

## **The Illusion of Culture-free Intelligence Testing**

**Michael Cole**

For almost as long as there have been IQ tests, there have been psychologists who believe that it is possible to construct "culture free" tests (Jensen, 1980). The desire for such tests springs directly out of the purposes for which tests of general intellectual ability were constructed in the first place: to provide a valid, objective, and socially unbiased measure of intellectual ability. Our society, founded upon the principle that all people are created equal, has never lived easily with the recognition of enormous de facto social inequality. We need a rationale for such inequality and our traditions strongly bias us to seek the causes of inequality in properties of the individual, not society. At the same time, we realize that social and economic conditions, by shaping people's experiences, can be the causes of individual intellectual differences, as well as their consequences. Can't we find universals in human experience and construct a test on this basis?

What would be more ideal than a psychological test that could measure intellectual potential independently of the specific experience provided by sociocultural and economic circumstance? Such a test would provide an excellent tool for insuring that unfortunate social circumstances would not prevent the identification of intellectual potential. Some psychologists have claimed not only that such tests are possible in principle, but have been applied in practice (Hernstein & Murray, 1994).

In this chapter, I will argue that the notion of culture-free intelligence is a contradiction in terms. I begin by reviewing the historical background of efforts to understand the relation between culture and thought that formed the scholarly background against which IQ testing came into being. After summarizing briefly the strategy developed by the pioneers of IQ testing, I will present a "thought experiment" to help clarify the issues and some empirical evidence from research which has sought to approximate the conditions of the thought experiment. I close by offering some comments on how to think about culture and IQ testing given the impossibility of a culture-free test of intellectual ability.

### **Beliefs About Culture and Cognitive Ability in the 19th Century**

The several decades just preceding this century provide a useful starting point from which to trace theories of culture and cognitive development, because it was during this period that both anthropology and psychology, the disciplines assigned the roles of studying culture and cognition, took shape as disciplines. Until this time there was no distinctive body of methods for the study of the "humane sciences," nor had scholars with different theories been institutionally divided into separate disciplines, each with its own methods of studying human nature.

Obvious differences in technological achievement between peoples living in different parts of the world were common knowledge among European scholars. Their theorizing about sources of these differences had produced rather general acceptance of the notion that it is possible to study the history of humanity by a study of contemporary peoples at different "levels of progress." The "father of anthropology," E. B. Tylor, summarized (in which he called a "mythic fashion") the general course of culture that most of his fellow scholars would have adhered to:

We may fancy ourselves looking on Civilization, as in personal figure she traverses the world; we see her lingering or resting by the way, and often deviating into paths that bring her toiling back to where she had passed by long ago; but direct or devious, her path lies forward, and if now and then she tries a few backward steps, her walk soon falls into a helpless stumbling. It is not according to her nature, her feet were not made to plant

### ***Cole: The Illusion of Culture-free Intelligence Testing***

uncertain steps behind her, for both in her forward view and in her onward gait she is of truly human type. (Taylor, 1958, p. 69)

Tylor made another assumption that also won general acceptance: there is an intimate connection between socio-cultural progress and mental progress. "...the condition of culture among various societies of mankind," he wrote, "...is a subject apt for the study of laws of human thought and action" (Tylor, 1874, p. 1). He even adopted the notion of a "mental culture," which he expected to be high or low depending upon the other conditions of culture with which it was associated.

Herbert Spencer, writing at about the same time, shared Tylor's belief in the fusion of mental and sociocultural progress. He argued that the circumstances under which the earliest human beings lived provided only a limited number and variety of experiences. "Consequently," he argued, "there can be no considerable exercise of faculties which take cognizance of the general truths displayed throughout many special truths." (Spencer, 1886, p. 521)

Spencer invites us to consider the most extreme case; suppose that only one experience were repeated over and over again, such that this single event comprised all of the person's experiences. In this case, as he put it, "the power of representation is limited to reproduction of the experience" in the mind. There isn't anything else to think about! Next we can imagine that life consists of two experiences, thus allowing at least elementary comparison. Three experiences add to the elementary comparisons, and elementary generalizations that we make on the basis on our limited (three) experiences. We can keep adding experience to our hypothetical culture until we arrive at the rich variety of experiences that characterizes our lives. It follows from this line of reasoning that generalizations, the "general truths" attainable by people, will be more numerous and more powerful the greater one's experience. Since cultures provide experience, and some cultures (Spencer claimed) provide a greater diversity of experience than others, a neat bond between cultural progress and mental progress is cemented.

