Issues in the definition and measurement of abilities*

David F. Lohman

The University of Iowa, USA

Paper presented at the Second Spearman Seminar, University of Plymouth, July 14, 1997.

* DRAFT DOCUMENT. DO NOT QUOTE OR CITE WITHOUT PERMISSION.

Probably the greatest satisfaction I obtain from the academic life is the experience of opening the door on a domain I had neglected, and discovering new systems of ideas which, upon reflection, helped me think about more familiar ideas in new ways. In the 1970's the door opened on information processing, and then later on cognitive psychology generally. Down the years, other doors have opened on anthropology, philosophy, history, and most recently on evolutionary biology. I make no claim, of course, to discovery. All that I have seen has been seen more clearly before. In fact, probably every idea I have had about human abilities is probably contained in some form in one of the thousands of journal articles, books, and book chapters devoted to the topic. Indeed, one lesson to this tale is that, after many months of reading and thinking about the implications of evolutionary biology for a theory of abilities, I discovered that Dick Snow had been there ahead of me. Nevertheless, it is useful--even necessary--to cross over periodically to an unfamiliar domain in order to gain perspective on one's own domain. This chapter, then, is the report of one journey of this sort.

More specifically, in this chapter I discuss different approaches to the definition and measurement of abilities. Following Mayr's (1982) summary of the biological sciences, I begin by distinguishing between population thinking and essentialist thinking. Variation and diversity are the stuff of population thinking: categories and typologies are the stuff of essentialist thinking. Population thinking characterized much of Darwin's work in evolutionary biology, particularly the Darwin-Wallace theory of natural selection, and later Galton's studies of the inheritance of mental and physical traits. Essentialist thinking, on the other hand, has ever guided experimentalists in both biology and psychology. Attempts to reduce these two types of thinking to one are briefly reviewed. I conclude this section by arguing for the legitimacy of a differential psychology that cannot be reduced to (or explained by) experimental psychology. I then discuss four ways in which the concept of ability has been defined in differential psychology: (1) as a latent trait inferred from patterns of individual differences across tasks, (2) as level of performance on a particular task or class of tasks, (3) as a latent cognitive process inferred from within-subject patterns of performance across trials within a task, and (4) as an affordance - effectivity relaxation (i.e., a joint property of the union of person and environment). I show how both population thinking and essentialist thinking have differentially influenced advocates of each of these definitions of abilities. I conclude with recommendations about how we might best conceptualize and measure human abilities.

World Views

Scholars trained in different disciplines conceptualize problems differently. Sometimes the differences in perspective and method are profound, as between the humanist and the radical empiricist. In other cases the differences are subtler, as for example when a psychologist steeped in developmental theory sees abrupt, stage-like transitions in the history of cognitive science (e.g., Gardner, 1985), or when a psychologist steeped in the categorical modes of thinking that dominate experimental psychology attempts to explain individual difference constructs of personality psychology (e.g., Cantor, 1990). I have come to believe that these general habits of thought, these characteristic ways of perceiving and organizing experience (or "world views," Pepper, 1942) are not just interesting epiphenomenon in the grand show of science, but are more like foundational elements that critically shape the sorts of theories we build--and, more importantly--cause conflict among those who adhere to different foundational assumptions within and between disciplines. Theories of human abilities are the product not only of data and argument but also of the personal proclivities and professional experiences of theorists, of their beliefs about what science is and how it should be conducted, and of the larger social, political, and religious themes that form the fabrics of the cultures in which they live (Lohman, in press).

There is also issue of the extent to which our methods -- particularly the statistical methods we use -- distort and mislead us. David Bakan (1973), Louis Guttman (1971) and many others have commented on this aspect of our enterprise. Statistical and psychometric methods both reflect and help perpetuate different modes of thinking. Indeed, I will argue that differential psychology requires a style of thinking quite unlike the style of thinking that serves us well in the physical sciences and in much of experimental psychology. At the outset it is important to note that I am not arguing that one style of thinking is better than the other, or that individuals can be typed by the style they prefer. Indeed, most of us move back and forth between these two ways of thinking. I will claim, however, that the essentialist or typological way of thinking is easier, seems more naturally to conform with our cognitive architecture, and thus both developmentally and historically precedes probabilistic thinking.

Essentialism Vs. Populations Thinking

In psychology--as in biology--one of the more pervasive differences in conceptual style is between essentialist or typological thinking and population or stochastic thinking. The distinction is suggested at in the cognitive style literature. Messick and Kogan (1963) discuss a style they call compartmentalization, which refers to the tendency to isolate ideas and objects into discrete and relatively rigid categories. But the obverse is a willingness to tolerate fuzzy concepts which they link to ideational fluency, not population or stochastic thinking.

Essentialism can be traced back to Plato, and surely earlier for anyone who cared to look. Objects in the world are but imperfect shadows of more perfect forms or ideas or

essences. These forms are more permanent and therefore more real than the particular objects through which we conceive and deduce them: Man is more permanent than Dave or Bob or Pat; the circle that I draw will someday fade but the form circle endures forever. Importantly, then, variation among category members reflects error or imperfection in manifestation of the essential form. The philosophy of essentialism has fitted well with the conceptual structure of the physical sciences. Carbon atoms are indeed alike; those that differ define new isotopes or ions (i.e., a new category). Closely linked with this type of categorical thinking is a deterministic (as opposed to probabilistic) view of causation. Essentialist thinkers typically work in worlds in which causal sequences may be described "IF A, THEN B."

