INTRODUCTION:
ISSUES IN THE MALLEABILITY OF INTELLIGENCE

Welko Tomic and Johannes Kingma

Intelligence is considered to be a very important and to some extent even a delicate human quality. When it comes to intelligent functioning, no one is eager to have their intelligence given a negative rating. After reviewing the literature on intelligence from the beginning of the century to the present day, we are faced with an enormous number of definitions of the subject. First, it is reasonable to assume that all definitions of intelligence are shaped largely by the culture, place, and time in which they originated.

In their exploration of the cultural underpinnings of cognitive abilities, Sternberg and Grigorenko (1997) state that it is negligent to deny the role of culture in the meaning and definition of intelligence. They give many examples from different cultures showing that the way intelligence and cognitive abilities are defined is very much determined by the particular cultural context. As a consequence, a common definition of intelligence across all cultures is an unreachable ideal.

Second, the fact that there are so many different definitions of intelligence means that we all know how hard it is to reach agreement about the essence of the subject, even within a particular cultural environment—that is, Western culture. For the time being, many theorists must be satisfied with the description
that "intelligence is what intelligence tests measure." It is rather questionable to claim that intelligence tests are measuring something while at the same time stating that what they measure can only be specified by reference to the test contents.

A compromise such as this, singularly lacking in theoretical underpinnings, is open to accusations of superficiality. Gigerenzer (1997) puts this pragmatic view into the following words:

A great error that has swamped research on intelligence tests is to start with no theory but seductive everyday concepts such as social skill and creative intelligence, then go on to design tests to measure these vague concepts, and to pray to heaven that these tests miraculously transform loose thinking into precise mechanisms of intelligence. Despite prayers that we backed up by statistics, some 90 years of factor analyzing and correlating IQ tests has not noticeably increased our understanding of the mechanisms of human intelligence (p. 284).

The possibility of measuring a concept definitely depends on our being able to formulate precisely what it is that is being measured. There has never been a satisfactory definition of intelligence, however.

It appears that theorists who confined themselves to a view of intelligence that compressed it to a single dimension—that is, classic g theory—and who developed intelligence tests are not interested in answering such highly interesting questions as "How do you make people more intelligent?" "What causes a child's intelligence to increase or decrease?" and "How can we reach durable results?"

There is nothing daring about considering intelligence to be a very important human characteristic, especially in developed societies. After all, a person's level of intelligence has consequences for his or her career and, therefore, social status. Our very survival depends on utilizing that particular characteristic.

According to Howe (1997), the predominant thought put forward about intelligence is quite easily understandable. Every human being possesses a quality of intelligence that can be considered the driving force which enables a few of us, but certainly not all of humanity, to acquire the necessary cognitive skills to flourish in our society. The predominant view of this mental characteristic is that it is inherited for the most part and has a biological basis, with the rather pessimistic consequence that it is largely unchangeable.

The mainstream view concerning this important issue is that variance in intelligence scores is produced largely by nature. Perkins (1995) calls this broad style in theorizing about intelligence the "neural view of intelligence." The danger of this view is its attractive oversimplification—that is, that differences in intelligence must come from either nature or nurture but not from both. Even well-educated people are tempted to draw the wrong conclusion. Wrangham and Peterson (1997) call this kind of reasoning "Galton's false dichotomy," or "Galton's error."

Bouchard avoids this trap. He has been involved in studies of twins for years and has examined the effects of adoption on a child's IQ, among other things. He does not feel forced to choose between an intelligence generated by nature or one gen-
erated by nurture. After many years of enthusiastic research, he has come to the following conclusion:

The results of twin studies, I believe, refute both biological and environmental determinism. They do not negate the effect of the environment on behavior, nor do they overglorify the role of genes (p. 57).

The predominant view of intelligence described above—that is, general factor or g, which is said to pervade performance on all tasks requiring intelligence—is rather depressing indeed, and still leads many theorists to stipulate that a person’s level of intelligence is largely unchangeable.