Although such evolutionary schemes seemed almost transparently obvious in the enthusiasm following publication of Darwin's *Origin of Species*, events toward the close of the nineteenth century proved that there could be a great deal of disagreement about the relation between culture and thought, despite the compelling story constructed by scholars like Tylor and Spencer. One set of disagreements arose when researchers started to examine more closely the data used to support conclusions about relations between cultures, especially claims for historical or evolutionary sequences. A different set of disagreements arose around conflicting claims about mental processes.

The source of these disagreements concerning sociocultural sequences can be found in Tylor's own work. The main criteria he used for judging the stage of a culture were the sophistication of industrial arts (including manufacturing techniques for metal tools, agricultural practices) and "the extent of scientific knowledge, the definitions of moral principles, the conditions of religious belief and ceremony, the degree of social and political organization, and so forth." However, in Tylor's words, "If not only knowledge and art, but at the same time moral and political excellence, be taken into consideration it becomes more difficult to scale societies from lower to higher stages of culture" (Tylor, 1874, p. 29).

This undeveloped theme in Tylor's work was taken up by Franz Boas, who submitted the cultural evolution position to a devastating critique at the close of the nineteenth century. On the basis of his own ethnographic work, Boas (1911) concluded that a great deal of the evidence apparently supportive of evolutionary schemes was so deeply flawed that no clear conclusions ranking one culture above another could be accepted. Boas did more than show the flaws in

### ***Cole: The Illusion of Culture-free Intelligence Testing***

evolutionists' data and arguments concerning culture; he also delighted in showing that examples of "primitive mind" produced as part of this argument were based on misunderstandings.

Consider the following example from Boas's classic, *The Mind of Primitive Man*, which repeats evidence used by Spencer to make some generalizations about properties of primitive mind:

In his description of the natives of the west coast of Vancouver Island, Sproat says, "The native mind, to an educated man, seems generally to be asleep....On his attention being fully aroused, he often shows much quickness in reply and ingenuity in argument. But a short conversation wearies him, particularly if questions are asked that require efforts of thought or memory on his part. The mind of the savage then appears to rock to and fro out of mere weakness." (Boas, 1911, p. 111)

Spencer's text goes on to cite a number of similar anecdotes corroborating this point. But Boas produces an anecdote of his own.

I happen to know through personal contact the tribes mentioned by Sproat. The questions put by the traveller seem mostly trifling to the Indian, and he naturally soon tires of a conversation carried on in a foreign language, and one in which he finds nothing to interest him. As a matter of fact, the interest of these natives can easily be raised to a high pitch, and I have often been the one who was wearied out first. Neither does the management of their intricate system of exchange prove mental inertness in matters which concern them. Without mnemonic aids to speak of, they plan the systematic distribution of their property in such a manner as to increase their wealth and social position. These plans require great foresight and constant application. (Boas, 1911, p. 128)

Thus, Boas tells us that the entire scheme was wrong. Cultures cannot be ranked using evolutionary age as a basis for comparison, and "mind" cannot be seen as rank in developmental age. (Boas also demonstrates the total hopelessness of deducing cultural differences from any differences, real or imagined, in genetic makeup.)

Finally, and very importantly, Boas was a leader in a subtle, but essential change in anthropological thinking about the concept of culture itself. Educated in Germany, Boas had begun his career imbued with the romantic concept of "Kultur," the expression of the highest attainments of human experience, as expressed in the arts, music, literature, and science. This is the conception of culture that allowed Tylor to talk about "the conditions of culture among various societies." Tylor, like Boas as a young man, conceived of culture as something groups and individuals had more or less of. It was a singular noun: one talked of higher or lower culture, not more or fewer cultures. By the same route that led him to deny the basis for ranking cultures in terms of a hypothetical, evolutionary sequence, Boas arrived at the idea that different societies create different "designs for living," each representing a uniquely adapted fit between their past and their present circumstances in the world. This point of view is central to contemporary anthropology, and it clearly has to be taken into account if we want to rank the intellectual achievements (levels of mental development) of people growing up with different cultural experiences. It renders simple more/less comparisons of cultures difficult and restricted, with parallel effects on our inferences about mind.