In psychology, those trained in experimental methods seem most comfortable with essentialist modes of thought. This is particularly evident in attempts of experimentalists to explain individual differences. Most, of course, do not get beyond the notion of individual differences as error, and thus see no need to explain them. But for those who do, there is usually (a) an attempt to impose a typology of some sort on the data (thus, we have not one type of person in the world but two types, which upon closer inspection, are further subdivided, ad infinitum, as in stage-theoretic models of development), and (b) an attempt to escape from the unstable bog of relative measurement onto the seemingly firmer ground of absolute measurement. For example, in their early paper in which they advocated an information-processing approach to the study of human intelligence, Hunt, Frost, and Lunneborg (1973) claim:

> The gist of our argument is that intelligence should be determined by absolute measures of aspects of a person's information processing capacity rather than by measures of his performance relative to the performance of others in a population (p. 119-120).

Hunt's research program then sought methods for measuring what were thought to be structural or mechanistic, information-free mental processes, typically on the absolute scale of response latency.

When confronted with questions of style or strategy, the experimentalist prefers an explanation that emphasizes qualitative rather than quantitative differences. This was clearly evident in the early work of Cooper on individual differences in visual comparison processes. Cooper (1982) identified two types of individuals: those who appeared to use a holistic strategy for comparing forms and those who appeared to use an analytic strategy. A similar preference for qualitative differences may be observed among more experimentally-oriented personality theorists. Most of these typologies do not survive close inspection. In the case of Cooper's typology is was, paradoxically, one of Hunt's graduate students who unmasked the continuum (see Agari, 1979).

Probabilistic thinking about populations takes the opposite tack. Population thinkers stress the uniqueness of each individual. There is thus no "typical" individual; mean values are abstractions. Rather, variation is the most interesting characteristic of natural populations. And this variation is multidimensional. Causal sequences are less mechanistic and more stochastic: If A, then B, but with probability C. Indeed, C may be perceptible only at the population level. Essentialists find this sort of thinking particularly difficult. Who has not heard an essentialist argument against the connection between smoking and lung cancer that rests on a single, octogenarian counterexample?

Differential psychology is, of course, grounded in population or probabilistic thinking. As such, it is more concerned with quantitative than with qualitative differences, and with relative rather than with absolute scales of measurement. Because the differentialist is often criticized for his reliance on relative measurement, he sometimes looks wistfully at the absolute measurements that his experimental colleagues have at their disposal. However, I believe that this envy is misplaced; measures of the relative fit between persons and situations is what his discipline is all about. Thus, even when absolute measures (such as latency) are available, it is information about the relative standing of individuals that is his special concern. This brings us closer to the heart of the matter, i.e., our differing conceptions of personality and ability constructs, particularly the latter.

Parts of Speech

One way to understand the source of our differing conceptions of abilities is to examine how terms that denote abilities are used linguistically. This is another of those ideas I was sure that I had discovered only later to find it clearly presented in a text I know that I had read years ago [in this case, Butcher's (1968) classic]. Perhaps I merely reconstructed a new version of these arguments from the kernel of a vaguely remembered idea. Or perhaps, as Dennett (1995) argues, it is not so much that great (and not so great) minds think alike as it is that we work within a design space that favors some moves and discourages others. Any serious consideration of the term ability will eventually have to consider whether it is a noun, an adjective, an adverb, or even a verb. Although this is a much more limited undertaking than Sternberg's (1990) discussion of the metaphors that underpin different theories of intelligence, there are interesting points of convergence.

The essentialists among us--that is, those who more strongly identify with experimental rather than differential psychology -- have sought to explain abilities in terms of the size or capacity of working memory, the speed or efficiency of information transmission within the system, or the attentional resources at the individual's disposal. In this way, intelligence is sometimes viewed as reflection of a structural difference in information processing systems. Hunt (1983) asks "What does intelligence do?" Although the question and the process view it

entails seems to invite the use of verbs, the measurement procedures used at best invoke the adverbs and adjectives -- larger working memories, faster processing, less proactive inhibition. In a similar, but more abstract vein Cronbach (1977) claimed, "Intelligence is not a thing; it is a style of work (p. 275)." In other words it is a way of characterizing how something is done, which inevitably involves a value judgment. To reason intelligently implies a different way of solving problems. So, for example, when a computer wins chess games by virtue of the brute force of computational algorithms we do not rate its performance as particularly intelligent.

Those steeped in traditional differential methods, on the other hand, seem most comfortable with the view that "intelligence" is best viewed as an adjective that describes a person or a particular class of behaviors (Anastasi, 1986). Unlike some personality and stylistic traits, it is a marked adjective: There is clearly a positive valence associated with being intelligent and a negative valence with being unintelligent. The adjectival use of the word also conforms well to the notion that--like beauty or tallness--intelligence is a relative concept. Who is considered intelligent depends on the range of intellectual competence in the group; the behaviors that are considered as intelligent depend on the demands and affordances of the environment, or, more generally, the culture (Sternberg, 1985).

