

## THE NATURE OF INTELLIGENCE AND ITS RELATION TO LEARNING

*by* ARTHUR R. JENSEN

Intelligence and learning are two of the most central topics in psychology and education. Educators, by and large, are interested spectators to what psychologists are thinking and finding in their researches on intelligence and learning. But a rather puzzling spectacle it must seem. For intelligence and learning have traditionally been investigated largely apart from each other.

Psychologists have had little success, so far, in getting intelligence and learning under the same roof, theoretically speaking; and the attempts to do so have been quite sporadic and half-hearted. This conspicuous lack of a unified theory of intelligence and learning, existing for so many years in the history of psychology, has resulted, unfortunately, in our possessing less abundant and definitive empirical knowledge concerning the relationship between intelligence and learning than one might imagine, in view of the vast amount of empirical investigation and theoretical development that have accumulated separately in each of these domains.

Explanations of this hiatus usually appeal to the historical accidents in the development of scientific psychology which resulted in the peculiar division of our field into what Cronbach (1975), over twenty years ago, referred to as the 'two disciplines of scientific psychology'. He was referring, of course, to the fields of differential psychology and experimental psychology, and the methodological and theoretical gulf that separates them. It has become a popular pastime at psychological conferences on learning or intelligence to deplore this separation, so well described by Cronbach. But this laudable acknowledgement and consensus have never led to our doing much of anything to solve the problem. Although I do not like to contribute still one more example to this old refrain, let me remind you it is

historical fact that intelligence has always been the central interest of differential psychology, and learning has long since become the central interest of experimental psychology; and the two fields still share very little in the way of methodology, vocabulary, or theoretical formulation. Research on the nature and measurement of intelligence has sprung from an interest in the nature and causes of *individual differences*, and, to a lesser extent, group differences. Research on learning has sprung from an interest in the effects of experience on *behavioural change* and the improvement of individual performance. But the theoretical integration of these two domains—the study of *individual differences in behavioural change*—has been the no man's land of behavioural science.

There is probably a more scientifically intrinsic reason than mere historical accident for this divided state of affairs. A quite fundamental distinction underlies our concepts of intelligence and learning.

Imagine, if you will, a genius, a potential Newton of psychology, abandoned as an infant on an uninhabited island. Assume a beneficent Nature provides all he needs for survival and healthy growth. Then, as a young man in his intellectual prime, he spends much of his time thinking profoundly and scientifically about all aspects of his experience, including his own behaviour and thoughts. As a necessarily self-taught psychologist, will he come up first with a concept of intelligence or with a concept of learning? We can never know for sure, of course, because this imaginary experiment will never be performed. But I would bet heavily on this outcome: our socially isolated Newton of psychology would discover the phenomena of learning and formulate the 'laws of learning' long before he would develop the concept of intelligence, if indeed that concept ever occurred to him at all. I quite doubt that he would ever come up with the notion of intelligence. How would he ever notice it, or any lack of it, in himself, that is, in his thoughts or observations of his own behaviour? We may doubt that anything resembling our present psychometric concepts of intelligence would ever cross his mind. Yet, I bet he would gain the concepts of learning and memory, and perhaps even concoct theories of learning and memory not much different from those of present day behaviourists. For example, he would see his skills improve with practice, and notice more rapid improvement in practice situations in which there was the most marked

informative feedback from the environment, and would find that his performance improved (up to a point) under increased hunger, and so on. Similarly, through self-observation he would arrive at the principles of memory and forgetting á la Ebbinghaus, who, as you may recall, served as his own sole experimental subject.

Now, in view of our imaginary experiment, assuming we agree on the more probable outcome, does it not seem paradoxical that in human recorded history the concept of intelligence predates the concept of learning? I have been able to find references to what we would today call intelligence (although usually not so labelled) in the earliest known literature of ancient India and ancient Greece. But mentions of learning, in the sense in which psychologists use that term, are hard to find, if indeed they can be found at all. I suspect that, to these ancient philosophers, what we mean by 'intelligence' was a much more salient phenomenon than 'learning'. Even in the comparatively much shorter history of scientific psychology, interest in intelligence predates interest in learning and memory.

Then why is the rank order of salience of learning and intelligence just the opposite for our isolated Newton than it apparently was for the rest of humankind? You probably already see the answer. Our imaginary Newton did not have any human companions. The presence of only one other person with him on the island would probably have been sufficient for our Newton quickly to form a concept of intelligence. Or to be safe, perhaps we should add the proviso that the other person not be an identical twin to our Newton. The behavioural differences between his companion and himself, and more so if his companion were not another genius, would be among the most salient features of all his observations. He would notice differences between them in the amounts of information each one gains about their island, given the same amount of time for exploration; differences in the ages at which each one picks up certain information and develops particular skills, given much the same opportunities; and differences in the speed and frequency of success in solving the many problems and challenges of survival. Our Newton would wonder about all of these behavioural differences; and if he tried to explain them, he would inevitably induce a concept of 'ability'. With many more observations of differences in a variety of situations, he would formulate a concept of 'general ability'. It might just

happen, too, that he notices his companion has better visual and auditory acuity and stronger muscles than he himself possesses, despite his own greater prowess in acquiring information and solving problems; and so he develops a concept of 'general mental ability'. And there you have it all. '*General mental ability*' is just what psychologists mean by 'intelligence'. It was recognized in the earliest writings of the ancients, as it is recognized by most people today, although I suppose there will always be a rare few professional dissenters who deride the concept.

The fact is, however, that the concept of general mental ability, which we now call intelligence, is just as warranted, empirically and theoretically, as is the concept of learning. Each concept arises from two different classes of variance. The concept of learning arises from our observations of behavioural changes during the course of practice. The concept of intelligence arises from observations of differences among individuals in a broad class of behaviours involving comprehension, reasoning, and problem-solving. The general 'laws of learning' could, in principle, be derived from observations and experiments on a single person. Our inferences about the nature of intelligence, on the other hand, depend on observations or measurements of differences among individuals. One of the important challenges for behavioural scientists is to gain a greater knowledge of each of these natural concepts—intelligence and learning—and to understand the relationship between them.