### **Enter Psychology**

The birth of psychology is usually dated back to 1879, when Wilhelm Wundt officially opened an experimental laboratory in Leipzig. The exact date is not important, because several laboratories opened almost simultaneously in different industrialized countries. But the reasons for these laboratory openings are important for understanding the problems of understanding the relation between culture and intelligence.

### ***Cole: The Illusion of Culture-free Intelligence Testing***

Boas's critique of developmental theories, whether of mind or culture, produced controversy in both domains of inquiry. Boas earned the enmity of anthropologists who believed his criticisms of their general theories unjust; they sought to rescue the more general theories, criticizing Boas and his students for "historical particularism" (Harris, 1968). Psychologists were people who took up the other half of the equation, the problem of specifying mental mechanisms.

The major difficulty facing psychologists was to devise methods for specifying pretty exactly what happens when an individual when some sort of "thinking" is going on. Competing claims were evaluated by constructing settings to control as exactly as possible the kinds of events a person experienced and to record the kinds of responses these experiences evoked. Since the presumed processes were not observable (they were, as we say, "psychological"), psychologists spent a great deal of time and ingenuity devising ways to pin down what these nonobservable processes might be.

The rapidly growing ability to control electricity and to build precision machinery was exploited to the fullest; the early psychology laboratories were marvels of inventions. Their instruments allowed psychologists to present people carefully controlled lights and tones for carefully controlled intervals and to measure precisely the time it took to respond. In their search for ways to make mind observable, they used electrophysiological devices to record internal, organic functioning. The discipline of "psychophysics" advanced appreciably in its quest to relate psychological phenomena of an elementary order (discriminating tones, judging hues). There were even hopes of uncovering a "cognitive algebra" by carefully comparing reaction times to stimuli of various complexities arranged to reveal steps in the thought process.

The activities of psychologists and anthropologists soon contrasted very dramatically. Psychologist brought people into the laboratory where behavior could be constrained, events controlled, and mind made visible. Whereas the anthropologists continued to concentrate on gathering data that would permit them firm statements about historical relations between cultures, scholars who came to identify themselves as psychologists concentrated on resolving arguments about thinking such as those illustrated in the passage quoted by Boas. Just as anthropology evolved careful field techniques to disambiguate competing claims about "culture," psychologists developed the laboratory experiment as a way to test competing claims about "mind."

There occurred, in effect, a division of labor in the "humane sciences," a division that was primarily a matter of scientific strategy in the beginning: progress required some concentrated work on specialized subtopics. The overall task remained the same for everyone: how do human beings come to be the way they are?

### **Enter Testing**

Despite an increasing gulf between scholars who called themselves psychologists and those who called themselves anthropologists, it was not long before those two areas of inquiry were brought together again. At the end of the nineteenth century, Francis Galton, in England, set out to test hypotheses about mental differences among people, using the newly devised psychological techniques. His concern was not differences between people growing up in different cultures. Rather, he studied people growing up in different families. Significantly, his tests were theoretically motivated; he believed that speed of mental processing was central to intelligence so he created tests to measure the rapid processing of elementary signals. Galton succeeded in finding differences among Englishmen on such tests as simple reaction time to a pure tone, but he did not succeed in relating these "psychological test" differences to human characteristics of greater interest to him such as scientific excellence or musical ability. Galton's tests, based on an oversimplified model of the human mind and the highly controlled procedures adopted from the laboratory appropriate to testing his theory, were not taken up by society. However, in creating an

### ***Cole: The Illusion of Culture-free Intelligence Testing***

early precursor of existing IQ tests, Galton did begin the development of the statistical techniques that would be necessary to show how test differences correlate with interesting behavioral differences.

Galton did all of his work in England, but other Englishmen, including W. H. R. Rivers (1901), traveled to the Torres Strait northeast of Australia, to see if psychological tests could be used to settle disputes over cultural differences in cognition. Rivers was in some senses an antique. He was both anthropologist and psychologist, which meant that he considered both the evidence of his tests and evidence provided by observation of the people he went to study when he made statements about culture and thought. His conclusions were consistent with Galton's data on individual differences; natives differed from each other on such simple tasks as their ability to detect a gap in a line, or their recognition of colors. But there were no impressive differences between the natives of the Torres Strait and Englishmen.