Ability theorists thus disagree whether intelligence is best characterized as a noun (e.g., a structural property of the brain or at trait possessed in a certain amount), an adjective (e.g., identifying certain types of people), a verb (e.g., denoting certain varieties of cognition or action), or an adverb (e.g., describing the qualities of cognition or behavior, such as its speed or efficiency). Those who search for those cognitive processes and knowledge structures that generate behavior labeled intelligent often assume that some nouns will be needed, but they place the most emphasis on verbs and adverbs (i.e., how and how well one thinks). Those who study social and cultural variations in intelligence generally assume that an adjective is needed. Sternberg's (1985) componential and contextual subtheories nicely capture this divergence. In contrast, trait-based theories of personality characterize the domain as a collection of adjectives, and when traits are thought to inhere in the individual, as nouns. The interesting question, though, is whether personality also can be understood using verbs and adverbs. Some (e.g., Cantor, 1990) see this as the wave of the future; others (e.g., Cervone, 1991) are less sanguine about the possibility of a rapprochement between the experimental and differential approaches. If recent attempts to apply cognitive theory to ability constructs are any guide, then bridges will be more difficult to build than initially seems possible. However, careful attention to issues that were insufficiently addressed in ability-process research--particularly those issues concerning the definition and measurement of constructs--will surely improve the changes of meaningful progress.

Definitions of Ability

Ability as Domain Referenced

Carroll (1993) noted that "although the term ability is in common usage in both everyday talk and in scientific discussions, its precise definition is seldom explicated or even considered." (p. 3). Indeed, some historians of science believe that scientific progress consists not only in the development of new concepts but also in the repeated refinement of definitions by which old concepts are articulated. "Particularly important" says Mayr (1982), is the occasional recognition that a more or less technical term, previously believed to characterize...a certain concept, was in reality used for a mixture of two or more concepts." (p. 43). I believe that such confusion attends discussions of ability, particularly intelligence.

I suspect that many of the would-be bridge builders between the separate kingdoms of personality theory and ability theory have failed because they assumed that one could build from the terra firma of ability theory into the less firmly grounded realm of personality theory. In brief, the problem is not that the supposedly terra firma of ability theory is terra incognita as much as it is terrae firmae. In plain English, the problem is that the term "ability" is used in quite different ways by many, but especially by experimental and differential psychologists. Attempts to link ability with personality will fare no better than attempts to link experimental and differential psychology unless we attend more carefully to the ways in which these terms are used and measured.

Ability as Trait

The first and by far most popular way in which ability is defined is as a latent trait inferred from consistencies in patterns of individual differences across tasks. In the limiting case of a single task, the latent variable is synonymous with the true score of classical test theory. When scores on multiple tasks are considered simultaneously, the latent trait is estimated from the covariation in individual differences across tasks.[1] Individual differences are thus central to this definition of ability. Further, the approach emphasizes transferable competencies, something often overlooked in task- and process-based definitions of ability.

What Darwin discovered, Galton applied to humankind, and Pearson and Spearman showed how to measure, was the importance of relative standing within the group. Although the trait definition is grounded in population thinking, it does not have much to say about the environment. Indeed, context effects, if they are included at all, tend to be treated as moderator variables in such models. In other words, contextual factors merely limit the scope of generalizations about abilities that can be made -- across types of stimuli (Is there more than

[1] Spearman's (1904) original formulation seemed closer to this conception than later formulations in which error was treated as a purely random variable rather than as non-generalizable individual difference variance, (as in modern generalizability theory). Indeed, the notion of error as noise usually rests on an essentialist mode of thinking .

g?) , across ages (Does the structure or meaning of intelligence vary across the life span?), across treatments (are there aptitude by treatment interactions?) or even cultures (Does the meaning of the construct vary across cultures?) This lopsided focus on the individual opens the door for the next definition of ability, which focuses on the task rather than on individual differences.

Ability as Task Performance

Ability is also sometimes defined in terms of performance on a particular task or class of tasks. For example, in the report of the Committee on Ability Testing of the U.S. National Research Council, Widgor and Garner (1982) define ability as "systematic observation of performance on a task." In an earlier draft the authors were even more focused: Ability is "how well a person performs a defined task if he does his best." There are thus as many different abilities as there are tasks that can be administered and on which performance can somehow be observed and scored. Since everyone could fail to accomplish a task, or could succeed at it, individual differences are not a necessary component of this definition of ability. Some efforts to export the tasks of experimental psychology into differential psychology use task-based definitions of ability. For example, some researchers use measures of overall performance on the Shepard-Metzler (1971) rotation task as a measure of "mental rotations ability." This extreme focus on particular tasks is thus diametrically opposed to Spearman's principle of the "indifference of the indicator."

If items (or tasks) can be ordered such that performance can be described by a Guttman Scale, then "ability" can be defined more precisely--for example, as the point at which the probability of a correct response is 50% (Carroll, 1990; Thurstone, 1937), or at which the function relating probability of a correct response to response latency intersects a particular latency value (Lohman, 1989). While such approaches sharpen the measurement of ability within a particular task, they do not address the issue of consistency in performance across tasks.

In educational measurement, criterion-referenced (or domain referenced) tests exemplify this definition of abilities. Linn and Gronlund (1995), for example, define a criterion-referenced test as "a test designed to provide a measure of performance that is interpretable in terms of a clearly defined and delimited domain of learning tasks" (p. 16). One can move from ability as performance on one task to ability as performance on many tasks only if the domain of tasks is clearly defined. Generalizability theory provides a particularly powerful method for doing this (see Kane, 1982). Educators are not the only ones who define abilities in this way. An employer, for example, is often more interested in whether the prospective employees can perform certain tasks at a given level rather than their relative standing. The task-based definition is thus an attempt to escape the relativistic world of norm-

referenced, trait-based interpretations of test scores. It is thus a retreat from population thinking. The focus in no longer on person variance but on the task and on the behaviors exhibited by the test taker. It is no accident that this type of interpretation was advanced by a psychologist steeped in behavioral learning theory (i.e., Robert Glaser). However, things are often what they seem to be. Scores on domain referenced tests are rarely interpretable without at least some reference to the behavior of others on the test. These implicit norms are embedded in the ascription rules (Rorer & Widiger, 1983) test interpreters use to make sense of even absolute measurements. Thus, as one wit put it, "Behind every criterion there lurks a norm."

