (i.e., uncorrelated) factors can be extracted from any large collection of diverse tests, each composed of highly homogeneous items. "Mental ability" comprises all the orthogonal factors that could be extracted from an indefinitely large collection of varied mental tests. Just how many factors might exist is not known and, in principle, may never be known for sure, because there is no foreordained limit to the discovery of mental abilities, any more than there is to the discovery of subatomic entities in the physical universe. By subjecting newly invented mental tests to a factor analysis that includes other tests of various known abilities, we might discover additional small factors that are orthogonal to all the rest. The ultimate number of orthogonal factors cannot be unlimited, of course, but we have no idea what that number is.

Of what use, then, is the word "intelligence" if it is just redundant with "mental abilities"? As such, it can't be measured by a single score, and we can't even say how many different scores would be needed to represent it. "Intelligence" is either redundant with the whole of "mental abilities," and is, therefore, an unnecessary term, or it is only a limited selection of certain abilities or factors (and also personality traits, in some notions). Empirical science, however, has no means for compelling agreement about which cognitive abilities (or noncognitive traits) should be included in a limited selection. This is probably the reason that, in more than 65 years, there has been no appreciable convergence of expert opinions on the meaning of "intelligence." There exists no objective method for deciding between our persistently differing opinions about which abilities or factors we should include in the concept. It could hardly be decided by fiat. After all, scientists can't live with doctrines or decrees. Even if we all decided to agree, say, to equate "intelligence" with only the *g* factor derived from a hierarchical factor analysis of a large number of diverse mental tests representing the whole domain of known mental abilities, it would be a considerable improvement certainly. However, making "intelligence" synonymous with *g* would still be another redundancy; and even when technically defined as *g*, "intelligence" would unfortunately still evoke the troublesome overtones of common usage.

For scientific purposes then, "intelligence" can best be thrown out altogether. Its use will continue in popular parlance, of course, to mean anything that speakers may feel it means; and *Intelligence* is perfectly acceptable as the name of this journal, just as *Nature* is an acceptable journal title. Scientists don't actually try to define "nature" or to construct "theories of nature." "Intelligence" ought to be regarded like "nature" in this respect.

The two classes of natural phenomena that are really the basic grist of our research are *abilities* and all the orthogonalized hierarchical *factors* that can be extracted from their intercorrelations. The *g* factor is but one among many, although it is an especially important one, for reasons I will mention. It is objectively based on the empirical fact of totally positive correlations among virtually all known mental abilities (when these are measured in an unrestricted sample of the population), and, therefore, can be studied as a natural phenomenon in the tradition of empirical science.

The facts that lend *g* so much scientific interest and importance have been documented in detail elsewhere (Jensen, 1986, 1987). They can be briefly summarized.

Psychometric *g* is by far more highly correlated with all tests conventionally called "IQ", "cognitive abilities," and the like, than any other single factor or combination of other factors independent of *g*. Also, *g* is by far the largest part of the validity of multitest batteries used for selection or prediction of success in school and college, in the armed forces, and in business and industry. This fact indicates that many "real life" kinds of performance, and not just psychometric tests, are substantially *g*-loaded.

Summing the separate scores on a number of different ability tests does, in fact, yield a meaningful total score which has practical validity, because most of the total variance consists of the covariance among all the subtests, of which *g* is the largest part. The greater the number of diverse mental tests in a battery, the closer the composite score approximates *g*.

The stability of *g* across entirely different batteries of diverse tests is remarkably high. For example, when each of 17 various "target" tests was included singly in six entirely different batteries, each comprising only eight other tests, and the *g* factor was extracted from each battery, the stability of the "target" tests' *g* loadings across the six different batteries was indicated by an average stability coefficient of +.83 (Thorndike, in press). As the number of different tests in each battery is increased, the stability of the extracted *g* rapidly increases as it asymptotically approaches its "true" value.

In addition to these psychometric attributes of *g*, we have discovered in recent years that *g* is also related to certain variables that lie entirely outside the area of tests and factor analysis, variables that cannot even be included in the category of socially recognized achievements or other such typical correlates of *g*. Whether psychometric *g* is measured as the first principal factor in a common factor analysis or as the most general factor in a hierarchical analysis, it appears to be more highly related to various *non*psychometric variables than are any other factors or combination of factors independent of *g* that can be obtained from any possible rotation of the primary factor axes. The degree to which various psychometric tests, such as the Wechsler subscales, for example, are correlated with certain nonpsychometric variables is found to be directly related to the tests' *g* loadings. This relationship has been found for the *heritability* of various tests; the *spouse correlations* and other *kinship correlations* on various tests; the degree of *inbreeding depression* of scores on various tests, and its converse, *hybrid vigor* (variables which, in genetic theory, have important implications concerning the evolution of *g*); *evoked potentials* of the brain; the size of the average *black-white difference* on various tests; and *reaction time* (averaging less than one second) on very simple tasks that reguire no knowledge or acquired specific skills. (For references to all these studies, see Jensen, 1986, 1987.) Therefore, *g* is no mere artifact of psychometrics or factor analysis as some psychologists have mistakenly believed, but it is a real phenomenon, a variable which links psychology to biology and evolution.