It would appear on the basis of this evidence that there are no cultural differences in thinking, at least no differences consistent with the pattern proposed by Tylor, Spencer, and others. However, it could be (and was) argued, that the important ways in which cultural differences cause mental differences were not even tested by Rivers and his associates. After all, Galton had found no relation between responses to his psychological tests and other presumed indicators of intelligence. Why would anyone, then, expect cultural differences in elementary sensory abilities since these depended on a physiological mechanisms common to all people? What seemed necessary were tests of higher psychological processes that could be used to compare people from different cultures or different people in the same culture.

This distinction between elementary and higher processes pinpoints a weakness in the basic foundations of experimental psychology, a weakness acknowledged by Wundt, its founder. It is impossible, Wundt believed, to study higher psychological functions in experiments because such functions always depend on prior, culturally organized, experience that differs from one individual and society to another, and these differences undermine the purity of the experiment. Wundt believed that scientists should use ethnological evidence and folklore if they want to discover the properties of the mind that get constructed on the basis of the elementary processes that he studied in the laboratory.

Wundt's doubts about the experimental method were not accepted because they put psychologists in a difficult bind. Psychology had been founded on the principle that carefully controlled environments are required to make legitimate statements about how the mind works. But a great many of the questions about how the mind works that interested psychologists and anthropologists alike clearly refer to "higher" psychological processes such as logical reasoning and inference. When Wundt gave up on the idea that such processes could be studied in the laboratory, he was, it seemed, robbing psychology of most of its interesting subject matter. For psychologists, the inability to study higher psychological processes in the laboratory meant that they could not be studied at all.

### **Binet's Strategy**

The major push for a way to measure mental ability apart from culturally-conditioned experience came from a source seemingly remote from theoretical disputes among anthropologists about the possibility of reconstructing history through a study of contemporary cultural variation or issues of cross-cultural experimentation among psychologists. Early in this century, Alfred Binet was asked to deal with a practical, social problem. With the growth of public education in France, there was a growing problem of school failure, or at least severe school under achievement. It seemed not only that some children learned more slowly than others, but that some children, who otherwise appeared perfectly normal, did not seem to benefit much from

### ***Cole: The Illusion of Culture-free Intelligence Testing***

instruction at all. Binet and his colleagues were asked to see if they could find a way to identify slow-learning children at an early stage in their education. If such identification were possible, special education could be provided them, and the remaining children could be more efficiently taught.

The subsequent history of IQ testing has been described too frequently to bear repetition here, but a sketch of the basic strategy of research is necessary as background to understand just how deeply IQ tests are embedded in cultural experience.

To begin with, early test makers had to decide what to test for. The decision seemed straightforward. They wanted to test people's ability to perform the kinds of tasks that are required by schools. They observed classrooms, looked at textbooks, talked to teachers, and used their intuitions to arrive at some idea of the many different kinds of knowledge and skills that children are eventually expected to master in school.

What Binet and his colleagues found was not easy to describe briefly, as anyone who has looked into a classroom can quickly testify (and all of us have done so, or we would not be reading these words). There was a very obvious need to understand graphic symbols, such as alphabets and number systems. So recognition of these symbols was tested. But mastery of the rudiments of these symbols was not enough. Children were also expected to manipulate these symbols to store and retrieve vast amounts of information, to rearrange this information according to the demands of the moment, and to use the information to solve a great variety of problems that had never arisen before in the experience of the individual pupil. Thus, children's abilities to remember and carry out sequences of movements, to define words, to construct plausible event sequences from jumbled picture sequences, and to recognize the missing element in graphic designs were tested (along with many other components of school-based problems).

It was also obvious that to master more and more esoteric applications of the basic knowledge contained in alpha-numeric writing systems, pupils had to learn to master their own behavior. They had not only to engage in a variety of "mental activities" directed at processing information; they also had to gain control over their own attention, applying it not according to the whim of the moment, but according to the whim of the teacher and the demands of the text.

It was clearly impossible to arrive at a single sample of all the kinds of thinking required by "the" school. Not only was there too much going on in any one classroom to make this feasible; it was equally clear that the school required different abilities from children of different ages. Binet realized that estimates of "basic aptitude" for this range of material would depend upon how much the child had learned about the specific content before he or she arrived at school, but he felt knowing the child's current abilities would be useful to teachers anyway.

In the face of these difficulties, Binet decided to construct a sample of school-like tasks appropriate for each year of education, starting with elementary grades, and reaching into higher levels of the curriculum. He would have liked to sample so that all the essential activities were included in his test and that tasks at one level of difficulty would be stepping stones to tasks at the next higher level. But because no firmly based theory of higher based psychological functions existed, Binet had to rely on a combination of his own common sense and a logical analysis of tasks that different classrooms seem to require (for example, you have to be able to remember three random digits before you can remember four; you have to know the alphabet before you can read). He also hit on the handy strategy of letting the children themselves tell him when an item selected for the test was appropriate.