Ability as Process

Whereas ability is inferred from the comparison of one individual's performance to that of other individuals (definition 1) or to an external standard (definition 2), process is inferred from the comparison of performance in one condition to performance in another condition. Since processes occur within individuals, the inference of process is not grounded in individual differences. Because of this, the measurement of process seemed to offer not only an insight into process but an escape from the relativistic world of traditional ability testing. Although I am still persuaded by the need for process-like analyses of ability constructs, I am less sanguine about the utility of the process measures derived from such analyses.

For example, consider the much-studied mental rotation task. In this task, subjects are shown two stimuli that differ in orientation. They must determine if the two stimuli can be brought into congruence. Shepard and Metzler (1971) proposed that subjects confronted with such problems form mental images of the stimuli, rotate one of these images the required distance, compare the two images, and then respond. They tested their model by regressing angular separation between stimuli on response latency. The slope of this function estimates the rate at which stimuli are rotated. The expectation has been that the slope parameter would provide a relatively pure measure of spatial ability. However, if anything, it is the intercept parameter that shows consistent correlations with other variables; correlations for the slope vary from highly negative to moderately positive (see Lohman, 1994). Such results have dampened enthusiasm for using estimated rate of rotation as a measure of spatial ability, but have not seriously challenged the fundamental assumptions of methods that rely on this sort of task decomposition.

To understand why component scores and other process measures are not what they seem to be, imagine a simple person by item data matrix whose entries X_{pi} represent the scores of n_p persons on n_i items or trials. Figure 1 shows how the variability in scores may be

---

Insert Figure 1 here

---

partitioned into three sources: the person source, the item source, and the residual. The person source represents variability in row means, that is, in the average performance of each person on the task. This would be the score ordinarily reported on a mental test. It thus represents the ability construct we hope better to understand. The item source represents variability in column means, that is, in average differences in item (or trial) difficulty. In the rotation example a large fraction of this variability can be attributed to the amount of rotation required. The residual is composed of the person by item interaction and other disturbances. In the language of reliability theory, it is the error variance. Individual differences in slope scores help salvage variance from this residual component. However, process scores defined by within-person contrasts of any sort do not decompose and therefore cannot help explain the typically much larger person variance component. In fact, mean scores for each person that are reflected in the p variance component will generally show high correlations with the intercept of the regression model, which is why the intercept often shows interesting and significant correlations with reference abilities while component scores show inconsistent correlations with such measures (see Lohman, 1994).

The problem that confronts us is actually much more complex than this simple two way classification suggests. Personality and style variables complicate the picture. Figure 2 shows a modified version of Cattell's (1966) covariation chart: persons x items (nested within tasks) x

---

Insert Figure 2 here

---

occasions (or situations). Differential psychologists typically worry about person main effects (or covariation of person main effects across several tasks). Experimental psychologists are less uniform. Those who follow an information-processing paradigm worry about variation over trials with a particular task. Situationalists, however, worry more about covariation of either task main effects (e.g., delay versus no delay of reinforcement) or person main effects across occasions; they typically emphasize the magnitude of the former relative to the magnitude of the latter. Developmentalists do the opposite. Then there are those who worry about interactions. The point is that any rapprochement between experimental and differential psychology has many dimensions, not just two. Person x situation is not the same as person x items within task. The bottom line is this: just because a construct has the same name in two different literatures does not mean that it refers to the same -- or even correlated -- aspects of variation. A more systematic accounting of which aspect of variability is represented by

different constructs may help us keep track of our constructs and keep in line our expectations for relationships among them.

Thus, in my opinion, the attempt to define ability by individual differences in within-subject component or process scores is unlikely to succeed. As noted, process-like scores generally do not capture much of the interesting individual differences on tasks (which is not to deny that they will sometimes show interesting and replicable correlations with other variables). Furthermore, like task-based definitions of ability, the process approach emphasizes absolute rather than relative measurement. The process explanations it offers are most informative when they uncover qualitative rather than on quantitative differences between individuals. Therefore, the primary contribution of an information-processing type of analysis of a task or problematic situation lies in the information such analyses provide about how subjects understood the situation or solved the task. Such analyses contribute not new scores, but new methods for addressing fundamental questions about construct validity. Returning to the mental rotation example, such analyses are helpful if they can tell us if subjects are indeed mentally rotating stimuli or are engaging in other strategies that might compromise the interpretation of their scores as measures of the construct we call spatial ability.

Ability as situated

The fourth and last definition of ability tries to bring in a vastly more relativistic definition of the stimulus environment. In other words, ability is seen not as the relative standing along some cognitive dimension of an individual within a group (definition 1), or of an individual's performance relative to some well-defined class of tasks (or performances) (definition 2), or even of facility in performing certain types of cognitive processes (definition 3), but rather is a joint property of the union of person and environment.

Snow (1994) has given the most articulate statement of this perspective. He begins by borrowing Gibson's (1979) concept of affordances to describe person-situation connections.