#### *Working Definitions of Learning and Intelligence*

*Learning* is a theoretical construct needed to explain changes in response probability associated with certain stimulus contingencies. The effective stimuli may either precede or follow the particular response, or both. If the probability of some response of the organism increases or decreases in the context of such stimulus events, we infer that learning has occurred. To put it another way, the effect of learning is a change in response probability. Of course, to make this inference, we must rule out all other factors that can alter the probability of a response: physical maturation of the nervous system, fatigue, changes in emotional state, changes in drive state or arousal, aging, drugs, or physical impairment to the organism.

This is the most general and fundamental operational definition of learning I know of. It seems impossible for me to think of any example of learning it would not include. It is what all

of our vast literature on the experimental psychology of learning and our theories of learning are all about. They are attempts to describe precisely the properties of the stimulus conditions (which include reinforcements) that govern changes in response probability, and to hypothesize models of one kind or another that will scientifically explain the relationship between the stimulus conditions and the behavioural changes.

Notice that my operational definition of learning intentionally says nothing about the size or complexity of the responses or behavioural units that undergo change in probability of occurrence. Also, it says nothing about the duration of the change in response probability. The duration of learning, i.e. a change in response probability, can last anywhere from a fraction of a second up to most of the individual's lifetime. This is the issue of retention or memory (and its opposite, forgetting) which is basically a question of the duration of the effect of learning. So much, then, for my definition of what I shall mean, from here on, when I use the terms *learning* and *memory*.

*Intelligence* is a theoretical construct needed to explain the covariation (or correlation) of individual differences in performance in a wide variety of situations. There are, of course, certain boundary conditions to these situations, diverse though they are, since they do not include every kind of behaviour in which there are individual differences. In general, they are situations in which the individuals' behaviour in response to an implicit or explicit demand of the situation can be reliably classified or graded in terms of an objective criterion of adequacy or goodness of the response, and in which very little, if any, of the differences among individuals in the 'adequacy of responses can be attributed to the individual differences in sensory acuity or muscular strength and agility found among the general population. Thus the behavioural situations from which intelligence is inferred are thought of as 'mental' or 'cognitive' rather than physical. And these behaviours are roughly distinguishable from emotional states, attitudes, and personality traits in that they can be judged in terms of some objective, more or less universally agreed upon, standard of response adequacy, such as pass-fail, right-wrong, can-can't, fast-slow, knows-doesn't know, etc.

To really determine objectively just which behaviours do and do not belong to this category, broadly defined as 'mental abilities', we must resort to correlational analysis. Behaviours

that do not show significant correlations among one another are excluded. This is a wholly objective criterion for determining a natural domain of individual differences. It reaches its culmination in the class of correlational analysis now known as factor analysis, which in general is a method for determining the fewest 'principal dimensions' or 'latent roots' that will account for the correlations among a much larger number of independently measurable variables.

Intelligence, or general mental ability, is best defined operationally, in terms of this type of correlational analysis, as the *first principal component of an indefinitely large number of highly diverse mental tasks.*

The first principal component extracted from the matrix of intercorrelations among a large number of diverse mental tests is essentially the same as Spearman's *g*, which stands for 'general factor'. (Approximately the same *g* can also be arrived at by extracting a higher-order factor from the intercorrelations among a number of obliquely rotated primary factors.) I will henceforth use *g* to refer to the general factor in a large number of diverse mental tests and as an operational measure of intelligence. Single tests or some relatively small number of tests with especially high loadings on *g* are often called intelligence tests or IQ tests, and may yield a fair index of *g*, which is itself merely an index of individual differences in the theoretical construct called general mental ability or intelligence. When we devise an IQ test, our aim is that it should measure an idealized *g* factor as accurately as possible. By 'idealized *g*' I mean the first principal component of an unlimited number of tests of unlimited diversity. Thus any particular test, or even subset of all possible tests, can provide only an imperfect indicator of idealized *g*.

But, thanks to the methods of factor analysis, by studying the varying *g* loadings of many different types of tests and of single test items, we can gain some psychological insight as to the nature of *g* by characterizing the kinds of behaviour in which it is most highly manifested.

#### *The Nature of g*

During the past year or so I have reviewed practically all of the studies of the factor analysis of ability tests of every imaginable kind that have ever been published, as well as many that have not been published. In many cases, I have myself had to

extract the first principal component from the original correlation matrix provided by the authors when they have failed to perform an analysis that could have revealed a general factor.

Let me make an aside at this point. Although it is just an aside, I think it is the most important thing I have to say in this paper, and I will say it again and again in other papers. I think it is a cardinal sin, psychologically, to immediately do an orthogonal rotation of an abilities matrix, thereby completely obscuring or eliminating the possibility of seeing the general factor. In the abilities domain, especially, the *g* should *always* be extracted, either as a first principal component or as a higher-order factor from a hierarchical analysis of the oblique rotation of the first-order factors. (I prefer the first principal component, because it accounts for the largest possible linear component of variance in the matrix and is considerably less sensitive to sampling errors than the *g* extracted hierarchically as a higher-order factor.) Whatever else is done, the *g* loadings of the original variables should always be reported. Jumping immediately to orthogonal rotation of factors, which has been made so easy by computer programmes like Kaiser's (1958) varimax method, has been a blight in this field and could only be tolerated by persons having no real theoretical interest whatever in intelligence. That big first principal component, let me emphasize, is the key to intelligence. It is not the last word, to be sure. It is certainly not an explanation or a theory of intelligence. But I am convinced it is the first, most fundamental step toward developing a theory of intelligence. By ignoring the first principal component, we can only be severely handicapped.

A number of quite important generalizations can be drawn from inspection of the *g* factor in dozens of large-scale factor analyses of literally hundreds of tests and test items of various kinds. We look at the *g* loadings of the various tests and items; these vary between zero and one. It is informative to notice the characteristics of test items that differ markedly in *g* loadings and try and figure out just how the high and low *g* loaded tests differ. Spearman, of course, did just that more than half a century ago. I have recently done the same thing again, but with a much greater data base than was available to Spearman. Spearman concluded from such examination of the *g* loadings of a diversity of quite homogeneous tests that *g* is characterized by 'neogenesis', that is, the 'education of relations and correlates', or 'inventive', as contrasted with 'reproductive', behaviour

I believe this characterization of  $g$  is both more specific and more complex than necessary, and tends to imply more of a clear-cut dichotomy than can actually be discerned by inspection of the  $g$  factor loadings of innumerable measurements of human ability. Spearman was by no means wrong. His characterization of the most highly  $g$  loaded tests was perfectly correct. But he was not general enough.