Don't think for one minute that, because psychologists have known about *g* for a long time (Spearman first discovered it in 1904), there are no new frontiers to conquer regarding its nature. The most basic questions about *g* are begging for answers. Recent research has been making some headway. Studies of reaction time (RT), inspection time (IT), and the average evoked potential (AEP) now leave little doubt that a substantial part of the *g* derived from conventional psychometric tests can be measured by means of knowledge-free tests that are so simple that the only reliable measure of individual differences is sheer speed of information processing. Spearman claimed, but far from proved, that tests of simple sensory discrimination could also measure *g*. This conjecture, for which there is already a little evidence (Deary, in press), remains to be systematically investigated with modern methods. It could prove as surprising as the recent findings on RT, IT, and AEP (e.g., Brebner & Nettelbeck, 1986; Eysenck, 1982; Vernon, 1987). We need to know more about the stability of *g* across varied test batteries, going beyond Thorndike's study (in press). Little yet is known about the constancy of *g* across test batteries devised to measure the salient abilities in quite different cultures. Probably a rough road lies ahead for systematic studies of the important societal implications of *g*, which we are just beginning to see (e.g., Gordon, 1987; Gottfredson, 1986).

One of the most critical questions now is whether the causal underpinning of psychometric *g* is some unitary process or a number of uncorrelated processes. Both are reasonable hypotheses at present; but the evidence is conflicting and even paradoxical, almost like the theories of light as being either waves or particles. Willerman and Bailey (1987) cite striking evidence against multiple-process or sampling theories of *g*. They note, for example, that severely impaired performance on a particular test, due to sensory handicaps or chromosomal anomalies, is not always accompanied by poor performance on other correlated tests, making it doubtful that the correlation is attributable to the sharing of certain cognitive processes among the tests. On the other hand, some researchers have found that when *g* factor scores derived from psychometric tests are predicted from scores on each of several different elementary cognitive tasks (ECTs) in a stepwise multiple regression, the multiple *R* is appreciably larger than the correlation between the *g* factor scores and factor scores based on the general factor of the several ECTs. This would suggest that a number of different elementary cognitive processes contribute to the total variance of the psychometric *g* factor scores. The apparent paradox awaits resolution. If factor scores of two or more orthogonal factors derived from a sizeable battery of diverse ECTs that tap different cognitive processes are entered into a stepwise multiple regression and account for significantly more of the variance in *g* factor scores based on a large battery of diverse psychometric tests than is accounted for by the first principal factor of the ECTs, would this prove that psychometric *g* does not have a unitary cause?

Fathoming the nature of *g*, this most important phenomenon in our domain, is a large order indeed. Like other fundamental and complex problems in modern science, its full accomplishment will most likely depend on a concerted focus of research effort by many investigators in different laboratories around the world. What a sure sign of authentic progress in our field it would be if we could achieve a scientifically worthy understanding of one of psychology's most venerable phenomena and the one which seems most central to the human condition.

REFERENCES

Brebner, J., & Nettelbeck, T. (Eds.). (1986). Inspection time. *Personality and Individual Differences, 7,* 603–749.

Deary, I. (in press). The nature of intelligence: Simplicity to complexity and back again. In M. Shepherd & D. Forshaw (Eds.), *Maudsley series in the history of psychology and psychiatry.*

Eysenck, H.J. (Ed.). (1982). *A model for intelligence.* Heidelberg: Springer.

Gordon, R.A. (1987). SES versus IQ in the race-IQ-delinquency model. *International Journal of Sociology and Social Policy.* London: Vade Meum Press.

Gottfredson, L.S. (1986). Societal consequences of the *g* factor in employment. *Journal of Vocational Behavior, 29,* 379–410.

Jensen, A.R. (1986). *g:* Artifact or reality? *Journal of Vocational Behavior, 29,* 301–331.

Jensen, A.R. (1987). 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: Erlbaum.

Sternberg, R.J., & Detterman, D.K. (Eds.). (1986). *What is intelligence?* Norwood, NJ: Ablex.

Thorndike, R.L. (in press). The stability of factor loadings. *Personality and Individual Differences, 8.*

Vernon, P.A. (1987). *Speed of information-processing and intelligence.* Norwood, NJ: Ablex.

Willerman, L., & Bailey, J.M. (in press). A note on Thomson's sampling theory for correlations among mental tests. *Personality and Individual Differences, 8.*