> The affordances of a situation are what it offers the person--what it provides or furnishes, for good or ill. The term implies a complementarity of person and situation, as in an ecological niche. A niche is a place or setting that is appropriate for a person...Affordances thus reflect the invitation, demand, or opportunity structure of a situation for those persons who are tuned or prepared to receive them. (p. 28)

Thus, particular affordances invite particular actions. In Gibson's terminology, these actions are called effectivities. "Abilities" Snow concludes "are properties of the union of person and environment that exhibit the opportunity structure of a situation and the effectivity

structure of the person in taking advantage of the opportunities afforded for learning" (p. 31). There is thus a reciprocity, a dance, between person and the situation. Change the demands or affordances of the situation, and you change the apparent abilities of the person--which is the repeated demonstration of the ATI literature. Change the effectivity structure of the individual, and you also change the ability of the person--which is the repeated demonstration of the literature on the effects of schooling and culture on cognitive competence. But the match is always relative: even when all are matched, some will be better matched than others; even when all are poorly matched, some will be less poorly matched than others. Indeed, as Snow notes and as Piaget demonstrated, we can learn much about what abilities are by studying those cases where there is a clear mismatch between the inner environment of the individual and the outer environment.

Although this view of abilities seems to derive from current (see Greeno, Collins, & Resnick, 1996) and past (Gibson, 1979) theories of situated cognition, it is also fundamentally Darwinian. Long before Simon discussed the interface between inner and outer environments in terms of artifacts, Spencer (1873) concluded:

> Regarded under every variety of aspect, intelligence is found to consist in the establishment of correspondences between relations in the organism and relations in the environment; and the entire development of intelligence may be formulated in the progress of such correspondences... (p. 385).

However, Spencer saw both intelligence and environment through unidimensional glasses. Indeed, most discussions of human intelligence speak as if intelligence means superior adaptability in <u>all</u> environments. Those who escape this chapter of the flat earth society seem at best able to see environments as arrayed along a unidimensional scale from best to worst. A careful reading of Darwin shows greater subtlety:

> The meritocratic selector and the experimental reformer alike missed the point of Darwin's theory. The theory did not posit that <u>generally</u> superior creatures evolve. [Rather, Darwin was]...concerned with fitness to survive in a <u>particular</u> ecology. To foster development of a wide variety of persons, then, one must offer a wide variety of environments. A social reform that would standardize the environment (whether to fit the average person, or the present elite, or the present proletariat) is inevitably procrustean, conservative, and self-limiting. (Cronbach & Snow, 1977, p. 11)

Although ability theorists have been guilt of ignoring situations, advocates of situated cognition have been guilty of the opposite fallacy. Many wrote as if there were no consistent individual differences across situations. Transfer became "a problematic issue" for those who advocated a situated view of cognition (Greeno, Collins, & Resnick, 1996, p. 24). Since abilities may be defined as transferable knowledge and skill, an approach which finds transfer at best "problematic" seems unlikely to provide much insight into human abilities. What Snow (1994) has done, however, is to bring individual differences back into discussions of situated cognition, and has done so in a way that goes considerably beyond early (and largely unsuccessful) attempts to define intelligence in terms of adaptability. Individuals perceive and create regularities across contexts that permit the application of old knowledge to new situations. Abilities thus allow "attunements to constraints and affordances .. that remain invariant across transformations of situations." (Greeno et al., 1996, p.24).

The Uniqueness of Differential Psychology

One of the more unfortunate consequences of the explosion of knowledge is that we rarely have time to step outside of the narrow confines of our own domains to see what is happening in our neighbors' backyards. The same wars that have raged between experimental and differential psychology have also plagued biology. Experimentally-oriented biologists (and their allies in the physical sciences) have scoffed at the observational-comparative methods of naturalists, paleontologists, and evolutionary biologists. Indeed, in philosophies of science written by physical scientists (or those who adulate their work), manipulative experiments are often referred to as the method of science. Yet, observation, classification, comparisons across individuals, groups, or time periods, and historical narratives are all legitimate scientific methods.

> Observation led to the discovery of foreign faunas and floras and became the basis of biogeography; observation revealed the diversity of organic nature and led to the establishment of the Linnaean hierarchy and to the theory of common descent; observations led to the foundations of ethnology and ecology. Observation in biology has probably produced more insights than all experiments combined. (Mayr, 1982, p. 32).

The physical sciences have been eminently successful. They also rest securely in a categorical, typological mode of thinking that we humans seem to find congenial. They traffic in clear concepts with clear boundaries that can be easily mathematized. However, the organisms that the biological and psychological sciences attempt to understand are vastly more complex than the systems that physical scientists study. Every organism is the product of a

history that dates back more than 3000 million years. Indeed, generalizations in the biological and social sciences are almost invariably probabilistic. As one writer put it: "There is only one universal law in biology: All biological laws have exceptions." Similarly, in psychology, test-theoretic models of abilities are invariably probabilistic rather than deterministic (Lord Novick, 1968, p. 23). We never say "If A, then B," but rather "If A, then maybe B."

The Darwinian revolution in biological thought was rooted in a shift from typological to population thinking. Darwin realized that members of a species differed importantly from one another, and that these individual differences were stuff upon which natural selection operated. Without a doubt the term "individual differences" or a synonym ("variation," "diversity") is the most frequently used term in chapter IV of The Origin wherein Darwin advances the theory of natural selection.

It was Galton, however, who first realized that such variation in human populations could be studied quantitatively. Quatelet, the Belgian astronomer and statistician whose work inspired Galton, believed that the mean of a distribution represented the ideal toward which nature was working; deviations from the mean were simply departures from this ideal.[2] Galton realized, however that such distributions could be used to document the extent of variability of human biological and psychological characteristics upon which natural selection operated. His interest, therefore, was not in the mean of the distribution but in its variance.

