A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework

Terence R. Taylor
AProLAB cc, P O Box 489, Highlands North, 2037, Republic of South Africa

Received October 1993, accepted March 1994

In this article, the author reviews three traditions of cognitive assessment: the structural approach (based on factor analytic research), the information-processing approach, and the dynamic approach (based on learning theory). The theoretical background of each assessment approach is sketched and the advantages and disadvantages of each discussed. The author then processes an integrated two-factor cognitive theory which accommodates all three traditions. Using this theory as a basis, he proposes a design for a test battery which would be appropriate for use in multicultural context and would identify individuals likely to benefit from developmental opportunities.

In this article the author reviews three traditions in psychological measurement: the conventional psychometric approach, the information-processing approach, and the dynamic or learning potential approach. Each tradition has developed fairly independently of the others; consequently there is relatively little theory linking the three and relating their perspectives to one another. The author will propose a theory. He will also suggest the outline of a test battery which will include measures from all three traditions and which seems appropriate to address the measurement needs which are arising in South Africa at present. These needs are not being adequately accommodated by currently available instrumentation. Hence the article has both theoretical and practical aims.

Psychological tests have been in existence for just over a century. In 1882 Galton established a laboratory in London where, for a small fee, people could be tested on visual acuity and reaction-time tasks. Galton (1883) regarded simple sensori-motor reactions to be indicators of the more profound intellectual capabilities of humans. However his view did not prevail, and by the early part of the present century, the sensori-motor approach had been replaced by another approach which attempted to measure broad-based psychological constructs such as abilities.

Tests based on this approach are still in the ascendancy today; in fact almost all published tests for use in industry and many of those for educational use are designed to measure such constructs. In the past two or three decades, however, two new approaches have begun to emerge, one focusing on information processing, the other on learning and modifiability. Interestingly, the former resurrects Galton’s project to search for the underpinnings of real-world problem-solving skills in processing speed and acuity on simple tasks. The latter continues the long tradition of regarding learning as the critical factor underlying cognitive competence and the mastery of problems and challenges. Neither of these ‘new’ approaches has yet yielded much in the way of practical measurement instruments, such as those which could be used for selection or vocational guidance. However, this situation could change in the future, especially since conventional ability measurement instruments have come increasingly under fire in certain quarters for being biased or ‘unfair’ to certain groups or classes of testees. Furthermore, the prospect of testing for ‘potential’ or ‘learning potential’ seems to be an idea whose time has come in South
Africa, as this is seen as helping to address some of the inequalities of opportunity which have characterised South African society.

In the next three sections, the author will review the three approaches to assessment mentioned above, sketching their theoretical background and their advantages and disadvantages. Thereafter he will propose a theoretical framework, which accommodates all three traditions, and a test battery comprising measures drawn from all three.

The conventional or structural approach
This approach attempts to measure performance along dimensions which are purported to constitute the fundamental structure of the psychological domain in question. The domains include cognitive, personality, interests, and others. The scientific desideratum of parsimony is served, because many different types of behaviour or mental phenomena may be reduced to scores on a limited number of psychological dimensions.

Almost all research and theory development in this tradition has been of the individual differences type (Cronbach, 1957; Taylor, 1987a; Verster, 1982). In the cognitive domain, individual differences have been mainly measured through the power score—a summation of correct answers to problems of moderate to high difficulty. In practice, power scores are often contaminated with an admixture of speed variance; but in contrast to the information-processing tradition, individual differences research has paid relatively little attention to speed in its own right, especially at the item level.

The individual differences research of the structural approach has made extensive use of correlational and factor analytic techniques in resolving theoretical and empirical questions. Spearman (1927) was the first researcher to use the factor analytic method to identify underlying structures in the cognitive domain. He concluded that there is one general factor, g, that underlies performance on all cognitive tasks, and a number of specific factors which contribute to performance on certain activities. Thurstone (1938), using a somewhat different factor analytic technique, came to a different conclusion. In a monumental project to map the structure of intellect, he administered a huge battery of 57 tests to a sample of students. The tests were selected to tap a wide variety of skills, but for the most part were rather scholastic in nature. Thurstone’s analysis identified between seven and nine ‘Primary Mental Abilities’, such as numerical ability, verbal ability, memory, and spatial ability.

The English cognitive psychologist Burt (1949), following in a tradition inspired by Spearman (1927), concluded from his factor analysis research that cognition is hierarchically ordered, from sensory processes to relational reasoning, a g-like construct.