What we actually see is a perfectly smooth continuum of the  $g$  factor loadings of various ability measures, ranging from close to zero to close to unity. When we arrange tests with their  $g$  loadings all in order of magnitude, we can begin to discern some of the characteristics of  $g$ . The most cognitively complex tests are at the top and the simplest tests are at the bottom. Do not confuse this test *complexity* dimension with *difficulty*. This is an important distinction, since the difficulty *per se* of test items is not related to  $g$ . Lifting a 200-pound weight, or recalling a string of ten digits, are extremely difficult, but they have very low  $g$  loadings. Difficulty is the probability of 'passing' a test item. Its relation to  $g$  is only incidental. We can find very highly  $g$  loaded items that range widely in difficulty.

Some of the complex tests that are highly  $g$  loaded are matrices, verbal analogies, number series, and figure analogies. The most  $g$  loaded tests are those yielding a single score based on a composite of a number of rather complex subtests, like the Stanford-Binet and Wechsler scales. But these are too complex to give us much insight into the nature of  $g$ .

Intermediate  $g$  loadings are found in tests involving delayed discriminations, word or figure recognition, speed of rote-learning, trial-and-error learning, pursuit rotor learning, and complex reaction time.

Quite low  $g$  loadings are found in simple addition, short-term memory, tapping speed, counting, speed of crossing out designated letters in words, simple reaction time, sensory acuity, muscular strength, and manual dexterity.

One of the most striking aspects of highly  $g$  loaded items is their infinite variety. It completely destroys the all-too-common notion that  $g$  is test-bound or culture-bound, that it is an artefact of a narrow class of tests reflecting scholastic and cultural attainments. Such scholastic and cultural items can indeed have high  $g$  loadings. But the measurement of  $g$  is not at all dependent on specific cultural or scholastic attainments. The gestalt completion test, for example, is highly  $g$  loaded. This test gets at



the ability to recognize mutilated silhouettes of common objects, all of which are highly familiar to the persons taking the test, and which virtually everyone would recognize instantly if the silhouettes were not mutilated. Note that, in this test, the person has mentally to fill a gap, turn something over in his mind, make a comparison, put one thing ahead of another, or search his memory store. In short, *g* appears to show itself as conscious mental manipulation either of stimulus inputs or of things retrieved from memory.

I have learned most about *g* from examining the *g* loadings of quite simple types of performance that differ in only one or two characteristics. Forward and backward digit span are a good example (Jensen & Figueroa, 1975). Persons are asked to repeat strings of digits either in the order in which they were presented (forward) or in reverse order (backward). Neither test is a very good index of *g*; but backward digit span is significantly more *g* loaded than forward digit span. Backward digit span obviously involves more mental manipulation, hence more *g*. On the other hand, longer, therefore more difficult, digit series are not more *g* loaded than shorter, therefore easier, digit series. So difficulty is not an intrinsic characteristic of *g*. We see this most clearly when scores on paired-associates and serial rote learning tasks are included in a factor analysis with good *g* reference tests. Under the condition of fast pacing of the stimuli to be learned, the *g* loadings of the paired-associate and serial learning scores are very low. Under slower pacing, making the learning task actually easier, the scores take on higher *g* loadings. Apparently the recruitment of *g* takes time, and with slower pacing of the learning materials, persons have more time to think, which is basically what *g* is all about.

Simple and complex reaction times (RT) are an even more clear-cut example. If persons only have to remove the index finger from a pushbutton the instant a light flashes on, the *g* loading of the RT is significantly less than if the persons have to look at two light bulbs and respond only when one but not the other goes on. Slightly more complex elaborations of this reaction time test show moderate *g* loadings. Decision and response in the face of uncertainty, however slight, such as the uncertainty as to which of two light bulbs will go on, bring out *g*.

The idea that what our best, most *g* loaded IQ tests measure is merely some narrow ability only important in school, or

'academic intelligence', as Ulric Neisser (1976) has termed it, is transparently false in view of this evidence. There is ample evidence that *g* is involved even in seemingly quite simple and commonplace activities that are exceedingly remote from school and academic pursuits. For example, it has been shown in on-the-job work sample tests given to U.S. army cooks, equated for months of experience in the kitchen, that the various routine tasks performed by cooks are differentially *g* loaded. Making jellyrolls, it turns out, is much more *g* loaded than preparing scrambled eggs! Whenever the task at hand, whatever it may be, involves stimulus complexity, uncertainty, mental manipulation, or retrieval of relevant information from memory, *g* is manifested. It is the very same *g* as is measured with useful accuracy by our present-day IQ tests.

Another thing we have found is that instruction, learning, and practice on a task first increases, and then decreases, its *g* loading. Whether the *g* loading first increases depends upon the initial difficulty level of the task. But if the task itself does not change over time, then instruction, learning, or practice on it invariably decreases its loading on *g*. That is why practice and overlearning of needed complex skills are a good thing: they result in the conservation of *g*. It is probably some such kind of observation that led Carl Bereiter to the clever definition of intelligence as 'what you do when you don't know what to do'. Persons can be taught strategies for solving certain types of highly *g* loaded intelligence test items and through practice on many such items they can improve their performance on them. But there is no evidence that this increases their overall standing on *g*. The highly practised type of items lose their *g* loading, and performance on them is a poor index of how the person will perform on different types of *g* loaded items. In the whole literature, I have not found any substantial evidence that intelligence, as *g*, can be expressly taught or learned. This is not to say, of course, that a multitude of micro-environmental factors does not influence the development of *g* from infancy to maturity, or that the behavioural manifestations of *g* may not be adversely affected in a person deprived of a normal social environment during his formative years.