Beginning with a large set of possible test questions, Binet hunted for items that half the children at a given age level could solve. An "average" child would then be the one who solved

### ***Cole: The Illusion of Culture-free Intelligence Testing***

problems appropriate to his or her age level. Keeping items that discriminated between children of different ages (as well as items that seemed to sample the activities demanded of kids in their classrooms), he arrived, with help from his colleagues, at the first important prototype of the modern IQ test.

Of course a great deal of work has gone into the construction of tests since Binet's early efforts, but the underlying logic has remained pretty much the same: sample the kinds of activities demanded by the culture (in the form of problems it requires that its children master in school) and compare children's performance to see how many of these activities they have mastered. Children who have mastered far less than we would expect given a comparable sample of kids their own age are those who will need extra help if they are to reach the level expected by the culture.

This strategy is perfectly reasonable, so long as we stay within the framework that generated the item selection procedures in the first place. However, much to the disapproval of Binet, people found new uses for the tests of school-based knowledge that carried with them the seeds of the current disputes over IQ testing.

Although Binet specifically warned against the procedure, his test and tests like it began to be used as a measure of an overall aptitude for solving problems in general, rather than samples of problem-solving ability and knowledge in particular. Those engaged in such extrapolations acknowledged that in principle it is important to make certain that everyone given the test has an equal opportunity to learn the material that the test demands. But in practice there was no way to guarantee this essential prerequisite for making comparative judgments about basic abilities.

These are important issues in thinking about applications of IQ testing, and they are extensively discussed in the psychological literature. However, it is not until we back up and examine the possible significance of Binet's work in the light of anthropological scholarship that we can see just how limited an enterprise IQ testing was at the beginning, and how restricted it remains today.

### **A Thought Experiment in Test Construction**

A good starting point for this reexamination is to think about what sort of activity Binet would have engaged in if he had been a member of a cultural group vastly different from his own. As a sort of "thought experiment" let us suppose that a "West African" Binet has taken an interest in the kinds of knowledge and skills that a child growing up in his part of the world would need to master as an adult. To make the thought experiment somewhat concrete, I will do my supposing about the tribal groups inhabiting the interior of Liberia, principally the Kpelle people, among whom I have worked and about whom a good deal of relevant information is available.

Following in the footsteps of his French model, our Liberian Binet would want to make a catalogue of the kinds of activities that children are expected to master by their parents and the village elders. People in rural Liberia make their living by growing rice and other crops, which they supplement with meat and fish when these scarce commodities can be obtained. Rice farming is physically difficult work that demands considerable knowledge and planning for its success, but as practiced by the Kpelle, it is not a technologically sophisticated enterprise. It is carried out using simple tools such as a machete to cut the underbrush; fire to burn the dry bush; vines to tie together fence posts in order to keep out animals, and slingshots to harass (Gay, 1973). Other aspects of Kpelle material culture are relatively simple, although in every case the proper use of tools requires a good deal of knowledge about how the tools are supposed to be used. There is division of labor among Kpelle adults (men hunt, women do most of the fishing; men cut the bush on the farms, women plant the seed, children guard the crops), but far more than is true of contemporary American, everyone pretty well knows what there is to know about adult economic

### ***Cole: The Illusion of Culture-free Intelligence Testing***

activities. There are some specialists (blacksmiths, bonesetters, weavers) whose work is an exception to the generalization, and study of their activities would certainly be important.

Of course, there is more to getting through life as a Kpelle than growing rice or weaving cloth. All descriptions of the social organization of Kpelle life stress that, as in America, knowledge of the social world is essential to adult status (Bellman, 1975). Kpelle people are linked by a complex set of relations that control how much of the resources available to the society actually get to the individual.

Faced with this situation, how should our West African Binet proceed? Should he sample all the kinds of activities valued by adults? This strategy is almost certainly unrealistic. Even allowing for the possibility that aspects of technology make it reasonable to speak of the Kpelle as a "less complex" society than our own, it is very complex indeed. No anthropologist would claim to have achieved a really thorough description of even one such society. Moreover, like Tylor, he would have to admit the possibility that in some respects Kpelle society provides members with more complex tasks than we are likely to face.