The factor analytic tradition was continued by a number of other cognitive researchers, including Vernon (1962), Guilford (1967), Cattell (1971), Snow and Lohman (1984), Horn (1986), and Ackerman (1988), each producing a different structure of intellect. The criticism may be directed against this tradition of research that it is too data-driven and rather theory-weak, that it relies too much on empirical findings which may be influenced by a number of fairly arbitrary circumstances or determinants, such as the nature of the sample, the choice of tests included in the study, the type of factor analytic procedure adopted, choice of criterion for acceptance of a factor, method of rotation, and so on. On the other hand, Guilford (1967), one of the more theoretically driven cognitive psychologists, has been criticized for ‘using’ factor analysis to confirm his theories (Horn & Knapp, 1973).

One of the best established theoretical positions is that of Cattell (1971), who offers a higher-order theory which distinguishes two forms of intelligence, fluid and crystallized. The former is a basic inherited capacity as developed by
an interaction with environmental characteristics which are found in any society, whereas the latter are specialized skills and knowledge promoted by and required in a given culture. Cattell’s culture-fair subtests are designed to measure fluid intelligence. Their medium of presentation is abstract-diagrammatic and the tasks involve such apparently universal activities such as series completion, pattern completion, classification, identification of conceptual odd-men-out, and identification of conceptual relationships. Evidence has been found in support of the cross-cultural validity of the fluid intelligence construct (e.g. Hakstian & Vandenberg, 1979). A further positive feature of the Cattell model is that it is amenable to dynamic, learning, or developmental interpretations. We shall return to it in the section on theory.

Most test construction tends to be based on a Thurstonian model. If one looks at the tests available in South Africa, for instance, there are numerous instruments for measuring verbal ability, numerical ability, mechanical ability, deductive reasoning etc. – constructs which are closely related to Thurstone’s (1938) so-called Primary Mental Abilities. Most, but not all, of these tests would be classified as measures of crystallized abilities, thus strongly affected by cultural influences and also schooling. There are some widely used tests such as the Figure Classification Test which are fluid intelligence measures (Werbeloff & Taylor, 1983); use of tests of this type for selection seems to be on the increase.

The defensibility of using measures of crystallized ability in a multicultural context such as obtains in South Africa is open to question, especially since certain cultural groups have had better quality schooling and more opportunities to develop specific skills. The first issue that has to be considered is the integrity of the constructs underlying the tests. Berry (1984) points out that many researchers tend to impose etics (universals) on members of cultures other than their own; these are often in reality emics (culturally specific constructs) of their own culture. Thus the etics are actually pseudo-universals. He suggests that the etic structures be used merely as a starting point and that these be modified systematically into emics; only through this process will true etics (if these exist) ultimately be uncovered.

In psychometrics, the issue of the integrity of constructs and the meaning of test scores in different cultures is dealt with under the rubric of comparability or bias. Jensen (1980), Van der Vijver and Poortinga (1982), Taylor (1987b), and others distinguish various types of comparability. Taylor and Boeyens (1991) conclude that there are basically three main varieties of comparability; construct, score, and predictive. According to Fridja and Johoda (1966), the fundamental concern when making comparisons is dimensional or construct comparability. ‘In the long run’, Irvine and Caroll (1980) comment, ‘construct validation is the only thing’ (p.212). The psychological constructs that conventional tests measure tend to be quite broad and are often not well defined. R.J. Sternberg (1977) points out that different tests which purportedly measure the same construct often tap only partially overlapping skills and processes. He uses this argument in support of his contention that the factorial construct is too gross for a viable theory of cognition and should be abandoned in favour of a more fine-grained analysis of cognitive functioning based on processes.

The techniques and procedures necessary to do methodologically adequate comparability research are quite elaborate and only a few thorough studies have been undertaken in South Africa (e.g. Holburn, 1992; Owen, 1989). Because of the expertise required and expense of such studies it is improbable that they will ever be widely conducted. Even the most conscientious of test users normally do little more than check means and reliabilities in different groups, and possibly compare the magnitudes of correlations of test scores with criteria. When test scores are not comparable across groups, there is a strong likelihood
of unfair treatment of a particular group (for instance in the awarding of scarce opportunities on the basis of test results).

However, even when no evidence is found of incomparability of test scores across groups, the question arises as to the defensibility of using scores which may maintain the status quo regarding privilege. If privilege (especially privilege conferred by a socio-political system which has been condemned almost universally as morally wrong) has endowed certain groups with better specific skills or abilities, is it acceptable to continue to award scarce opportunities (such as desirable work positions and places in educational or training programmes) on the basis of measures of these skills and abilities? Or should we abandon these measures and look for others which may be more compatible with a process of redressing the wrongs of the past, such as measures of potential for development?