Spearman extended this sort of probabilistic thinking to the selection of tasks that served as indicators of intelligence. In doing so he shed the categorical chains of thinking about thinking that haunted Binet to the end. But probabilistic thinking about populations cannot stand alone. Darwin was unable to get beyond a Lammarkian theory of inheritance of acquired characteristics because the black-box of genetics had not yet been opened. After Wiseman and Fisher and most especially Watson and Crick, evolutionary biology has made great strides. But the new experimentally-based theories of genetics have not and will not somehow supplant evolutionary biology. In much the same way, early theories of human intelligence were unable to move beyond a belief in innateness because they lacked a cognitive theory of learning and development. Experimental studies of thinking and its development thus usefully inform, but do not dispense with the need for the study of human cognitive diversity. Sometimes I think we need to be reminded that differential psychology need not find justification outside of itself. This does not mean that it should ignore the work of experimentalists. It means, rather, that differential psychology cannot be reduced to or explained away by experimental psychology.

Personality-Ability Connections

---

[2] Quatelet's conception of the mean as ideal type endures in psychometrics as a "Platonic true score" (Lord & Novick, 1968).

Thus far I have tried to argue that the differences between experimental and differential psychology are real. The two are not only grounded in different world views, but explain quite different aspects of variation. Differential psychology will be informed by, but cannot be reduced to, experimental cognitive psychology, or neurology, or any other discipline. I have also tried to argue that the most profitable way to understand abilities is not to view them in a domain-referenced, or trait, or even in a process way, but rather, as Snow (1994) suggests, as the joint property of person and context. This situated view of abilities has a number of interesting implications. For the purposes of this conference, I will note two that have direct implications for the ability-personality question that confronts us: the role of volition, and the affective match between student and mentor.

One aspect of ability that was easily overlooked in trait definitions and the factor-analytic research it inspired is that all abilities are developed through an extensive transaction with the personal and social environment. Indeed, the most important non-biological factor in the development of what we call intelligence is formal schooling. The more schooling, the greater the gains in intelligence. Correspondingly, the single most important factor in predicting absolute gains in narrower ability and skill constructs is the amount of focused practice. In a recent review of expert performance, Ericsson and Charness (1994) concluded:

> Expert performance is predominantly mediated by acquired complex skills and physiological adaptations. For elite performers, supervised practice starts at very young ages and is maintained at a high level for more than a decade. The effects of extended, deliberate practice are more far-reaching than commonly believed. (p. 725)

Although Ericsson and Charness can rightly be accused of understating the influence of genetic factors, it is fair to say that most ability theorists even more dramatically underestimate the cumulative effects of five, ten or even twenty years of guided practice. The important point here, however, is that one must not only be so fortunate as to have high quality instruction available throughout this long period, but one must somehow persist. And therein, I think, lies one of the chief connections between ability and personality. The central construct is "volition." Since I expect Corno will discuss this in greater detail in her presentation, my summary will be brief.

Volition

"Volition" is an old term new lease on life. To do something of one's own volition means to do it "by one's own resources and sustained efforts, independent of external source or pressure" (Corno, 1993, p. 14). But early in this century, those who

studied motivation claimed this was the proper purview of their discipline. Also, associations of volition with free will and other "prescientific" concepts from an earlier psychology lead to the abandonment of this construct, at least by most U.S. psychologists.

The German psychologists Kuhl and Beckmann (1985) revived U.S. interest in volition with their theory of action control. Corno and Kanfer (1993) elaborate an educational view of this work. The basic idea is straightforward: motivation concerns those affects and processes that initiate behavior, that move us from wishes to wants to actions. From purely cognitive perspective, motivation is about goal setting. In Kuhl's view, it concerns the "predecisional" phase of action.

Volition, on the other hand, concerns those processes whereby one actively maintains an action, often in the face of competing action tendencies and negative affect. It is post-decisional. Kuhl describes several aspects of volition. Two of the most important are: 1) strategies for the protection of goals against competing goal tendencies, and 2) strategies for the management of affect, especially negative affect.. Corno and Kanfer (1993) list a variety of volitional control strategies, many of which are designed to regenerate positive affect or to control negative affect. The development of high levels of competence requires extended, guided practice over many years. Thus, in my view, understanding how some are able to protect their goals and maintain their efforts to achieve these goals is a crucial topic for understanding the development of abilities. Many start the journey, but few finish it.

It is interesting that attempts to integrate modern work on volition into older stylistic or trait view of human performance use words like "responsibility," "dependability," and "conscientiousness" to describe the individuals who exhibit these characteristics in many situations. Yet these are the same trait labels that are included in definitions of intelligence that go beyond mere cognitive competence. For example, in the same 1921 symposium in which Thorndike gave the oft-cited definition of intelligence as “the power of good responses from the point of view of truth or fact” (p. 124), he also noted:

> It is probably unwise to spend much time in attempts to separate off sharply certain qualities of man, as his intelligence, from such emotional and vocational qualities as his interest in mental activity, carefulness, determination to respond effectively, persistence in his efforts to do so; or from his amount of knowledge; or from his moral or esthetic tastes. (p. 124)

The Nichols and Holland (1963) study of 10,000 National Merit Finalists in the U.S. showed this clearly. More than 150 measures were obtained on each participant,

including personality and biographical data, and correlated with 14 criteria of success in the first year of college. The best non-cognitive predictor of first-year grades for these highly able students was a factor called "perseverance and motivation to succeed."