Since it is unreasonable in Liberia, as it is in the United States to think that we can come up with a test that samples all types of Kpelle adult activities, why not follow Binet's example and sample an important subset of those activities? From an anthropological perspective, schools are social institutions for assuring that adult knowledge of highly valued kinds gets transmitted to a society's next generation (it must be transmitted, or there would be no later generations!). While the school is not likely to be a random sample of life's tasks, it is certainly a convenient place to sample activities that adults consider important, activities that are complex enough to make it unlikely that kids would learn what they need to know simply by "hanging around."

So, our Liberian Binet might decide to search for some institutions in his society that correspond roughly with the basic goals of schooling in ours. Not all societies readily manifest such institutions, so that anthropologists are led to speak of "socialization" as the broadest relevant category. Fortunately for discussion, in the case of Liberia, he would undoubtedly discover the existence of institutions called "bush schools" in the Liberian English vernacular.

There are no detailed accounts of the curriculum of the bush school. The three or four years that youngsters spend are organized by town elders who are leaders in the secret societies that control a variety of esoteric information. This material cannot, on pain of death, be communicated to outsiders. However, we know enough about aspects of bush school activities to continue our hypothetical research (Bellman, 1975; Gay, 1973); we know that youngsters learn to farm, construct houses, track animals, shoot birds, and carry out a variety of adult economic activities (children live apart from their home villages in something like a scouting camp during their time in bush school). They are also interested in the important lore of the group. This lore is communicated not only in a variety of ceremonies, but in stories, myths, and riddles. So, let us suppose that our West African Binet decided to use "successful execution of bush school activities" as the abilities he wanted to sample.

Again, like Binet, our researcher would not be able to sample all such activities for his test, nor would he want to. He would not, for example, want to sample activities that all children knew how to accomplish before they got to school, nor would he want to sample activities considered so universally accessible that everyone mastered them well before the end of schooling. This information would not help him pick out those children who needed extra instruction. Instead, he would seek those activities that discriminated among children, activities that some mastered far earlier than others, and perhaps activities that some mastered only in later life. Once these Binet-like restrictions had been placed upon the activities selected for study, our hypothetical researcher could begin selecting tasks on which he could base test items.

### ***Cole: The Illusion of Culture-free Intelligence Testing***

In considering what sort of test would emerge, it is useful first to consider what activities would be excluded as well as those included. Cutting brush or sowing rice seed probably would not be the test; everyone knows how to do that before he or she gets to school. Nor would anyone spend time explicitly teaching children common vocabulary. However, there would be explicit instruction in such tasks as constructing houses and identifying leaves that are useful in different kinds of medicine. There would also be some mechanism for insuring that the history of the group and its laws and customs were taught to everyone in the form of stories and dances. Finally, some children would be selected for specialist roles that would require special tests (bonesetter, weaver, midwife, blacksmith, hunter, and so on). These children would receive additional instruction.

Looking at those areas where instruction might be considered important, we can see many candidate activities for testing. We might want to see if children had learned all of the important leaf names for making medicine (Bowen, 1964). Riddles are often important parts of stories and arguments, so we could test to see how many riddles children know and how adept they are at interpreting them (Kulah, 1977). The specialties would be a rich source of test material, especially if we thought that rational testing of ability to perform like adults would improve the quality of our cloth or machetes. In short, it seems possible, in principle, to come up with test items that could perform functions in Kpelle society similar to the way that Binet wanted to use IQ tests.

Could we carry out such a program of research in practice? There is no simple answer to this question, but it is useful to consider the obstacles. For some activities such as naming leaves or remembering riddles, it should be relatively easy to make the relevant observations because the Kpelle have already arranged for them: several researchers have described children's games that embody precisely these activities (Cole, Gay, Glick, and Sharp, 1971; Lancy, 1977). We could also test people's skills at constructing houses, weaving designs, and forging sturdy hoes. However, from a Kpelle point of view, tests of such skills would not be particularly interesting. The real stuff of using one's wits to get along in the world has been excluded.