#### *The Psychometrics of Intelligence*

IQ tests are comprised of many items that involve specific cultural content, like general information and vocabulary, all the

elements (or 'fundamentals', to use Spearman's term) of which could only be acquired in a particular culture at a particular period of history, and so it has often been said that IQ tests are culture-bound. But it is a serious mistake to believe, therefore, that IQ tests measure *nothing but* the specific cultural information and skills that persons have had the chance to acquire, and that individual differences in the scores on such tests merely reflect differences in opportunity for acquiring the specific bits of information or skills included in the various items of the IQ test. If the test is highly *g* loaded, it gets at something much more basic than what can be seen in the surface content of its items. It should not be forgotten that most of the individual differences variance on IQ tests, and *all* of the variance in *g*, is due not to the single test items *per se*, but to all of the correlations among a diversity of items. The total variance of an IQ test is composed of the sum of all the item variances plus the sum of all of the covariances among items. In psychometrics, the part of the total variance attributable to items is error variance (or unreliability); the part of the total variance attributable to covariance among items is the so-called true-score variance, or the reliability, of the test, and it usually amounts to more than 90 per cent of the total variance. In the case of most IQ tests, some 90 per cent of the true-score variance (or about 80 per cent of the total variance) in IQs is attributable to *g*. The remaining 10 per cent or so of the variance is attributable, in varying amounts, to small 'group factors', as they are called, such as verbal ability, numerical ability and spatial ability, which are linked more closely to specific types of test items. An ideally constructed intelligence test should balance out these group factors, thereby minimizing their contribution to the variance in the total scores on the test.

#### *The Evolution and Phylogeny of Intelligence*

The mistaken notion that the *g* measured by IQ tests resides in the specific item content and is therefore only an index of culture-specific learnings is most strongly contradicted by the study of the evolution and phylogeny of intelligence. I believe the de-biologizing of our conceptions of intelligence has been the greatest error of a whole generation of psychologists. But it appears that this de-biologizing of the science of human behaviour, in general, has already seen its day, with its few ardent

exponents now producing only ineffectual rear-guard skirmishes in their reluctant retreat from the scientific arena.

If psychologists' concept of intelligence were really just a specifically cultural artefact of modern Western civilization, as many of its critics seem to imply, then it should seem surprising indeed if a similar concept of intelligence arose independently in the field of zoology, from the comparative study of animal behaviour. In trying to gain greater insight into the nature of intelligence, it should be instructive to note the features that characterize differences in the behavioural capacities of various phyla and species of animals. There is virtually universal assent that some animals are more intelligent than others. By what criteria do we judge the dog to be more intelligent than the chicken, the monkey more intelligent than the dog, and the chimpanzee more intelligent than the monkey? Zoologists, ethologists, and comparative psychologists have amassed a great deal of information regarding this question.

The main indices of intelligence in animals are the speed of learning and the complexity of what can be learned, the integration of sensory information to achieve a goal, flexibility of behaviour in the face of obstacles, the amount of insightful as contrasted with trial-and-error problem-solving behaviour, transfer of learning from one problem situation to somewhat different situations, and the acquisition of abstract concepts.

It turns out that there is a definite relationship between ratings of animals' performances along these dimensions and the animals' phylogenetic statuses. Numerous ingenious behavioural tests have been devised to investigate this relationship, tests which permit equivalent comparisons of behavioural capacities of quite markedly differing animals despite their often vast differences in sensory and motor capacities. It is possible to give such diverse species as fish, birds, rats, cats, dogs, monkeys, and apes essentially equivalent forms of the same test problems.

Behavioural differences among species, like physical differences, are largely a product of evolution. Natural selection, by acting directly upon the behaviour involved in the organism's coping with its environment, indirectly shapes the physical structures underlying adaptive behaviour, of which the nervous system is the most important.

There is much evidence for evolutionary continuity in the behaviour of organisms, just as there is in their morphology. The phylogenetic differences in the complexity of behavioural

capacities are clearly related to brain size in relation to body size, and to the proportion of the brain tissue not involved in vegetative or autonomic and sensorimotor functions. Development of the cerebral cortex, the association areas, and the frontal lobes phylogenetically parallels behavioural complexity (Jerison, 1973). Also, the higher the animal ranks in the phyletic scale, the more seriously do lesions of the cortex of the brain affect its objectively measured behavioural capacity. Cerebral development, as reflected in cranial capacity, is known to have increased markedly over the five million years of human evolution, almost tripling in size from *Australopithecus* up to modern man.

Behaviourally, the phylogenetic hierarchy shows an increasing complexity of adaptive capabilities and an increasing degree of intersensory integration and greater breadth of transfer and generalization of learning as we move from lower to higher phyla. The degree of complexity and abstractness of what can be learned, given any amount of time and training, shows quite distinct differences between phyla. Each phyletic level in general possesses all the learning capacities (although not necessarily the same sensory and motor capacities) of the levels below itself, in addition to new emergent abilities, which can be broadly conceived of as an increase in the complexity of information processing. This increase of complexity of information processing is the common dimension along which animal tests of intelligence can be ordered. They represent a wide range of complexity: simple S-R conditioning, discriminant conditioning, trial-and-error learning, habit reversal, detour problems, learning-set acquisition, simple and double oddity problems, cross-modal transfer, transposition or relational learning, and insightful problem-solving that requires the animal to 'catch on' to a set of relationships, such as pulling food into the cage by means of strings attached to it, using a rake or joining sticks together to get out-of-reach food. (Many of these tests have been described by Viaud, 1960.)

The important point of this in the present context is that practically all of these tests of animal intelligence can be and have been given to humans, usually to children and the mentally retarded. Not only do these tests show correlations with our standard intelligence tests, such as the Stanford-Binet, but the degree to which they are *g* loaded, as indicated by their correlations with other highly *g* loaded human tests, is closely related to the tests' standings on the complexity dimension and the

levels of the phylogenetic hierarchy that are most discriminated by the tests. For example, both normal children and children with varying degrees of mental retardation have been given the battery of tests that Köhler used with chimpanzees. Exactly the same rank order of difficulty of these problems emerged for human children as for chimpanzees and lower primates (Viaud, 1960, pp. 44-5). The complexity factor common to these experimental animal problems, that differentiates species of primates, rank-orders human children the same as do standard IQ tests. Thus, the *g* factor of IQ tests reflects the same kind of ability to deal with complexity that is measured by the animals tests which most clearly reveal phylogenetic differences in behavioural adaptive capacity. The *g* factor thus is not just peculiar to individual differences among persons within a particular culture, but is continuous with broader biological aspects of neural organization reflected, as well, in individual differences within other primate species, and even in the evolutionary differences in behavioural capacities between various species. In this sense, intelligence is as much a biological reality, fashioned by evolution, as are the morphological features of organisms.