Nevertheless, as Hunter and Schmidt (1981) and others have pointed out, conventional tests do have economic utility; and at present well proven alternatives are not available. If these tests are to continue to be used in a context of affirmative action, new ways of using the scores will have to be adopted, without totally sacrificing the economic utility of the instruments. Certain options are available, including the use of separate norms before compiling candidates into a consolidated top-down list as proposed by the U.S. National Academy of Sciences (Hartigan & Wigdor, 1989), or the application of Cascio, Outtz, Zedeck and Goldstein’s (1991) sliding bands methodology.

The information processing approach
Information processing began to establish itself as the basis of cognitive psychology in the 1960’s. The emergence of computers in this period as a tool to assist human problem solving strongly stimulated this development. The computer and associated concepts (hardware and software, temporary and permanent data storage, processing bottlenecks, etc.) began to be seen as fruitful modes of understanding human perception, thinking, and problem solving. Human functioning was interpreted in a highly cognitive-rational term, and man came to be seen primarily as an information processor. Psychometrics and cognition had over the years been closely associated, but this was a relatively barren time for theory. Information processing seemed to offer cognitive psychology a more scientific basis and the promise of theoretical development.

Three main information processing paradigms were in place by the end of the 1960’s: Those of Hick (1952), Posner, Boies, Eichelman, and Taylor (1969). A fourth was added in the 1970’s: the inspection time paradigm of Vickers, Nettlebeck, and Willson (1972 ) and Nettlebeck and Lally (1976). Whereas most conventional tests are power tests (i.e. evaluate the person on the basis of the correctness of response), the first three of the above paradigms are chronometric in nature (the primary measure is response latency). Computer presentation is necessary for the Posner, S. Sternberg, and Nettlebeck (inspection time) paradigms; in the Hick paradigm, the computer is normally used as a recording device, but presentation is on a separate apparatus. In contrast to conventional or factorially grounded tests, information processing tests are fine-grained: they tap one or a few specified cognition activities or approaches. In the Hick paradigm (as elaborated by Jensen, 1988) the subject presses a button to extinguish a light from a set of one, two, four, or eight lights, the sets imposing processing demands corresponding to zero, one, two and three bits of information respectively. Scores include decision times, movement times, and the slope of the regression line drawn through the data for the four conditions (which expresses the number of milliseconds needed to encode one bit of information). The Posner paradigm assesses long term memory access latency. Subjects are required to make same-different decisions regarding pairs of letters under two conditions – physical identity and name (conceptual) identity.
The mean difference between response latencies in the two conditions constitutes long-term memory access time.

The S. Sternberg paradigm measures short-term memory scanning. Subjects are presented briefly with between two and seven digits or letters and are then presented with a probe stimulus which may or may not have been in the original set. The primary measure is the slope of the regression line drawn through the latencies for the different string lengths, which represents the amount of time required to scan one element in short-term memory. Levy (1992) defines inspection time (IT) as the minimum exposure time for an observer to identify reliably a highly evident feature of a stimulus or display; hence it refers to stimulus apprehension as opposed to stimulus encoding which is measured in the Hick paradigm. In the classic Nettlebeck or IT paradigm, two vertical lines, one appreciably longer than the other, are presented for a brief period, then the stimulus is masked. The subject is required to indicate which line is longer. The on-time of the stimulus is progressively shortened until the point is reached where the subject can no longer respond correctly. These four paradigms are in the tradition of the limited capacity theories of cognitive competence, as articulated by Broadbent (1958), Eysenck (1986), Jensen (1982), Vernon (1987), and others.

The information processing tradition was further developed in a 1972 landmark book by Newell and Simon, Human problem solving. The authors introduced the concept of the production system, which they regarded as the basic unit of cognitive thinking. The production system is not unlike a line of computer programming, such as an ‘if-then’ statement, and the solution of a problem involves the assembly of a number of such statements.

Newell and Simon’s (1972) computer program conception of human problem solving inspired a whole branch of modern information processing theory known as the cognitive components approach (Pellegrino & Glaser, 1979). Two of the main exponents of this approach are R.J. Sternberg (e.g. Sternberg, 1982, 1988) and Carroll (1976). The aim of this approach is to uncover the components of cognitive thought, such as the components that make up analogical problem solving (R.J. Sternberg, 1977).

The cognitive components approach has identified a large array of processes, which does not accord well with science’s predilection for parsimony. R.J. Sternberg (1986) proposed a strategy for identifying core processes or components, and most information processing theorists have tried to secure parsimony by developing the idea of meta- or executive processes, of which there are apparently a limited number. However, these higher level processes are, unlike the first level processes, not free of cultural influences (Verster, 1986). In fact they are probably as conditioned by culture as conventional test items. Hence an attempt to assess metaprocesses would open the researcher to many of the problems suffered by conventional testing regarding cultural bias.