On further consideration, the advocates of this view of intelligence take a narrow approach to intelligence, for in most cases they use intelligence test scores as a dependent variable. If intervention programs do not induce large gains in intelligence test scores, then the conclusion is drawn that intelligence is unchangeable. Nevertheless, it should not be forgotten that most intervention studies were not specifically designed to change IQ.

However, this is no longer a valid observation given the evidence produced by several research sources on the changeability or plasticity of human intelligence. This view of intelligence, called reflective (Perkins, 1995), refers to our knowledge about thinking and reflective self-guidance and what they contribute to intelligence.

Researchers who adopt elements from the Genevan school and action psychology approach when designing interventions can explain why the above-mentioned narrow view of intelligence is unproductive. To bring about a change in intelligence—or, to use an expression employed by the Genevan school, cognitive representation—one must provide a number of elements in instruction that teach intelligence (Tomic & Kingma, 1996).

One significant factor is that interventions must be designed in a way that brings about a change in the cognitive structure of the child. An essential requirement is that skills taught during training should be transferred to concept areas in which the children have not been trained. It is precisely the range of the transfer which indicates whether intervention has brought about a fundamental change in a child’s cognitive structure. If the child is unable to generalize, then intervention has in all likelihood generated an isolated structure, so that the child is unable to use the principle it has learned to solve other problems. This means that there has to be evidence of transfer to conceptual areas in which the child has not been trained. For example, if the child has been trained to solve problems involving conservation of length, then after training, that child should also be able to solve other related problems, for example, those involving seriation and transitivity of length.

It is important that a training effect should be durable as well. If durability is not warranted, the effects will diminish over time. To assess durability, the child should be re-tested on the problems taught during intervention a few weeks to a
few months later. If the child is still able to solve these problems correctly even after a substantial period of time, for instance, four months after the first posttest, then training may be assumed to have been successful.

Only if the criteria regarding range of transfer (far-far transfer preferred) and durability are fulfilled can one determine whether intervention has induced a change in the child’s cognitive structure (Tomic, Kingma, & Tenvergert, 1993; Tomic & Kingma, 1996).

Besides durability and range of transfer, the length of intervention is another and rather decisive factor when it comes to increasing intelligence. A very important observation is that short intervention programs do not lead to a change in the cognitive structure and, therefore, do not change the level of intelligence positively. Long-term interventions do successfully influence the level of intelligence (Kingma & Tomic, 1997). If this problem is not seriously tackled, the majority of children will drop back down to the level of their untrained classmates. We must prevent the phenomenon of fade-out. However durable, large, and stable the effects, they are not permanent unless there is ample opportunity to use and practice them. Without this, no cognitive change will be brought about.

The fact that some interventions have produced large gains may be attributed to the duration of training, the degree of systematization of the applied instruction method, and the range of transfer induced by intervention programs. We agree with Howe (1997) that it is reasonable to suppose that the periods of special intervention needed to produce major improvements would be at least as long as the amounts of time necessary to reach high degrees of expertise in narrower areas of competence such as music, chess, and sports. However, most of the programs that have been evaluated on their effects on children’s IQs have been of much shorter duration.

In contrast, the interventions that increase intelligence substantially are based on various theoretically founded methods of instruction. A theoretically derived methodology is evidently not a panacea for increasing intelligence. On the contrary, more general aspects, present in every successful method, are decisive. In each successful method, the curriculum is constructed systematically. For example, suitable curriculum sequences are determined, the partial actions which underlie the training concepts are taught systematically, and regular repetition of these partial actions in various situations are spread out over a period of several weeks or more. These are features of successful methods that can also be found in Adey and Shayer’s intervention program (1994) and in Perkins and Grotzer’s review (1997).