The now well-established fact of the substantial heritability of individual differences in scores on *g* loaded tests is further evidence of the biological basis of intelligence. There is also evidence of considerable genetic dominance for higher intelligence, which should be expected if intelligence is a fitness character, in the Darwinian sense, that has been subject to natural selection in the course of human evolution. The pattern of correlations among IQs of persons of various degrees of genetic relatedness, whether reared together or reared apart, must be adequately accounted for by any complete theory of intelligence. Genetic models based on general principles of genetics and applicable to metric traits in all plants and animals fit the observed kinship correlations for IQ remarkably well. No theoretical model that ignores genetics has ever been formulated that can explain the distinctive pattern of kinship correlations found for intelligence test scores, a pattern that is typical of many hereditary physical traits as well. There is no escaping the fact that an adequate theory of intelligence will have to be closely linked to genetic theory.

The *g* factor is correlated not only with behavioural measures, but also with anatomical and electrophysiological brain measurements. The relationship between brain size and IQ has been

greatly played down in most modern psychology textbooks. But a thorough and methodologically sophisticated recent review of all the evidence relevant to human brain size and intelligence concludes that the best estimate of the within-sex correlations between brain size and IQ is about + 0.30, taking proper account of physical stature, birthweight, and other correlated variables (Van Valen, 1974). Such a correlation is considered quite important from a biological and evolutionary standpoint, considering that much of the brain is devoted to noncognitive functions. Van Valen argues that there has been a direct causal effect, through natural selection in the course of human evolution, between intelligence and brain size. The evolutionary selective advantage of greater brain size was the greater capacity for more complex intellectual functioning. Also, it is now quite well established that IQ is correlated with various indices involving the speed and amplitude of electrical potentials in the brain, evoked by visual and auditory stimuli, and measured by the electroencephalogram (Callaway, 1975).

#### *Explanation of g*

As yet we have no well developed scientific theory of the neurophysiological nature of *g* itself. Any scientifically respectable theory of intelligence surely must reckon with the problem. A commentator at a recent conference on the nature of intelligence stated that the main consensus of the conference was that 'the understanding of intelligence must proceed via the relating of intelligence to psychologically meaningful processes' (Voss, 1976, p. 310). From what I have read of the attempts in this vein, I believe the psychological process approach will prove to be a blind alley. It is really a kind of faculty psychology in modern disguise, which attempts to explain intelligence in terms of learned strategies of information processing, verbal encoding, learning hierarchies of transferable skills, and the like, mostly describable in terms of a subjective analysis of what one does when faced with various cognitive problems. But I see no good reason to believe that the operation of the brain should resemble the subjective verbal description of our thought processes or the psychological analysis of cognitive experience.

I doubt that the nature of *g* can be characterized in terms of an analysis of the mental processes that are involved in any particular cognitive task. I believe that it is possible to analyse the cognitive process, as this term is generally used, in response

to specific problems. But I doubt that such a process analysis can explain individual differences in *g*. I have seen no hint of promise in any process analysis that it can explain the commonality of individual differences in such highly diverse *g* loaded mental measurements as vocabulary, the increase in reaction time with increased stimulus complexity, intra-individual variability in RT, oddity problems, figure copying, Piagetian conservation tests, embedded figures tests, and latency of evoked potentials.

David Wechsler (1958), who devised one of our best known IQ tests, has noted that *g* 'is independent of the modality or contextual structure from which it is elicited; *g* cannot be exclusively identified with any single intellectual ability and for this reason cannot be described in concrete operational terms' (p. 121). Wechsler further remarks that *g*

is not a factor at all in the sense that verbal comprehension, memory, etc., are factors of the mind . . . One should note . . . that unlike all other factors [*g*] cannot be associated with any unique or single ability; *g* is involved in many different types of ability; it is in essence not an ability at all, but a property of the mind (p. 124).

There is no time here to describe the various theories of *g* that have been propounded. But the one that seems to provide the most promising springboard for further developments in this field is the notion elaborated by Sir Godfrey Thomson, in his famous work *The Factorial Analysis of Human Abilities* (1939). He showed mathematically that the results of factor analysis, with its hierarchical structure involving a general factor and a number of group factors, could be explained in term of a 'sampling' model in which the elements sampled by particular tests are small, numerous, diverse, and need not be specified as to their physiological nature, a knowledge of which must depend upon investigations into the brain itself.

The action of the brain can be thought of as involving a large number of elements of various types: the number and amount of branching of brain cells, synaptic conductivity, thresholds of activation of neural elements, production of neurochemical transmitters, the blood supply to the brain and the richness of the capillary network, and so on. If various kinds of behavioural tasks involve different samples of these many elements, the degree to which performance is correlated across tasks will



depend upon the number of common elements they involve. It is a well-established empirical finding that the size of the correlations between various tests of diverse content is directly related to their degree of complexity.

#### *The Connection between Intelligence and Learning Ability*

For many years, definitions of intelligence usually included the 'ability to learn'. This would seem to accord with commonsense notions of intelligence and with simple observation. Is not the 'bright' or high IQ pupil a 'fast learner' and the 'dull' or low IQ pupil a 'slow learner'? Studies of learning ability in the experimental laboratory, however, apparently failed to uphold the commonsense connection between learning ability and IQ, and most psychologists dropped 'ability to learn' from their definitions of intelligence. Ability to learn was still regarded as a mental ability, to be sure, but not the same ability as intelligence, at least as it is measured by IQ tests.

A series of studies by Woodrow (1938, 1939, 1940, 1946) were the most influential for the belief that intelligence and learning are quite distinct abilities, and to this day it has left psychologists puzzled or at best unclear as to the connection between learning and intelligence. Woodrow's studies yield two main findings: (1) Measures of performance on a large variety of rather simple learning tasks showed only meagre intercorrelations among the various learning tasks, and between learning tasks and IQ. Thus factor analysis did not reveal any large general factor of learning ability. (2) Learning ability, measured as the rate of improvement with practice or gains in proficiency as measured by the difference between initial and final performance levels, showed little or no correlation among various learning tasks, or with IQ. Even short-term pretest-posttest gains, reflecting improvement with practice in certain school subjects, showed little or no correlation with IQ. Woodrow's results seemed to indicate that the speed of learning of simple skills and associative rote learning, and the rate of improvement with practice, have little in common with the *g* of intelligence tests.