A second main tradition in information processing was developed by Hunt (e.g. Hunt, 1985). This is the cognitive correlates approach. Rather than attempt an exhaustive breakdown of the elements of a task, the cognitive correlates approach attempts to find the critical processes which underlie performance on a given conventional measure (such as verbal ability). Thus the model accepts that conventional test scores have some value as it treats these as a kind of criterion against which processes are evaluated.

Measures derived from four paradigms mentioned earlier have been used in the cognitive correlates approach, with the
Ravens Progressive Matrices (a fluid intelligence measure) or other global measure used as criterion. In a number of studies, quite substantial correlations have been found (e.g. Jensen, 1982, 1988; Kirby & Nettlebeck, 1989; Kranzler, 1992; Kranzler & Jensen, 1989; Larson & Alderton, 1990; Neubauer, 1990; Saccuzzo, Larson, & Rimland, 1986; Snow, Marshalek, & Lohman, 1076; Vernon, 1983, 1986, 1989, 1990; Vernon & Jensen, 1984).

After reviewing a number of information processing variables (in particular those derived from the paradigms mentioned above) and the relationship of measures of these variables to cognitive abilities, Sen (1991) concludes: ‘Thus it appears that the rudiments of mental abilities of human beings can be tapped by measuring information receiving, processing and retrieval speeds that would be relatively free from the scope of availability of knowledge and other environmental variables’ (p.213). Vernon (1990) comes to a similar conclusion.

Apart from the measures derived from the four paradigms discussed above, certain other processes have shown promise as cognitive correlates. In particular, working memory has been shown to correlate strongly with fluid intelligence measures (Baddeley, 1986; Larson & Saccuzzo, 1989). Baddeley defines working memory as short-term memory (STM) ‘in action’ rather than the mere temporary storage of information. It consists of two aspects, storage and processing; the processing draws on data stored in STM, and once having manipulated it, returns it to STM.

If one employs Sen’s (1991) division of cognitive processing into receiving, central processing, and long-term storage, then the four paradigms plus working memory may be associated with specific aspects of these functions. The IT and Hick paradigms tap two sequential aspects of information receiving: apprehension and encoding. Working memory is related to ‘active STM’ in the central processing function, and the S. Sternberg paradigm to a more passive (storage) aspect of STM.

Finally, the Posner paradigm taps long-term memory access time. This model (see Figure 1) guided the construction of tests, as described in Nell and Taylor (1993), to be used for neuropsychological diagnostic purposes.

The information processing approach promises a fruitful new era in assessment. Its constructs are much more precisely defined and the measures are so simple that the problems of cultural bias are likely to be relatively minor. However this cannot be said with certainty, as very little, if any, cross-cultural research has been done on information-processing tests. As mentioned above, many studies have shown that certain information processing variables correlate strongly with general measures of ability, and optimally weighted combinations can share more than 50 % of variance with fluid intelligence tests. To date, though, no research has been done on the effectiveness of an information processing battery for predicting practical criteria such as work performance or educational success. The correlations with Raven and other measures are promising, but such tests in themselves are not interesting as criteria.

A possible disadvantage of the information processing approach is that computer administration is essential; this limits the number of people that can be tested at a time. But nevertheless, Nell and Taylor (1993) have developed tests which may be administered to individuals at all educational levels (including illiterates). For instance, in the Posner paradigm they replaced upper case and lower case letters
with pictures of standing and sitting animals.

**The learning or dynamic approach**

Resnick and Neches (1984) point out that there has long existed in psychology two meanings of the concept of intelligence. One tradition views intelligence (or a set of intelligences or abilities) as a stable and possibly innate endowment which different people have in different quantities. The other tradition identifies intelligence with the capacity of humans to adapt to circumstantial demands – in other words to learn to function effectively in their environment. The former tradition has been more guilty of imposing etics on non-Western peoples and thus has had to cope with both the polemic surrounding the frequent finding of lower scores in such groups, and the negative sentiment towards testing which has developed in many non-Western cultures. The latter approach to intelligence may be more appropriate as a basis for cross-cultural assessment, a point which Biesheuvel (1972) has made strongly in conceptualizing intelligence as, fundamentally, the capacity for adaptation. The dynamic approach addresses itself to the assessment of adaptation to novel tasks, as evidenced in mastery or increased speed and accuracy as a result of repeated exposure, instruction, examples, or hints. Bransford, Delclos, Vye, Burns and Hasselbring (1987) claim that dynamic tests may also provide information on learning processes (as opposed to learning products, which are assessed by conventional tests) and diagnostic information to guide the design of intervention techniques.