In order to learn how to persist, one must be challenged. Ultimately, then, attempts to develop transferable volition-control strategies are attempts to develop what used to be called character. But do volitional skills learned in one context transfer to other contexts? Certainly the better than average performance of endurance athletes in college suggests that this might be the case. But such correlational evidence is open to multiple interpretations. Indeed, most modern students of transfer would agree with Thorndike and Woodworth (1901) that transfer is generally quite limited. However, in his 1913 text Thorndike cautioned differently: "Some careless thinkers have rushed from the belief in totally general training to a belief that training is totally specialized." (p. 365) He then gave examples of "general" S-R bonds:

> Of special importance are the connections of neglect. Such bonds as 'Stimuli to hunger save at meal times - neglect them'; 'Sounds of boys at play save at playtime - neglect them'; 'Ideas of lying down and closing one's eyes save at bed time - neglect them,' and the like are the main elements of real fact meant by 'power of attention,' or 'concentration' or 'strength of will.' In so far as a certain situation is bound to the response of neglect it is prevented from distracting one in general. (p. 419)

In modern jargon, Thorndike (1913) would agree with Kuhl that volitional control strategies are among the most transferable mental competencies.

The upside of volition is that it helps an individual maintain focus; but the downside is that it may be hard to disengage these processes once they are firmly entrenched. There is a thin line between persistence and rigidity. Athletes ignore pain at their peril; the body can and does break down. Workers can persist at their tasks until work is all they have, or they have burned out. Yet some do learn to manage the tradeoff. Good athletes do learn to listen carefully to some pains while disregarding others. This higher level of adaptation is well captured in Sternberg's (1985) concept of "mental self government."

Ericsson and Charness (1994) also note that the attainment of high levels of competence requires more than persistent practice; it also requires the timely assistance ad feedback of parents, teachers, coaches, and other mentors. Good teachers are not just good technicians. Somewhere Augustine remarks that the most important thing a teacher brings to students is the example of his character. In other words, the mentoring process

is greatly facilitated if the student cares for and identifies with the teacher. One cannot merely provide the external support and expect that it will work. There must be a match between the internal environment of the learner -- that is, of the learners abilities, needs, wishes, wants, and temperament -- and the external environment -- particularly of the mentoring provided. Using the ability-as-effectivity model, the outer environment offers various affordances for action that must mesh with the inner environment of the learner. And the nature of this coupling, this dance, changes over time. What works for one will not necessarily work for another. What works well at one time may be quite inappropriate later. What works over the short haul may not be best over the long run. Most importantly, the affordance-effectivity match has a large -- and largely overlooked -- affective component (see Snow, Corno, & Jackson, 1996). Good teachers are not just good technicians.

Summary

Snow (1994) argues that abilities are best understood as "affordances -- properties of the union of person and environment that exhibit the opportunity structure of a situation and the effectivity structure of the person in taking advantage of the opportunities afforded for learning." (p. 31) Abilities are thus situated. Some persons succeed in learning in a given situation; they are in harmony with it. Others do not, because they are not tuned to the opportunities the situation provides or to produce what it demands. Over the long haul, then, affect and volition are probably as important in the development of talent as are entry level of ability and opportunities provided. The potential for great accomplishment may indeed be in significant measure a gift from one's ancestors. However, the attainment of domain expertise comes only after much learning and practice.

We work in Darwin's shadow -- not the shadow of Wundt or Leeuwenhoek or Boaz or Ward. And although the disciplines that were given shape by these luminaries inform our efforts, in the end the study of individual differences concerns the adaptation of individuals to the environments in which they are placed, or which they select, or which they help mold (see Sternberg, 1985). Adaptation -- or person-environment fit -- occurs simultaneously and interactively along many dimensions, that include not only the cognitive but also the affective and conative. Although there is nothing that prohibits the expansion of other definitions of ability to include these dimensions, only the situated definition demands -- or, better, affords -- their inclusion at the outset. Furthermore, a situated view of abilities brings us back to Darwin's insight that context matters. Therefore, as we go about the business of trying to forge alliances among the separate fiefdoms of what Cronbach (1957) dubbed the Holy Roman Empire of differential

psychology, I suggest that we consider the advantages of defining ability in this way. I also suggest that, whatever definition we use, we be wary of the Siren call of essentialism and its cousin reductionism, even though advancing a discipline based on probabilistic thinking about populations means always sailing into the wind.

Figure Captions

Figure 1. Top panel shows a schematic of the basic person by item data matrix, with entries X_{pi}. Variation in row means ($\bar{X}_{p.}$) captures differences among individuals in overall performance whereas variation in column means ($\bar{X}_{.i}$) captures differences in item or trial difficulties. Component or process scores capture neither of these sources of variation, but instead salvage some portion of the p x i interaction, as shown in the Venn diagram in the bottom panel.

Figure 2. A person by task (with items nested within tasks) by occasion data matrix.

REFERENCES

Agari, T. T. (1979). Individual differences in visual processing of nonverbal shapes. Unpublished Masters Thesis. Seattle: University of Washington.

Anastasi, A. (1986). Intelligence as a quality of behavior. In R. Sternberg & D. Detterman (Eds.). What is intelligence? Contemporary viewpoints on its nature and definition (pp. 19-21). Norwood, NJ: Ablex.

Bakan, D. (1973). On method. San Francisco: Jossey-Bass.

Butcher, H. J. (1968). Human intelligence: Its nature and assessment. London: Methuen.