Concerning the changeability of intelligence, there is also evidence from studies conducted nationwide. There are well-known examples of research studies that have looked at evidence suggesting that in certain nations, the average IQ level may have changed from one generation to another. Flynn (1987) conducted a series of investigations aimed at finding out what would have happened had the rescaling of intelligence tests not taken place. Supposing that the same tests of
approximately 50 years ago were still used now, what would have happened to people’s IQ scores? Flynn’s studies revealed that there have been unexpectedly large changes, with IQs constantly increasing, sometimes considerably. On average, students nowadays score more than 30 percentile ranks higher than the group which produced the most recent generation of government leaders and captains of industry.

If the intelligence tests given throughout the U.S. are measuring any of the factors the general public includes in its definition of “smart,” we are now smarter than we have ever been before (Berliner, 1993, p. 634).

Flynn’s explanation is that it is quite impossible for genetic mechanisms to bring about noticeable changes between one generation and the next. His view is that the causes must be entirely environmental, and he is skeptical about the present state of our knowledge of these causes (Flynn, 1994). We believe a number of factors are likely to be involved, including the increasing complexity of life, better communications, and more exposure to information. Nutritional improvements may also have had an influence. Improved schooling is almost certainly one reason for the changes.

Moreover, there are an increasing number of recent studies which give us the empirically obtained results of educational intervention programs particularly designed to provide training for young children and adolescents and to enhance children’s thinking. These intervention programs all contained features emphasized by the reflective intelligence view. According to Howe (1997), almost all the evidence from training studies aimed at the changeability of intelligence shows a substantial increase not only in intelligence test scores but in school subject achievement as well. An illustrative example for many of these interventions is the CASE program designed by Adiey and Shayer (1994), which demonstrated that secondary school students could become more intelligent learners. The conclusion is that it can be fruitful to teach children to think better.

Howe (1997) reviews American, German, Israeli, South African, and Swedish studies to examine the possible effects on intelligence test scores of variations in the amount of schooling that children receive. These studies provide unequivocal evidence that intelligence can be raised (Brody, 1992).

The conclusion is that the level of intelligence can in fact be increased, regardless of the theoretical underpinnings. But lasting change will only be effected if the environment in which the newly acquired “skills” are to be exercised has also changed more or less permanently. It is important to note that recently designed interventions aimed at increasing intelligence utilize elements recommended by the Genevan and action psychology school—the reflective intelligence view. Consequences and reviews of some important elements in successful intervention programs are found in Adiey and Shayer (1994), Perkins and Grotzer (1997), and Tomic and Kingma (1996), among others.
The question to be answered remains: why we should undertake so many efforts to increase intelligence. The question is important in both theoretical and practical terms. The answer to this question can be broken down into four parts. First, many writers on intelligence have pointed out that people with low IQs do not flourish in our society (Howe, 1997). It is likely that they will stay unemployed. Raising their level of intellectual functioning can help them to prosper better in society. Second, the greater the mastery of cognitive skills, the more children benefit from regular education. Third, human beings are persistent in their curiosity to know more. The question as to whether—and if so, to what extent—intelligence can be increased is simply an expression of this curiosity. Fourth, the point is not only to acquire knowledge, but knowledge that can be applied. The application of this knowledge must also be generally useful. It is, hence, precisely the results of well-designed interventions that can help the educational psychologist and the educationalist to develop instructional technology further.

The practical importance of modifying intelligence is more obvious but independent of the theoretical importance. The degree to which intelligence is thought to be a biological characteristic in theoretical models has no bearing on the practical importance of its alterability. All interventions must be environmental. The development of a technology of education depends on exactly understanding how to achieve gains in intelligence (Detterman, 1982, p. vii).

Before dealing with each chapter in detail, we will provide an aerial photograph.