A major theoretical underpinning for an adaptability interpretation of intelligence was provided by the Russian psychologist Vygotsky (1978), whose work *Mind in Society* was originally published in 1926. Vygotsky took the view that the acquisition of cognitive competence is a social phenomenon. Adults and older peers pass on to children the knowledge and skills required in their culture. In effect they play the role of metacognitive or executive processes in assisting children to conceptualise and interpret the world and solve problems; later these processes are internalised by the recipients of this mediation who then make the intelligence or skill their own. Vygotsky acknowledged that individuals differ in their capacity to benefit from mediated learning experiences, and defined a concept, the zone of proximal development, to indicate the difference between unassisted performance and performance attained through mediation. Conventional tests in contrast measure only unassisted performance and thus do not reflect developmental capacity. Feuerstein’s (1980) theoretical position on learning potential and his method of assessing it is strongly in the Vygotsky school.

Useful as Vygotsky’s (1978) perspective is, it does not thoroughly explicate the relationship between intelligence, as observed on intelligence tests, and learning ability. Many earlier studies showed us a modest relationship between learning and intelligence (Estes, 1982; Woodrow, 1938). These early studies, though, involved very simple learning tasks; only fairly recently have psychologists started studying learning in more complex knowledge domains (Taylor, 1987a). Marshalek, Lohman, and Snow (1983) have found that a radex map describes the relationship of abilities effectively, with the more complex and g-loaded abilities clustering around the centre (with, fluid intelligence, g right at the centre) and specific abilities and low complexity speeded activities falling further out in verbal, spatial, and numerical sectors of the radex. Snow, Kyllonen and Marshalek (1984) produce evidence that the topography of learning tasks may be similar. Learning in complex learning tasks correlates more highly with g-type tests than learning in simple tasks. However, in their data, even the highest correlations did not exceed about 0.5 or 0.6. Ackerman (1988), using a learning task of moderate complexity, found that general intellectual ability correlated about 0.4 with performance on the first session, but less than 0.1 by the 12th session.
(no correlations with gain scores were supplied).

Ferguson’s (1954, 1956) theory of learning and ability contributes to an understanding of these findings. Ferguson holds that abilities are attributes of behaviour which have attained a certain stability in the adult through a process of learning that has reached, or is approaching, an asymptote. Cultural factors prescribe what shall be learned; consequently different patterns of abilities emerge in different cultures. Abilities emerge through a process of transfer, in which existing abilities contribute to the development of new abilities (e.g. computer programming skill may emerge as a result of transfer from verbal, numerical, and reasoning abilities). The positive correlation between abilities and learning scores is attributable to this process. Hunt (1980) points out that fluid intelligence is a function of the cognitive strategies available to the person; presumably many of those which play a role in a given ability also assist in the process of transfer. R.J. Sternberg (1984) makes a distinction between knowledge acquisition processes and problem solving processes; however, many processes may play a role in both activities. Furthermore, Sternberg sees coping with novelty and automatization – learning concepts – as fundamental to intelligence.

Hebb (1949), P.E. Vernon (1962), and more recently, Carlson and Wiedl (1992) presents a model which assists in the understanding of the inter-relationships among learning, ability and actual performance on conventional tests.

These authors distinguish three types of intelligence, A, B and C. Intelligence A is the individuals biological potential. This potential may be reflected in information processing speed, STM span, working memory, LTM accessing latency, and other variables mentioned in the previous section. However, information processing measures are incomplete indices of biological intelligence and contaminated with error or unique variance. Intelligence B is the individual’s ‘ actual ‘ intelligence or problem-solving capacity, although this cannot be measured directly. Finally, intelligence C is the person’s intelligence as measured by tests.

Intelligence A is apparently determined and not modifiable. Carlson and Wiedl (1992) believe that intelligence B is affected by environmental and experiential factors such as advantagement, the quality and quantity of past mediation by caregivers, etc. This type of intelligence is modifiable, but only by very extensive cognitive mediation, probably running to hundreds of hours. Intelligence C is a function of intelligence B, but is also affected by the cultural content of the test (which can be a source of bias), familiarity with the test material, and also testing conditions.

