Fluid intelligence, inductive reasoning, and working memory: Where the theory of Multiple Intelligences falls short

David F. Lohman
The University of Iowa
March 2001

Paper presented at the annual meeting of the American Educational Research Association in Seattle, WA

An earlier version of this talk also appears in:
N. Colangelo & S. Assouline (Eds.), Talent Development IV: Proceedings from the 1998 Henry B. & Jocelyn Wallace National Research Symposium on talent development. Scottsdale, AZ: Gifted Psychology Press.

Abstract

In this paper, I argue that empirical investigations of human abilities show that abilities are not only correlated, but also organized hierarchically. Of particular import was Gustafsson's (1988) demonstration that General ability (G) is largely synonymous with General Fluid ability (Gf) which in turn is a stand-in for Inductive Reasoning ability (IR). Theories of human abilities that ignore the central role of inductive reasoning are fundamentally flawed. Cognitive research suggests that the ability to maintain, transform, and coordinate information in working memory is at the core of reasoning abilities. The Theory of Multiple Intelligences (Gardner, 1983) misleads because it denies the existence of a central working memory and the importance of inductive reasoning abilities. It turns out, however, that these are two aspects of the same thing.

In Julius Caesar, Shakespeare observed that there is "a tide in the affairs of men," and presumably of women too. Scientific theories are one such class of affairs. However, the rate at which theories ebb and flow varies across disciplines. In the physical sciences, change in established paradigms proceeds at a pace that is frantic at the base but glacial in the aggregate. When they do occur, however, changes in established paradigms can be cataclysmic. Social scientists live in an even less stable world. Theories sometimes rise, fall, and rise again in new garb several times within the span of a single career. But the world of the social scientist cannot even begin to compare with that of the practitioner. Fads come and go like waves on a choppy sea. Some practitioners gladly ride each new wave as it comes along, ever hopeful that it might be the one that will carry all closer to the shores of Utopia. Their initial enthusiasm for the theory is boundless, but their rejection can be equally abrupt when hopes are dashed. Other practitioners, often an older and more seasick lot, cling grimly to the gunwales of their smaller and more weather-beaten boats, hoping to survive each new wave with the least upset.

Theories of human abilities, and the responses of educators to them, show both the relatively slow ebb and flow that characterizes other theories in the social sciences and more abrupt cycle of enthusiasm and disenchantment that attends other popular movements. For example, biologically grounded explanations of human differences waxed strong in the years after Darwin until well in to the twentieth century. Then Boaz in anthropology, Ward in sociology, and Watson in psychology championed the cause of culture and environment. This view gained hegemony and held sway until well past mid century. Biological explanations for human differences did not rise strongly again until more recent advances in brain sciences and genetics (see Degler, 1991, for a readable summary). At the level of popular theory, early views of intelligence –- abetted by overstatements from zealous enthusiasts -- were strongly hereditarian. By the late 1960's, however, biologically based explanations had been completely discarded. In significant measure, this one-sided enthusiasm or rejection comes about because the enthusiast is rarely willing to consider – much less search for – ways in which the theory is inadequate or even wrong.

Many early theories of human intelligence also emphasized the multiplicity and independence of human abilities (Thorndike, 1914; Kelley, 1928). The pendulum then swung to the side of those who emphasized a unitary intelligence (Spearman, 1927; Terman & Merrill, 1937), and oscillated back and forth over the years. British theorists (and those who administered individual intelligence test on both sides of the Atlantic) emphasized a unitary intelligence, whereas American theorists emphasized multiple abilities. Some rapprochement was achieved when both sides agreed on the value of a hierarchical model. Advocates of the unitary view have always been happier with this solution than advocates of the multiple view. A hierarchical model gives precedence to the general factor. What is left over is divided among broad group factors. Once these have had their say, then the primaries are admitted.

Those interested in understanding what abilities might be have always been troubled by this. The primary factors at the base of the hierarchy are invariably psychologically more transparent than higher-order factors. In fact, psychological clarity decreases as one moves up the hierarchy. There is far greater agreement on what phonemic awareness might be than what verbal ability might be; and greater agreement on what verbal ability might be than about general crystallized ability (Gc); and there is least agreement on what G might be. A hierarchical model thus gives parsimony precedence over psychological clarity.

It is of some import, then, when one claims that it is possible to understand the topmost factor (G) not just by understanding a broad group factor at the second level (Gf) but by a "primary" factor [Inductive Reasoning (IR)]. Yet, this is exactly what Gustafsson, (1988) now argues: G = Gf = IR. Neither G nor Gf are seen as abstractions of some sort, but as "stand-ins" for inductive reasoning. One may quibble about whether the relationship between factors is one of identity or of an almost perfect overlap. Either way, it says we can both understand essential aspects of what "G" and especially "Gf" might be and measure them clearly by understanding and measuring inductive reasoning abilities.

The Theory of Multiple Intelligences

Why is this of any import in the current clime? Of late, the most popular theory of human abilities has been the Multiple Intelligence (MI) view of Gardner (1983). The original presentation of the theory of multiple intelligences artfully combined information from a half dozen fields into an engaging set of essays on the initial set of seven intelligences. Each account was liberally illustrated with events from the lives of famous individuals or clinical cases. Indeed, insightful psychobiography seems Gardner's particular forté.