**PLAN OF THE BOOK**

We now turn to the plan of the book. In chapter 2, Flynn states that IQ tests have many virtues. Thanks to Spearman and Jensen, they have engendered the first serious theory of intelligence, and that theory must not be dismissed because of problems about defining intelligence. The Spearman-Jensen theory solves these in a way parallel to the hard sciences. IQ tests are an important instrument for investigating group differences, particularly within industrial societies and within generations. However, school personnel deal primarily with individuals rather than groups and here, IQ tests often play a counterproductive role. Massive IQ gains over time show that IQ test performance includes a potent nonintelligence factor. Teachers should not use them to label students as "intelligent" or "dull" or to make placement decisions. IQ gains also show that there is no body of evidence in favor of any particular IQ criterion of mental retardation. Psychologists should use behavioural criteria to make such assessments. Differential IQ gains between tests encourage false diagnoses about whether students are overperforming or underperforming and whether they suffer from learning disabilities or reading disorders. The literature on the relationship between IQ and female underperformance in mathematics, the educational potential of the elderly, and the counselling of high school graduates contains doubtful assumptions. Teachers would be better off if
they were ignorant of their students’ IQs, psychologists would probably be better off, counsellors perhaps better off.

According to Mayer and Mitchell (chapter 3), three frameworks have traditionally been employed to integrate personality with intelligence: a differential psychology approach, a theory-by-theory approach, and a research topic-by-topic approach. They introduce an alternative systems approach because each of these earlier ones have limitations as integrative procedures. The systems approach they developed identifies eight components of intelligence that also belong to the larger personality system: energy, inputs, processors, arenas, outputs, knowledge base, enhancers, and constraints. Mayer and Mitchell show how several touchstone theories of intelligence (e.g., Spearman’s g, Gardner’s theory of multiple intelligences) can be organized in such a component-by-component basis. They further apply the framework to their own work on hot (emotional) processing. Finally, the authors reexamine the intelligence components (i.e., energy, inputs, processors, etc.) as components of personality with specific attention to how they may define, and be defined by, the personality system.

Campbell and Nabors (chapter 4) examined longitudinal predictors of Verbal and Performance IQ scores in a sample of students from low-income families (98% African American). Predictors representing personal characteristics of the students and their mothers, interactions between them, and variations in the context of development were identified. Study participants were randomly assigned to receive five years of early childhood educational intervention in a childcare setting or to a preschool control group; half of each preschool group was assigned to receive three years of educational intervention in the primary grades. Wechsler Verbal and Performance IQ scores were obtained on five occasions between the ages of 5 and 15 years. Predictors were arrayed in four developmental stages: infancy, preschool, elementary school, and adolescence.

Maternal IQ, a consistently powerful correlate of child IQ, and sociodemographic scores (High Risk Index) were covaried in the authors’ analyses in order to better understand the relationship of other factors to child IQ. Given these covariates, Verbal IQ appeared to be more heavily influenced by the developmental context, especially by early childhood educational intervention and the quality of the preschool-age home environment. Neither school-age intervention nor the quality of the school-age home environment contributed significantly to the maintenance of Verbal IQ. Although students with preschool treatment maintained an advantage in Verbal IQ over preschool controls, both groups declined in Verbal IQ after school entry.

In contrast, Performance IQ scores remained relatively stable. Although mean levels of Performance IQ were not as strongly affected by contextual factors as those of Verbal IQ, individuals who had school-age treatment and also had better-quality home environments were more likely to maintain Performance IQ across time.
In chapter 5, Van Peet reviews the potentiality theory of intelligence by A.D. de Groot. A key concept in this theory is potentiality: the level of achievement that is within an individual’s capacity to attain if his or her environment is optimal. Other key concepts are: learning, (life) intelligence, specific capability, abstract-verbal intelligence, potentiality at age A, talent, and ceiling. These concepts are defined and their relationships are described.

Potentiality theory argues that many people are capable of, or can learn to, estimate their own level of intelligence and how close they are to their intelligence ceiling, and that they are also able to make these estimates about family members and close friends.

To validate these suppositions, parents of large families were asked to estimate their own intelligence and those of their children. These estimates were compared to the results of the general intelligence tests administered to both the parents and their children. The parents were also asked two questions about intelligence ceilings.