This point was made very explicitly by a Kpelle anthropological acquaintance of mine who was versed in the more esoteric aspects of Kpelle secret societies and medicine (or magic, according to American stereotypes). We had been talking about what it means to be intelligent in Kpelle society (the most appropriate term is translated as "clever"). "Can you be a clever farmer?" I asked. "No," came the reply. "You can be a hardworking farmer, or you can be a lucky farmer, but we couldn't say that someone is a clever farmer. Everyone knows how to farm. We use 'clever' when we talk about the way someone gets other people to help him. Some people always win arguments. Some people know how to deal with strangers. Some people know powerful medicine. These are the things we talk about as clever."

In this bit of dialogue we see an emphasis on activities that require social interaction as the arena where intelligence is an appropriate concept. (Among the Kpelle and many other nontechnological groups, display of a good memory for use in discussions is often considered an important component of intelligence, Dube, 1977) This usage is quite consistent with Binet's analysis; it is those activities that differentiate among people in terms of the way they manipulate information that the Kpelle, like the French, use to mark intelligence.

However, once we reach this point, we face two important difficulties. First, the situations that we have selected for our study of Kpelle intelligence are exceedingly difficult to describe. Second, these contexts are very difficult to arrange. It is not enough to know riddles, everyone knows riddles. What is important about riddles is how they are used to get one's way with other people. Riddles are a resource to be used in a variety of social interactions where people's status's and rights are at issue.

### ***Cole: The Illusion of Culture-free Intelligence Testing***

Consider the first difficulty. Bellman (1978) recounts an occasion when an elder member of a secret society told a long story about how he came to be a high ranking shaman. He followed this (presumably autobiographical) story with a long riddle, which was also in story form. A novice such as myself would have no way of figuring out what part of the story was true, and I certainly would not have responded to the riddle as if its interpretation depended upon the autobiographical story; the two monologues appear to be about quite different topics. Bellman succeeds in demonstrating, however, that the riddle is closely linked to the autobiography. Not only are there formal, structural similarities (once one understands the basic categories of the relevant Kpelle belief systems). There is a rhetorical link as well. The autobiographical story actually represents a bit of self-aggrandizement by the person who told it. The man is claiming special knowledge and special power in a covert manner. The riddle reinforces the main point of the story (which raises the teller above his fellow shaman), giving the story "logical" as well as "historical" validity. The fact that listeners are constrained to agree with the riddle also gets them to agree, at least in part, with the message of the autobiographical story.

By almost any account, this man's autobiographical account plus riddle is a clever bit of behavior. It is exactly the kind of thing that our West African Binet ought to be sampling. But, at precisely this point, our cross-cultural thought experiment in IQ testing comes apart. As I have already pointed out, in order to construct a test Binet needed to be able to select a large number of items. But the "item" we have just described (very loosely) is not easily constructable. The participants in this scene were doing social work on each other; the shaman, in particular, was attempting to establish his preeminence using an account of his past history that would be difficult to check up on, a riddle whose structure was designed to reinforce his account, and his knowledge of his listener's state of knowledge concerning both the shaman's past and Kpelle social structure. This was one item; it was constructed by the subject, not the "tester." It is very difficult for me to imagine how to insure that a test includes one or more items "of this type." Furthermore, because the example's structure and content depend upon the special circumstances surrounding it, how could I insure that I would be able to present the test to the subject since it was the "subject" who did a lot of the presenting in the example I have described?

Here the contrast with Binet's situation is very strong. Like Binet, we have proceeded by figuring out what sorts of activities differentiate people according to some notion of what it means to behave intelligently. Unlike Binet, the activities we need to sample in West Africa to accomplish this goal lead us into domains that are systematically absent from Binet's tests. These domains involve interactions among people in which flexibly employed social knowledge is of paramount importance. They are not domains of hypothetical knowledge; rather, they always involve some real operations on the world, operations that require a great deal of care simply to describe. We have no good notion of how to make such activities happen in a manner analogous to the way that teachers make vocabulary tests and multiplication problems happen. Furthermore, even if we solved all these problems, we would have no real theory of the psychological processes that our subject engaged in. Such problems have not been studied by cognitive psychologists.

On both practical and theoretical grounds, then, it appears virtually impossible to come up with a way of testing Kpelle intelligence in a manner really equivalent to what we understand to be intelligence tests in our society. So long as we restrict our attention to Kpelle culture, this conclusion should not cause much consternation. After all, the idea of a West African Binet is rather absurd; Kpelle people have managed to pass on their culture for many years without IQ tests to help them select clever children and give extra assistance to the dull.