Many other studies, following Woodrow's, essentially confirmed his findings. (Good reviews of this literature are presented by Zeaman and House (1967) and by Estes (1970).) It is a fact that the rate of acquisition of conditioned responses, the learning of motor skills (e.g. pursuit rotor learning), simple

discrimination learning and simple associative or rote learning of verbal material (e.g. paired-associates and serial learning) are not highly correlated with IQ, although the correlations are reliably greater than zero.

The studies that have led to these conclusions are justly liable to numerous methodological criticisms. A large part of the problem has been that 'learning ability' is much less precisely defined, delimited, and measured than intelligence. The psychometric features of most measures of 'learning ability' are not directly comparable to tests of intelligence. Each learning measure is more like a single item on an intelligence test in the narrowness and specificity of what it measures. Learning measures have not been developed with the same degree of psychometric sophistication that has been applied to the measurement of intelligence and other abilities. Gain scores, that is, differences between initial and final performance levels in the course of practice, are notoriously unreliable and should probably never be used in such research. Covariance analysis involving initial and final performance levels, or better yet, the use of estimated parameters of individual learning curves, are preferable to raw gain scores as measures of learning ability.

But even with such improvements of the indices of learning ability, I doubt that the picture would be markedly changed from the general impression left by the early studies of Woodrow. My examination of virtually the entire literature on individual differences in learning ability, memory, and intelligence has led me to the following tentative conclusions:

1. I believe the evidence forces us to reject the view, which has been quite popular with experimental psychologists in the field of learning, that intelligence is merely a product of learning, a kind of learned achievement, and that the IQ simply indexes the rate of learning certain types of more or less universally available information during the period from infancy to maturity. According to this view, the mental age (MA) is an index of how much the individual has learned by a given chronological age (CA); and the ratio of MA/CA (which is the IQ) is the *rate* of information acquisition.

2. A more tenable position, which has been the prevailing view among developmental psychologists, is that the brain is more than just a general learning machine; it undergoes organic growth and development of more complex functions from birth to early maturity. The MA, as measured by tests such as the

Stanford-Binet, is not essentially a measure of what has already been learned up to a given CA, but an index of the level of the individual's cerebral development, a level which limits the complexity of what can be learned and indicates the child's readiness for more complex forms of learning. Some concepts or skills that can be learned only with inordinate difficulty, if they can be learned at all, at one age, are learned easily, almost spontaneously, as a result of mere exposure, at a later age. The concept of learning 'readiness' springs from this common observation. The IQ, according to this view, is an index of the child's developmental status in learning readiness, in relation to his CA.

3. There appears to be no general factor of learning ability or of memory which is independent of the *g* of intelligence. A quite small general factor (first principal component) is found when a number of learning or memory tasks is factor analysed. (By 'small' I mean the general factor does not account for much of the total variance in all of the learning measures.) But here is the important point: this small general factor is not at all distinguishable from the *g* of intelligence tests. General factor scores from a diverse battery of learning and memory tests are as highly correlated with the *g* of a battery of intelligence tests as their reliabilities will permit. Whatever general factor can be extracted from learning and memory measures, provided they are more diverse than merely equivalent forms of the same task, is the same *g* as is found in cognitive tests of the kind that comprise most standard intelligence tests. There are small group factors in various learning and retention measures, and a great deal of task specific variance. It is uncertain why there is so much task specific variance in learning measures, but it is a fact and cannot be explained away in terms of measurement error. Learning tasks for which individuals closely approach asymptotic levels of performance, even when these levels show highly reliable individual differences, generally show increasing factorial specificity with increasing practice as the asymptotic level of performance is more closely approached. Practice and overlearning tend to divest learning tasks of nearly all factorial communality as well as of *g*. In terms of Thomson's sampling theory of *g*, we would conclude that fewer and fewer elements of cerebral functioning are involved in the performance of a complex task throughout the course of practice and especially as the skills involved become overlearned. It is as if the cerebral

mechanisms involved in a task become more highly specialized, more narrowly focused, as a result of continuing practice on the task.

4. I believe the brain mechanisms underlying learning and memory, independently of *g*, are phylogenetically more primitive than *g*; they evolved earlier and reached more or less maximal efficiency long before the greatest evolutionary developments of the neural mechanisms involved in *g*. Those behavioural capacities characterized by *g* have been superimposed by evolution upon the more basic, relatively unintegrated capacity of neural tissue for learning and retention at the molecular, cellular, and synaptic levels.

Until we have a well developed theory of the relation between intelligence and learning, we can only at this point answer with empirical generalizations the question of what are the conditions that govern the magnitude of the correlation between measures of learning and intelligence. Some types of learning clearly show higher correlations with IQ than other types of learning. There seems to be no clear-cut distinction between types of learning, in this respect, but a continuum, which in general can be characterized as going from the simple to the complex. Individual differences in learning proficiency show higher correlations with IQ directly in relation to the following characteristics of the learning task:

1. Learning is more highly correlated with IQ when it is *intentional* and the task calls forth conscious mental effort and is paced in such a way as to permit the subject to 'think'. It is possible to learn passively without 'thinking', by mere repetition of simple material; such learning is only slightly correlated with IQ. In fact, *negative* correlations between learning speed and IQ have been found in some simple tasks which could only be learned by simple repetition or rote learning but were disguised to appear more complex so as to evoke 'thinking' (Osler & Trautman, 1961). Persons with higher IQs engaged in more complex mental processes (reasoning, hypothesis testing, etc.) which in this specially contrived task only interfered with rote learning. Persons of lower IQ were not hindered by this interference of more complex mental processes and learned the material by simple rote association.

2. Learning is more highly correlated with IQ when the material to be learned is *hierarchical*, in the sense that the learning of later elements depends upon mastery of earlier

elements. A task of many elements in which the order of learning the elements has no effect on learning rate or level of final performance is less correlated with IQ than a task in which there is some more or less optimal order in which the elements are learned, and the acquisition of earlier elements in the sequence facilitates the acquisition of later elements. The dependence of learning rate on IQ, however, can be lessened to some degree by making the hierarchical structure of the learning sequence extremely explicit for all individuals, so that relationships between levels of the hierarchy do not have to be spontaneously discovered, induced, or inferred by the learner. Explicitly emphasizing the hierarchical aspects of the to-be-learned material is a prominent feature of programmed learning, computer assisted instruction, and 'mastery learning'.