Thus there appear to be two sources of suboptimal performance; that arising from past genotype-environment interactions and that arising from the assessment procedure used. Equivalently there are two types of learning potential. Although most students of dynamic assessment do not use the above classification, it appears that some approaches focus more on the potential (which this author shall call potential type I) which can be actualized through extensive mediation and the teaching of thinking skills, while others concentrate more on the potential ( which the author shall call type II ) which is revealed in performance on tests that measure learning, either as a result of repeated exposure and instruction or prompts. It is potential type II which has been shown by Snow, Kyllonen and Marshalek (1984) and others to have moderate correlations with static test scores (these scores may be derived from an external intelligence test or initial performance on a learning test). The relationship between test score and type II potential may be accounted for by citing the relationship between intelligences B and C as defined in the Hebb-Vernon model, or by using Ferguson’s (1956) transfer concept, or by taking Sternberg’s (1984) view that intelligence largely comprises the capacity to respond effectively to novelty and to automatize.
These alternative explanations are probably mutually supportive rather than antagonistic. It must be remembered, however, that the relationship between static and dynamic cognitive performance (learning or type II potential) is only moderate. Hence the latter gives additional information on the individual, which seems to be particularly valuable in cross-cultural assessment exercises and where testees differ in advantagement or past opportunity.

Several approaches have been devised for assessing learning potential. Laughton (1990) identifies three main approaches: those of Budoff, Feuerstein and Campione / Brown. Budoff’s (1968, 1974, 1987) preferred approach is to pre-test subjects (children in various IQ brackets ranging down to retarded) on a block design or Ravens type task, and then to retest after simple practice and task-specific training. Feuerstein (1980), in contrast, assesses pretest to post-test improvement in response to elaborate mediation designed to develop thinking skills. More than other methods, Feuerstein’s attempts to assess what has been designated here as learning potential type I. Campione, Brown, Ferrara, Jones and Steinberg (1985) concentrate on learning potential as transfer and assess it as an inverse function of the number of hints required. Butterfield and his co-workers have adopted a similar transfer-centred approach (e.g. Ferretti & Butterfield, 1992). Dynamic measures have been found to have considerable variance comparable to those of static tests. In addition, certain individuals who produce poor showings on static tests produce considerably better performances on dynamic measures (Budoff, 1987; Laughton, 1990). From a conceptual and theoretical point of view, dynamic tests have several advantages. They seem to be less susceptible to the effects of cultural bias (Gupta & Coxhead, 1988). To a large extent, learning tests equate the amount of relevant learning experience across people and even cultures. This is achieved by selecting learning tasks which are initially unfamiliar to all (Hegarty, 1988). Even if there is some variation in initial familiarity with the material, the improvement score may be less contaminated with bias than a conventional test; the reason for this is that the bias may be roughly a constant factor throughout the learning process, and will thus be effectively cancelled out when the improvement score is calculated. If the gain score is expressed as a normed percentage improvement over the individuals initial performance, such a score may even be used as a method of implementing affirmative action in a way that does not require the tester to record the testees cultural background, race, or any other biological information. The reason for this is that it is ‘easier’ to obtain a high percentage improvement over an initial score depressed by disadvantagement than it is to achieve a similar improvement on a high initial score. Another approach which may achieve a similar effect is to measure learning potential as the difference between the individual’s actual final score and the predicted final score as determined by the regression of the final score on the initial score. This index reflects the extent to which the person underperformed or overperformed relative to others in the sample with a similar initial score. Embretson (1987) points out the psychometric superiority of this residual score over a simple difference score.

Certain questions remain concerning learning potential, however. As yet there is relatively little solid empirical evidence in support of the predictive and construct validity of learning potential tests, although there are some promising results (Campione & Brown, 1987; Embretson, 1992). Test-train-test formats may be somewhat problematical because subjects may become confused by inconsistencies between self-generated strategies developed in the initial test and the strategies taught in the training session (Boeyens, 1989). Gain scores may be unreliable unless special effort is made to optimise the reliability of pre-test and post-test measures and maximize the size and variance of different scores. Finally, there is an apparent paradox concerning the interpretability of a difference score as its reliability is inversely related to the
correlation between pre-test and post-test; hence a reliable difference score can be seen as the result of comparing ‘apples and pears’ (Cronbach & Furby, 1970). However, traditional psychometrics is not fully appropriate for learning tests, and Embretson (1992) has developed a model which largely resolves the apparent paradox mentioned above. The model treats certain components of variance as systematic which would traditionally be regarded as error.