### **Some Implications for the Notion of a Culture-Free Test**

Our characterization of what one has to do to be clever in Kpelle culture and what it would take to sample such cleverness in a test must be discomfiting for anyone who imagines that one can

### **Cole: *The Illusion of Culture-free Intelligence Testing***

construct a culture-free test of intelligence. Imagine, for example, that by some quirk it was our imaginary Liberian Binet who constructed the first IQ test, and that other West African tribal people had adopted it. Next, imagine that American children were posed items from the West African test. Even items considered too simple for Kpelle eight-year-olds would cause our children severe problems. Learning the names of leaves, for example, has proven too difficult for more than one American Ph.D.<sup>14</sup> Our children know some riddles, but little use is made of such knowledge in our society except for riddling, which would put them at a severe disadvantage on more "advanced" items.

If our children were forced to take a test constructed by a West African Binet, we might object that these Kpelle-derived items were unfairly biased toward Kpelle culture. If the eventual incomes of our children depended in any way on their ability to interpret Kpelle riddles, we would be outraged. Nor would we be too happy if their incomes depended upon their use of their own riddles as rhetorical devices. At the very minimum, we would want a culture-free test if real life outcomes depended upon test performance. However, what kind of test is a West African Binet likely to dream up that we would consider culture-free? It would not involve a set of drawings of geometrically precise figures, because Kpelle, a preliterate group, do not engage in much graphic representation and they have no technology for drawing straight lines. It would not be recall of lists of nonsense syllables or even lists of words, because there are no corresponding activities in Kpelle adult life. We might try a memory test like recalling all of one's family, but here the Kpelle, who teach their children genealogies, would have a distinct advantage: what is the name of your grandmother's father on your father's side of the family? In fact, if we run down the list of presumably culture-free items that our mental experiment on Kpelle IQ testing turned up, we would almost certainly find none of the subtests that have been claimed as culture-free tests of intelligence in our society. The reason is very simple; our West African Binet, having scientifically sampled his culture, would have come up with items that reflect valued activities and that differentiate people in his culture, while Binet and all his successors have come up with items that do the same job in their culture. They are different kinds of activities.

Our imagined study of cross-cultural test construction makes it clear that tests of ability are inevitably cultural devices. This conclusion must seem dreary and disappointing to people who have been working to construct valid, culture-free tests. But from the perspective of history and logic, it simply confirms the fact, stated so clearly by Franz Boas half a century ago, that "mind, independent of experience, is inconceivable."

### **References**

- Bellman, B.L. (1975). *Village of Curers and Assassins: On the Production of Fala Kpelle Cosmological Categories*. The Hague: Mouton.
- Bellman, B.L. (1978), "Ethnohermeneutics: On the Interpretation of Subjective Meaning," in W.C. McCormack and S. A. Wurm (Eds.), *Language and the Mind*. The Hague: Mouton and Co.
- Boas, F. (1911), *The Mind of Primitive Man*. New York: Macmillan,
- Bowen, E.S. (1964). *Return to Laughter*. New York: Doubleday, 1964).
- Cole, M., Gay, J. Glick, J.A., Sharp, D.W. (1971). *The Cultural Context of Learning and Thinking*. New York: Basic Books
- Dube, E.F. (1977). *A Cross-cultural Study of the Relationship between 'Intelligence' Level and Story Recall*. Doctoral dissertation, Cornell University.
- Gay, J. (1973). *Red Dust on the Green Leaves* (Thompson, Conn: Inter-Culture Associates, 1973)

**Cole: *The Illusion of Culture-free Intelligence Testing***

Gibbs, J.L., (1965), The Kpelle of Liberia, in J.L. Gibbs (Ed.), *Peoples of Africa*, New York: Holt, Rinehart and Winston.

Hernnstein, R.J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Jensen, A. (1980). *Bias in mental testing*. New York: Free Press.

Kulah, A.A. (1973), *The Organization and Learning of Proverbs among the Kpelle of Liberia*.  
Doctoral dissertation, University of California, Irvine

Lancy, D. (1977), Studies of Memory in Culture, *Annals of the New York Academy of Science*, 307, 285-297.

Rivers, W.H.R. (1901). Vision. In A.C. Haddon (Ed.) *Report of the Cambridge anthropological expedition to the Torres Straits, Vol. 2*. Cambridge: Cambridge University Press.

Spencer, H. (1886). *The Principles of Psychology, vol. 5*. New York: D. Appleton.

Tylor, E.B. (1958). *The Origins of Culture*. York: Harper and Row.