3. Learning is more highly correlated with IQ when the material to be learned is *meaningful*, in the sense that it is related to other knowledge or experience already possessed by the learner. Rote learning of the serial order of a list of meaningless three-letter nonsense syllables or coloured forms, for example, shows relatively little correlation with IQ. In contrast, learning the essential content of a meaningful prose passage is more highly correlated with IQ. The IQ is much more highly related to comprehension than to memorization.

4. Learning is more highly correlated with IQ when the nature of the learning task permits *transfer* from somewhat different but related past learning. Outside the intentionally artificial learning tasks of the experimental psychology laboratory, little that we are called upon to learn beyond infancy is entirely new and unrelated to anything we had previously learned. Making more and better use of elements of past learning in learning something 'new'—in short, the transfer of learning—is positively correlated with IQ.

5. Learning is more highly correlated with IQ when it is *insightful*, that is to say, when the learning involves 'catching on' or 'getting the idea'. Learning to name the capital cities of the fifty states of the United States, for example, does not permit this aspect of learning to come into play and would therefore be less correlated with IQ than, say, learning to prove the Pythagorean theorem.

6. Learning is more highly correlated with IQ when the material to be learned is of *moderate difficulty and complexity*. If a learning task is too complex, everyone, regardless of his IQ,

flounders and falls back on simpler processes such as trial-and-error and rote association. Complexity, as contrasted with sheer difficulty due to the amount of material to be learned, refers to the number of elements that must be integrated simultaneously for the learning to progress.

7. Learning is more highly correlated with IQ when the *amount of time* for learning a given amount of material is fixed for all students. This condition becomes increasingly important to the extent that the other conditions listed above are enactive.

8. Learning is more highly correlated with IQ the more the learning material is *age-related*. Some things can be learned almost as easily by a nine-year-old child as by an eighteen-year-old. Such learning shows relatively little correlation with IQ. Other forms of learning, on the other hand, are facilitated by maturation and show a substantial correlation with age. The concept of 'learning readiness' is based on this fact. Tests of 'readiness', which predict rate of progress in certain kinds of learning, particularly reading and mathematics, are highly correlated with IQ.

9. Learning is more highly correlated with IQ at an *early stage* of learning something 'new' than is performance or gains later in the course of practice. That is, IQ is related more to rate of acquisition of new skills or knowledge rather than to rate of improvement or degree of proficiency at later stages of learning, assuming that new material and concepts have not been introduced at the intermediate stages. Practice makes a task less cognitively demanding and decreases its correlation with IQ. With practice the learner's performance becomes more or less automatic and hence less demanding of conscious effort and attention. For example, learning to read music is an intellectually demanding task for the beginner. But for an experienced musician it is an almost automatic process which makes little conscious demand on the higher mental processes. Individual differences in proficiency at this stage are scarcely related to IQ. Much the same thing is true of other skills such as typing, stenography, and Morse code sending and receiving.

It can be seen that all of the above listed conditions which influence the correlation between learning and IQ are highly characteristic of much of school learning. Hence the impression of teachers that IQ is an index of learning aptitude is quite justifiable. Under the above listed conditions of learning, the

low IQ child is indeed a 'slow learner' as compared with children of high IQ.

Very similar conditions pertain to the relation between memory or retention and IQ. When persons are equated in degree of original learning of simple material, their retention measured at a later time is only slightly if at all correlated with IQ. The retention of more complex learning, however, involves meaningfulness and the way in which the learner has transformed or encoded the material. This is related to the degree of the learner's understanding, the extent to which the learned material is linked into the learner's pre-existing associative and conceptual network, and the learner's capacity for conceptual reconstruction of the whole material from a few recollected principles. The more that these aspects of memory can play a part in the material to be learned and later recalled, the more that retention measures are correlated with IQ.

#### *Scholastic Achievement and g*

It is well known that scores on highly *g* loaded tests, such as the better standard IQ tests (whatever they may now be called), predict scholastic achievement better than any other single item of information we can obtain about a child after the age of five. At any one point in time, a good IQ test will correlate about .70 to .80 with objective measures of scholastic achievement. The reliability of both IQ and achievement measures can be considerably enhanced by obtaining repeated measures throughout the course of the child's schooling, thereby averaging out the fluctuations in performance that commonly occur between any two test scores obtained on the same individual tested at different times (Keys, 1928). Such cumulated IQ and cumulated achievement scores can correlate as much as .90 with each other. The important fact is that the *g* factor of any large battery of nonscholastic cognitive tests is the very same *g* factor as can be extracted from a comprehensive battery of scholastic achievement measures.

The identity of the *g* of intelligence tests and the *g* of scholastic achievement cannot be convincingly explained as being the result of children's learning the content and skills involved in IQ tests in school. IQ predicts achievement in school, not because the IQ measures merely an accretion of scholastic achievements, but because the nature of scholastic achievement in the long run is so highly characterized by, and is dependent upon,

the nature of *g*. This is true, at least, for scholastic achievement as we now know it, and it is virtually impossible to imagine there could ever be a much different state of affairs as regards the more academic aspects of the curriculum, especially at the advanced levels.

A statistical technique known as cross-lagged correlation analysis has recently been used to investigate the question of whether IQ causes achievement or vice versa (Crano, Kenny, & Campbell, 1972; Crano, 1974). The key question asked in this research is: Do individual differences in achievement measured in the earlier grades of school predict IQ in later grades more, or less, than early differences in achievement predict later differences in IQ? That is to say, is the correlation higher going from early IQ to later achievement, or from early achievement to later IQ? In a sample of 5495 elementary school pupils who were given both IQ and scholastic achievement tests in the fourth and sixth grades, the cross-lagged correlations were significantly higher going from IQ<sub>4</sub> to Achievement<sub>6</sub> than from Achievement<sub>4</sub> to IQ<sub>6</sub>, indicating greater causality from IQ to achievement than the reverse. The results were somewhat different in subsamples of middle and low socio-economic status children, and in subsamples of high and low IQ children, for various types of scholastic achievement. But overall, the predominant direction of causality, surprising as it may seem to many, is in the direction going from the more abstract and *g* loaded tests to the acquisition of the more specific and concrete skills. Crano *et al.* (1972) conclude:

Taken together, these results suggest that the more complex abstract abilities depend upon the acquisition of a *number* of diverse, concrete skills, but these concrete acquisitions, taken independently, do not operate causally to form more abstract, complex abilities. Apparently, the integration of a number of such skills is a necessary precondition to the generation of higher order abstract rules or schema. Such schema, in turn, operate as causal determinants in the acquisition of later concrete skills (p. 272).