A theoretical position accommodating the three traditions, and a suggested test battery based on it

In the previous section the author began drawing together theoretical threads from the ability and learning literature. In this section, a theory integrating the ability, information processing, and learning traditions in cognitive psychology will be more fully spelled out. This theory forms the basis of a suggested test battery incorporating measurement approaches from all three traditions, which is intended to be appropriate for application in multicultural contexts. Certain information-processing theorists claim that processing capacity and speed is the entire foundation of intelligence and problem solving (e.g. Jensen, 1982; Vernon, 1986, 1987). On the other hand, Spearman (1927), Cattell (1971), Snow, Kyllonen and Marshalek (1984), Horn (1986, 1989) and others conceptualise intelligence as being founded on a broader construct or set of constructs, although some of these authors accept the importance of speed as an aspect of cognitive competence. Spearman defines intelligence as the education of relations and correlates. To a large extent, Cattell’s fluid intelligence is accommodated by the first half of this definition – the education of relations, or the inference of rules and formation of concepts. Cattell’s crystallised intelligence involves rule application, which is related to the second half of Spearman’s definition, the education of correlates. Guttman (1965), representing test scores on a circumplex, makes a similar distinction in identifying tests at the core as analytic or rule inferring and those remote from the centre as rule applying in nature. Snow, Kyllonen and Marshalek, using a similar method to represent scores, have also found fluid intelligence measures at the core.

The present author takes the position that information-processing speed and capacity are not the complete foundation of intelligence, although these form one or two main fundamentals. The other is the potential to infer concepts or rules and thus to think abstractly. This potentiality is not independent of processing speed and capacity – the two factors are related; but processing variables do not fully account for the individual’s potential to think abstractly.

Both factors are biologically or genetically endowed; they set unalterable upper limits on performance. Jointly they may be thought of as Hebb’s (1949) intelligence A. As such they are not directly measurable, although empirical information processing measures may provide a reasonably accurate reflection of fundamental processing efficiency; conventional ability test scores, however, are further removed from the fundamental potential to form concepts and think abstractly.

The genotype intelligence A is manifest in the phenotype as intelligence B. This intelligence emerges as the product of learning and other interactions with the environment. The potentiality to think abstractly and form concepts develops as fluid intelligence. It consists of a set of general cognitive tools and strategies for application to novel problems. Ackerman’s (1988) cylindrical elaboration of Snow, Kyllonen and Marshalek’s (1984) circular cognitive model best accommodates the theoretical position presented here. Competencies near the core of the cylinder are more general and closely related to the genotypic potential. Progressively larger concentric rings contain skills which are even more specific and remote from fundamental potential. These rings also reflect the process of transfer in development and learning (Ferguson, 1954; 1956). The
vertical dimension of the cylinder is a speed dimension. Starting from the top, each successive ‘slice’ through the cylinder contains skills which are of an increasingly speeded nature.

As development proceeds, skills and knowledge accumulated in prior learning have a growing impact on the emergence of new skills. Kyllonen and Christal (1989) in fact distinguish four sources of individual differences in learning: Knowledge and skills (the enablers) and processing speed and processing capacity (the mediators). The enablers thus would be expected to play a critical role in the transfer required for the development of later emerging skills (those on the periphery of the Ackerman cognitive cylinder). The mediators would be expected to impact on the development of the speeded skills in the lower reaches of the cylinder.

Several authors (e.g. Anderson, 1983; Shiffrin & Schneider, 1977) have distinguished three phases of learning: conceptual understanding of the task, compilation of execution procedures, and the automatization of processing. The boundary between the second and third phases is not distinct and the second phase merges into the third. It seems likely that the abstract thinking factor will play the major role in the first phase, whereas the processing speed and capacity factor will play an increasingly important role as learning progresses to the phase of automatization. As automatization progresses, skills shift outwards and downwards in the cognitive cylinder.

Measures in the core of the cylinder provide the best estimate of the individual’s fundamental potentiality. Those skills at the periphery are the product of a longer process of learning and transfer. The impact of developmental opportunities is more apparent at the periphery than at the core. If developmental opportunities are equalized, individual differences in skills at the periphery will be largely a function of personal endowment of the two fundamental cognitive factors. In South Africa, of course, there has been great disparity in the distribution of opportunity. Thus there are fairness problems in using tests of specific skill to make selection decisions, although in some applications their use may be justified.

At this point it is appropriate to make a distinction between learning performance and learning potential. The author conceptualises them as learning analogues of crystallized and fluid intelligence. Learning performance is demonstrated when an individual acquires specialized skill through transfer from other fairly specialized skills or abilities. The more elaborated and developed a person’s skills repertoire, the more effectively and swiftly he or she is likely to acquire the new skill.

Learning potential is shown when a person comes to grips with a novel learning task involving unfamiliar stimulus material; in this case previously developed specific skills are of relatively little help to him or her, and the learner has to use very general transfer and skill acquisition strategies; similarly general strategies are needed to solve abstract problems of the sort encountered in fluid intelligence tests.