The parents were able to estimate the intelligence of their children better than their own intelligence. Furthermore, their estimates of their children’s intelligence correlated significantly with the children’s results on the intelligence tests.

A significant majority of parents (73% to 83%) said that their estimates of their children’s intelligence and of their intelligence ceiling were equal. A smaller percentage of parents (66% to 68%) said that the same was true of their children’s schooling and their intelligence ceiling. Assuming the most favorable life circumstances, then, parents were more likely to say that their children should have had more schooling than that they could have been more intelligent.

Not all results supported the potentiality theory, however. The last section discusses the meaning of the validation study results.

In chapter 6, Carver relates reading achievement to intelligence and memory capacity. A causal model of reading achievement was used as a framework for investigating how intelligence and memory capacity are related to reading. Study I involved 62 college students who were administered tests that measured constructs in the causal model as well as general intelligence and memory capacity. Simple correlations and a hierarchical regression analysis provided evidence that measures of general intelligence and memory capacity did not have an important causal impact upon reading achievement. Study II involved 36 students in grades 3 to 10 who were given the same tests as the college students in Study I; the results of Study II were interpreted exactly the same as in Study I. Study III involved 49 students in grades 3 to 11, who were given exactly the same tests as Studies I and II except that two memory capacity tests administered by the computer were modified slightly so as to be more appropriate for younger students; the results were again interpreted exactly as for Study I. The overall results were interpreted as providing support for the causal model. Subsequently, the model was expanded to include teaching and learning experiences as well as three aptitude factors that purportedly have a major impact upon reading achievement—verbal knowledge
aptitude $g_v$), decoding aptitude ($g_p$), and cognitive speed aptitude ($g_s$). General intelligence ($g$) seems to be primarily composed of fluid intelligence ($G_f$), along with other subsidiary aptitudes such as $g_v$, $g_p$, and $g_s$. There was no compelling evidence that fluid intelligence has a major causal impact upon reading achievement, and since $g$ is composed primarily of $G_f$, it should not be considered as a primary causal factor either.

In chapter 7, Clements and Nastasi discuss nontraditional assessment of intelligence and the development of intelligence in specially designed educational environments. The theory of intelligence they employ takes an information-processing perspective. First, the authors discuss two approaches to measuring executive-level processing of young students—naturalistic observation and dynamic assessment—that address shortcomings of traditional approaches to intellectual assessment. Studies designed to assess the validity of these approaches are briefly reviewed. Results provide support for the use of these techniques as valid measures of general executive-level functioning, but mixed support for the delineation of specific metacomponents. They also suggest the criticality of considering contextual factors in assessment and of examining processes as well as products of thinking. Second, the development of these intellectual processes within collaborative educational computer environments is discussed. Using both approaches to the measure of the metacomponents, their research program demonstrated that unique characteristics of the computer environment enhances students’ development of intellectual functioning. The research program also focused on the effects of collaboration within different educational computer environments, using naturalistic observations of students’ interactions during problem solving and post-treatment measures of perceived competence and teacher ratings of social competence. These studies demonstrated that certain social and motivational processes mediate computer environments’ effects on higher-order thinking. Based on their research, we discuss implications for assessment of intelligence and development of educational programs to foster higher-order thinking.

According to Soraci, Carlin, and Chechile (chapter 8), the particular structural aspects of visual arrays importantly influence both preattentional and attentional processes. Perceptual variables such as symmetry and contiguity are involved in the initial detection of interstimulus relations, and can be manipulated in ways that facilitate attentional functioning. Because basic properties of multi-stimulus visual arrays (e.g., similarity-dissimilarity and novelty-familiarity) are intrinsically relational and are embedded in ubiquitous environmental contexts, the ability to detect interstimulus relationships is a critical factor in mediating attentional functioning. Soraci, Carlin, and Chechile provide converging evidence suggesting that one locus of intelligence-related differences may be a differential sensitivity to relational information. The present research involves an examination of the structure of inter- and intrastimulus relations, and how detection of such relations influences cognitive processing across a range of paradigms. Experimental tasks discussed explore encoding, rule-acquisition, pattern recognition, and verbally
mediated relational structure. Their "front-end," perceptually based approach of "guiding" attention is demonstrated to rapidly facilitate the effective detection and discrimination of relevant stimulus relations across a range of individual and task differences. The authors claim that a delineation of visual structure, and its effects on cognitive processes, can have important implications with respect to advancing our understanding of the relationship between intellectual functioning and attention and in providing a direction for ameliorative procedures. The results are thus informative for issues of an applied and theoretical nature in the study of attention, cognition, and intelligence.