It is this integrative capacity for the formation of higher order schema that characterizes *g* and facilitates the acquisition and retention of specific scholastic knowledge and skills.

In summary, I believe the dominant importance of *g* for scholastic learning is a central fact of nature which cannot be avoided. The enormously wide range of individual differences in



*g* poses tremendous dilemmas and challenges to enforced universal education. The ideal of equality of educational opportunity becomes problematic although, one hopes, still viable. But equality of performance, or anything even close to it, is a wholly unrealistic and hence unwarranted ideal. Future educational innovations that ignore the wide range of individual variation in *g*, or pretend that such variation can be overcome by improving educational interventions at some stage in children's development, are doomed to disappointment.

This is not to say, however, that we cannot manipulate the methods and age-pacing of instruction and the contents of the curriculum so as to reduce somewhat the correlation between *g* and scholastic achievement, when scholastic achievement is conceived more broadly as the acquisition of a variety of useful knowledge and skills, not just of an academic nature, for future citizens ranging widely in abilities. The dependence of school learning on *g* can be lessened to some extent by simplifying tasks, ensuring that prerequisite skills are mastered in some optimal sequence, providing all the time needed for mastery of each successive skill, and discovering motivational techniques that will keep students on the task.

Leading educational psychologists, such as Cronbach (1975) and Bereiter (1976), are now advocating that school may be made more rewarding for many youngsters in the lower half of the IQ distribution by designing instruction in such a way as to put less of a premium on *g* in scholastic learning. But I see little substantial evidence, as yet, that we really know how to go about doing this. Educational researchers are now just at the beginning of the research and development that must precede the attainment of this goal on a mass scale. I think we now see the direction in which we must move, but we are too inexperienced as yet to offer a definite and detailed prescription of this sort to the educators who run our schools. It is the main challenge, in the next decade, to educational psychologists, as well as to educational researchers in curriculum development and school policy planning, and to teacher training institutions. The survival of universal public education will depend to a large extent upon the successful outcome of our efforts in this direction.

#### *References*

Bereiter, C. (1976). IQ differences and social policy, in N. F. Ash-

- line, T. R. Pezzullo and C. I. Norris (eds), *Education, Inequality, and National Policy*. Lexington, Mass.: Lexington Books.
- Callaway, E. (1975). *Brain Electrical Potentials and Individual Psychological Differences*. New York: Grune & Stratton.
- Crano, W. D. (1974). Causal analyses of the effects of socioeconomic status and initial intellectual endowment on patterns of cognitive development and academic achievement, in D. R. Green (ed.), *The Aptitude-Achievement Distinction*, pp. 223-53. Del Monte Research Park, Monterey, California: CTB/McGraw-Hill.
- Crano, W. D., Kenny, D. A., and Campbell, D. T. (1972). Does intelligence cause achievement? A cross-lagged panel analysis. *Journal of Educational Psychology*. 63: 258-75.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*. 12: 671-84.
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*. 30: 116-27.
- Estes, W. K. (1970). *Learning Theory and Mental Development*. New York: Academic Press.
- Jensen, A. R., and Figueroa, R. A. (1975). Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. *Journal of Educational Psychology*. 67: 882-93.
- Jerison, H. J. (1973). *Evolution of the Brain and Intelligence*. New York: Academic Press.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*. 23: 187-200.
- Keys, N. (1928). *The Improvement of Measurement through Cumulative Testing*. New York: Bureau of Publications, Teachers College, Columbia University.
- Neisser, U. (1976). General, academic, and artificial intelligence, in L. B. Resnick (ed.), *The Nature of Intelligence*, pp. 135-44. Hillsdale, New Jersey: Erlbaum.
- Osler, S. F. and Trautman, G. E. (1961). Concept attainment: II. Effect of stimulus complexity upon concept attainment at two levels of intelligence. *Journal of Experimental Psychology*. 62: 9-13.
- Thomson, G. H. (1939). *The Factorial Analysis of Human Ability*. Boston: Houghton-Mifflin.
- Van Valen, L. (1974). Brain size and intelligence in man. *American Journal of Physical Anthropology*. 40: 417-23.
- Viaud, G. (1960). *Intelligence: Its Evolution and Forms*. London: Hutchinson.
- Voss, J. F. (1976). The nature of 'the nature of intelligence', in L. B. Resnick (ed.), *The Nature of Intelligence*, pp. 307-15. Hillsdale, New Jersey: Erlbaum.
- Wechsler, D. (1958). *The Measurement and Appraisal of Adult Intelligence*. 4th edition. Baltimore: Williams & Wilkins.
- Woodrow, H. (1938a). The effect of practice on groups of different initial ability. *Journal of Educational Psychology*. 29: 268-78.
- Woodrow, H. (1938b). The effect of practice on test intercorrelations. *Journal of Educational Psychology*. 29: 561-72.
- Woodrow, H. (1938c). The relation between abilities and improvement with practice. *Journal of Educational Psychology*. 29: 215-30.

- Woodrow, H. (1939a). Factors in improvement with practice. *Journal of Psychology*. 7: 55-70.
- Woodrow, H. (1939b). The application of factor analysis to problems of practice. *Journal of General Psychology*. 21: 457-60.
- Woodrow, H. (1939c). The relation of verbal ability to improvement with practice in verbal tests. *Journal of Educational Psychology*. 30: 179-86.
- Woodrow, H. (1940). Interrelations of measures of learning. *Journal of Psychology*. 10: 49-73.
- Woodrow, H. (1946). The ability to learn. *Psychological Review*. 53: 147-58.
- Zeaman, D. and House, Betty J. (1967). The relation of IQ and learning, in R. M. Gagné (ed.), *Learning and Individual Differences*, pp. 192-217. Columbus, Ohio: Merrill.