In the light of the comments made above regarding the issue of fairness and the consequent importance of assessing potential, it is clear that in the learning domain it is potential rather than achievement that should be assessed. The critical learning aspects to measure appear to be the implementation of general transfer strategies in dealing with novel material, and the early stages of proceduralization and automatization. Although it may be unwise to attempt a precise one-to-one mapping of the two basic intellectual factors on to these learning variables, it would appear that transfer and abstract thinking capacity are related and processing efficiency and automatization are related.

The purpose of the exposition presented above is to establish a theoretical basis to guide test design and the compilation of a battery for the purpose of fair assessment in a multicultural situation, where the emphasis is on evaluating potential rather than existing specific skills. We shall now discuss the matter of the tests.
It is proposed that a battery to achieve the purpose stated above should consist of tests measuring in the following four domains: fluid intelligence; information processing; transfer in tasks designed to assess learning potential; and automatization in such tasks. The question may be asked, ‘Why are four domains being assessed when the theory posits just two factors?’ It must be remembered that measures of fluid intelligence and information processing speed and capacity are not direct or pure indices of intelligence A or biological potential. These are phenomena which have already been conditioned to some extent by interaction with the environment. Hence it seems wise to include other measures which the theory identifies as assessing learning potential directly, by virtue of actually giving the testee a learning task to perform. Performance on each of the measures will include an admixture of what might be called ‘experiential variance’ or particular learned skills and strategies, and this admixture will differ from measure to measure. The four types of measures will thus provide a broader and better rounded picture of the individual’s potential for development. It should also be borne in mind that the learning potential assessment instruments envisaged for inclusion will be of the kind designed to measure what the author has previously called learning potential type II, or potential to benefit from repeated exposure or factual feedback (type I in contrast is the potential to benefit from mediation intended to develop thinking skills). Thus the potential being assessed is of a particular type which appears to be more relevant to practical selection and replacement exercises.

What information processing measures should be included in the battery? The information processing model presented in Figure 1 can be used as a guide in the selection of tests. Measures from each of the identified processing phases could be included. If the battery is to be of a pencil-and-paper variety, however, the choices become more restricted. Nevertheless one or more speed tests could be included. Copious practice should be given to minimize any cultural difference in stimulus familiarity. Verster (1982), Horn (1986), and others have identified different types of speed tasks, varying from psychomotor to conceptual. Verster (1982; 1986) has also pointed out the relevance of assessing accuracy of information processing. Working memory is an important index of information processing capacity, possibly the most important index. The author has devised a way of measuring this construct using paper-and-pencil methods.

Conclusion

This article has two aims, one theoretical and one practical. In addressing the former, the author evaluated the existing measurement traditions and the theory underlying them; then he proposed an integrated theory which makes it possible to compile a test battery in a more orderly and reasoned way than is normally done. The practical aim involved finding a way of going about assessment that addresses the needs confronting South Africa as it struggles to create a more just society - in particular, finding a way to identify those who have potential for development, even though they may have gaps or limitations in their skill repertoire due to past disadvantage.

The two-factor theory here attempts to integrate the information processing theory and constructs of the experimental tradition in psychology with the trait or ability theory and constructs of the individual differences tradition. (Cronbach, 1957, has pointed to the need to combine these traditions in order to make progress in cognitive psychology.) Learning was incorporated into the model in the form of the constructs of transfer and automatization; these help explain the development of the cylindrical structure of skills and abilities as proposed by Ackerman (1988).

It was concluded that the sorts of measures which should be included in a battery to fulfil the requirement of assessing potential rather than crystallized skill should be from the core of the cylindrical...
cognitive space. Stimulus material should therefore not be drawn from any specific skill or ability area and in particular should avoid the numerical and verbal domains as much as possible, as competence in these is very much a function of quality and quantity of schooling. Abstract-diagrammatic material such as is encountered in the Ravens and Figure Classification tests should rather be used.

Employers and educationalists may justifiably want to know about a candidate’s verbal or numerical skills, for these skills may have a bearing on whether the person will be able to cope in the job or training course. In such cases, test materials very close to the actual skill demands should be administered and the scores interpreted as pure skill scores rather than measures of intellectual capability or potential.

If we are to address the inequalities of the past in South Africa, employers and educationalists will have to place more emphasis on potential rather than skill or specific ability and will have to be more prepared to give those with high potential the opportunities to develop specific skills through educational, training and other development programs. The battery proposed above could be used as one of the tools for the identification of such people.

References
cultural context (pp. 105-145). Cambridge: Cambridge University Press.