Cantor, N. (1990). From thought to behavior: "Having" and "doing" in the study of personality and cognition. America Psychologist, 45, 735-750.

Carroll, J. B. (1990). Estimating item and ability parameters in homogeneous tests with the person characteristic function. Applied Psychological Measurement, 12, 109-125.

Carroll, J. B. (1993). Human cognitive abilities. Cambridge, UK: Cambridge University Press

Cattell, R. B. (1966). Handbook of multivariate experimental psychology. Chicago: Rand McNally.

Cooper, L. A. (1982) Strategies for visual comparison and representation: Individual differences. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1., pp. 77-124). Hillsdale, NJ: Erlbaum.

Corno, L. (1993). The best-laid plans: Modern conceptions of volition and educational research. Educational Researcher, 22, 14-22.

Corno, L., & Kanfer, R. (1993). The role of volition in learning and performance. Review of Research in Education, 19, 301-342.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. American Psychologist, 12, 671-684.

Cronbach, L. J. (1977). Educational Psychology (3rd ed.). New York: Harcourt Brace Jovanovich.

Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington.

Darwin, C. (1860) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. New York: D. Appleton and Company.

Dennett, D. C. (1995). Darwin's dangerous idea: Evolution and the meanings of life. New York: Simon & Schuster.

Ericsson, K. A., & Charness, M. (1994). Expert performance: Its structure and acquisition. American Psychologist, 49, 725-747.

Gardner, H. (1985). The minds new science: A history of the cognitive revolution. New York: Basic Books.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton Mifflin.

Greeno, J. G., Collins, A. M., & Resnick, L. (1996). Cognition and learning. In D. Berliner & R. Calfee (Eds.). Handbook of Educational Psychology (pp. 15-46). New York: Macmillan.

Guttman, L. (1971). Measurement as structural theory. Psychometrika, 36, 329-347.

Hernstein, R. (1971). I.Q. The Atlantic Monthly, 228, 43-64.

Hunt, E. (1983). On the nature of intelligence. Science, 219, 141 -146.

Hunt, E., Frost, N., & Lunneborg, C. (1973). Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.). Advances in learning and motivation (Vol. VII, pp. 87-122). New York: Academic Press.

Kane, M. T. (1982). A sampling model for validity. Applied Psychological Measurement, 6, 125-160.

Kuhl, J., & Beckmann, J. (Eds.) (1985). Action control from cognition to behavior. New York: Springer-Verlag.

Linn, R. L., & Gronlund, N. E. (1995). Measurement and assessment in teaching (7th edition). Englewood Cliffs, NJ: Merrill.

Lohman, D. F. (1989). Estimating individual differences in information processing using speed-accuracy models. In R. Kanfer, P.L. Ackerman, & R. Cudeck (Eds.) Abilities, motivation, and methodology: The Minnesota symposium on learning and individual differences (pp. 119-156). Hillsdale, NJ: Erlbaum.

Lohman, D. F. (1994). Component scores as residual variation (or why the intercept correlates best). Intelligence, 19, 1-12.

Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

Mayr, E. (1982). The growth of biological thought: Diversity, evolution, and inheritance. Cambridge, MA: Harvard University Press.

Messick, S., & Kogan, N. (1963). Differentiation and compartmentalization in object-sorting measures of categorizing style. Perceptual and Motor Skills, 16, 47-51.

Nichols, R. C. & Holland, J. L. (1963). Prediction of the first year college performance of high aptitude students. Psychological Monographs, 77 (no. 7), Whole Number 570.

Pepper, S. C. (1942). World hypotheses, a study in evidence. Berkeley: University of California Press.

Rorer, L. G., & Widiger, T. A. (1983). Personality structure and assessment. In M. Rosenzweig & L. Porter (Eds.). Annual review of psychology (Vol. 34, 00. 431-465). Palo Alto, CA: Annual Reviews.

Snow, R. E. (1994). Abilities in academic tasks. In R. J. Sternberg & R. K. Wagner (Eds.), Mind in context: Interactionist perspectives on human intelligence (pp. 3-37). Cambridge, UK: Cambridge Univ. Press.

Snow, R. E., Corno, L., & Jackson, D. (1996). Individual differences in affective and conative functions. In D. Berliner & R. Calfee (Eds.). Handbook of Educational Psychology (pp. 243-310). New York: Macmillan.

Shepard, R. N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171, 701-703.

Spearman, C. E. (1904). "General intelligence" orjectively determined and measured.
American Journal of Psychology, 15, 201-293.

Spencer, H. (1873). The principles of psychology (2nd edition). New York: D. Appleton and Company.

Sternberg, R. J. (1985). Beyond IQ: A triarchic theory of human intelligence. Cambridge, England: Cambridge University Press.

Sternberg, R. J. (1990). Metaphors of mind: Conceptions of the nature of intelligence. Cambridge, UK: Cambridge University Press.

Thorndike, E.L. (1913). Educational psychology: Vol. 2. The psychology of learning. New York: Teachers College, Columbia University.

Thorndike, E. L. (1921). Intelligence and its measurement: A symposium. Journal of Educational Psychology, 12, 124-127.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. Psychological Review, 8, 247-261.

Thurstone, L. L. (1937). Ability, motivation, and speed. Psychometrika, 2, 249-254.

Widgor, A. K., & Garner, W. R. (Eds.) (1982). Ability testing: Uses, Consequences, and Controversies. Part 1. Report of the Committee. Washington, DC: National Academy Press.