In chapter 9, Resing reviews intelligence and learning potential. She describes the latest developments in research on intelligence and intelligence tests, focusing on learning potential tests. Before describing these developments, Resing examines the concept of intelligence. In addition, she discusses the relationship between intelligence, learning, and the ability to learn.

After defining intelligence in terms of the ability to learn, the chapter discusses research on learning potential, the measurement of learning potential, and the diagnostic use of learning potential tests in testing practice.

Instruments that measure learning potential appear to be quite effective, in general. The predictive validity of learning potential test scores is better than that of traditional intelligence tests. Additionally, learning potential test scores contribute independently to the prediction of school success. Learning potential and intelligence appear to be inter-related but at the same time, in terms of their contents, different concepts, both fairly closely related to different external criteria. Each concept can be said to complement the other in the diagnostic intelligence examination. The learning potential test collects extra quantitative and qualitative information about the intellectual potential of the child in a very standardized way. This testing method focuses more specifically on the child's learning abilities, both the diagnostic and the solution processes, than do standard intelligence measurement procedures.

In chapter 10, Klauer describes how the close relationship between inductive reasoning and (fluid) intelligence has been established by factor analytic research and corroborated by cognitivistic process analytical research. A prescriptive theory of inductive reasoning specifies the processes enabling one to solve problems requiring inductive reasoning. According to this prescriptive theory, three training programs have been developed. A number of experimental training studies shows that the programs improve fluid intelligence. Trained subjects outperform untrained by about half a standard deviation and the effects lasted for several months. However, the programs improved the learning of school-type subject matter even more than fluid intelligence. Theoretical implications of these results concerning the theory of inductive reasoning and the relationship between inductive reasoning, fluid intelligence, and learning are discussed. Suggestions for future research are also given.
The issue of accelerating intelligence development through an inductive reasoning training is described by Tomic and Kingma in chapter 11. The study investigated the effects of an inductive reasoning training program for teaching children. The experiment assessed the effects of training and the range of transfer of the training achieved. The subjects were 47 third grade primary school children of average ability. Children from the training condition \((N = 23)\) received a three-week course of training (10 half-hour sessions). The results demonstrated a significant, positive training effect on childrens’ performance of inductive reasoning tasks. A near-far transfer was observed—that is, children were able to solve tasks in which they had not been trained. These effects persisted for four months. Far-far transfer was also observed, because the children were able to solve arithmetic problems which related to inductive reasoning, in which they had received no training. Implications for training children within the context of regular schools and the range of transfer are discussed.

The effects of test preparation are investigated by Van der Molen, te Nijenhuis, and Keen (chapter 12). The first goal of this study was to investigate the effects of reading a book concerning intelligence tests and the effects of a specific test-training program on numerical and verbal intelligence items. The second goal was to investigate to what extent the acquisition of test-specific problem-solving strategies affects the ability to solve items on different but comparable tests (transfer). In the experimental design, two factors were included: practice (pretest or no pretest) and (level of) preparation (none, book, or training), so there were six conditions. Each condition consisted of about 26 subjects, who had been randomly assigned to one of the conditions. The results showed a strong effect for preparation, especially on the numerical intelligence test and, to a lesser degree, on the verbal intelligence test. No practice or pretest effects were found. Positive transfer was demonstrated for the numerical test. The results for the verbal test were less clear.

REFERENCES