MI theory has always been more popular with lay readers and practitioners than with either cognitive or differential psychologists. Cognitive psychologists were troubled by the theory's extreme modularity, particularity the claim that a central working memory was unnecessary (Messick, 1992). Rather, each "intelligence" was seen as having its own working memory. This may sound like a relatively minor point, but it is not. Working memory is a pivotal construct in all models of cognition. Theorists debate how working memory functions, and what sort of modality-specific slave systems it might have. But most do not debate whether it is a useful or needed construct.

Differential psychologists, on the other hand, were troubled by Gardner's dismissal of 80 years of research on the organization of human abilities (Carroll, 1993; Gustafsson & Undheim, 1996). The reason given is that, as Gardner sees it, evidence for G is provided almost entirely by short-answer multiple choice, paper-and-pencil tests of the sorts of linguistic and logical intelligence that are at best useful for predicting success in the narrow domain of conventionally structured schools (Gardner, 1993, p. 39). Yet even the most cursory examination of human abilities literature shows that every one of these claims at best overstates and at worst is simply false. But there is enough truth to them, and enough dissatisfaction with standardized tests of all sorts that even those who were troubled by the claims were willing to wait and see what sort of evidence would be produced by the new assessment procedures Gardner advocated. They are still waiting.

Lessons learned from performance assessments in other areas of education will probably hold here. We have found that so-called authentic tests have a beneficial effect on the curriculum and can measure aspects of knowledge and skills not tapped by surrogate measures. But we also have discovered that there is vastly more overlap with conventional tests than difference from them, that performance measures are less reliable, more time consuming to administer and score, and vastly more expensive. Most importantly, there is little reason to think that performance assessments show no overlap with each other, and thus support the notion of an independent set of intelligences.

Gardner's theory appeals to teachers and parents because it reinforces the idea that at root all children are special, and that giftedness is a multidimensional—not undimensional—affair. Surely both of these are noble goals. However the theory also appeals the perverse human tendency to think categorically rather than probabilistically. To see two types or seven types of people in the world is surely an advance over seeing one type. But typologies mislead more than they lead when the underlying structure in the domain is not categorical. Indeed, the critical failure of Gardner's theory is not just that it fails to acknowledge or explain why abilities are correlated, but why this correlational structure is implies a hierarchy. In other words, it does not explain why some "intelligences" are more intelligent than others.

Abilities as hierarchically ordered

An example may help. Many phenomena that we treat as categorical are at root continuous. There is invariably some loss in precision when a continuum is described by a series of categories. The loss is not fatal and may even be helpful in that it enables us to communicate with one another using category labels. Thus, we speak of wealthy and poor, more able and less able, and of green, blue, and violet rather than of the continuous scales that underlie each. Color perception is also a good example of how correlation need not imply hierarchy. If we ask people to rate the similarities among colors that span the visible spectrum, we invariably find that some colors are perceived as more similar than others e.g., blue is perceived as more similar to violet than it is to yellow. If we then treat these ratings as distances with the simple transformation that greater similarity implies smaller distance, then we can plot all of the colors in a single space. When we do this we find that the ratings scale as a circle (or, more formally, a circumplex). This is shown in Figure 1. Colors are related to one another. But none is superordinate. Blue is no more or less important than yellow.

This is not what happens when we scale scores on different ability tests—even when abilities are defined as narrowly as a particular cell in Guilford's (1985) elaborate scheme. Rather, what we find is that some tests invariably fall near the center of the plot whereas others fall near the periphery. In other words, some abilities are more general, more important, and more pervasive than others. Both scaling studies (like that illustrated for colors) and factor analyses show that tests that require reasoning, or reflect the products of past reasoning, fall in the central parts of the scale, or load most highly on the general fluid ability (Gf) factors. Figure 2 shows a summary of the scalings obtained for many different ability tests and learning tasks to different groups of people. Note that tests are not all equidistant from the center of the plot. Tests near the center are a variegated lot. Some are difficult and lengthy tests (Raven Progressive Matrices). Others are short (a 12 item Letter Series Test). Some are verbal, some figural, and some quantitative. But all require reasoning.

A good example is the Test Necessary Arithmetic Operations. This test was devised by Guilford to measure a specific cell in his Structure of the Intellect Model. Each item presents a short word problem. The examinee's task is not to solve the problem, but to say which two operations she would use, and in what order. There are four operations: add, subtract, multiply, and divide. Thus, problems do not require advanced mathematics. Yet in the sample of over 100 Stanford undergraduates who were administered most of the tests in Figure 2, Necessary Arithmetic Operations had one of the highest loadings on the Gf factor (Marshalek, Lohman, & Snow, 1983). How can we explain why this test (or indeed some of the others near the center of Figure 2) behave as they do? They do not simply combine several primaries, as Thurstone (1938) hypothesized. In other words, they are not simply mixtures of pure types. They are not just “linguistic” and “mathematical” -- as Gardner (1993) hypothesized. Figural tasks such as the Wechsler Block Design test or Thurstone’s Paper Folding test fall within the innermost circle. Nor are they particularly speeded, as Gardner also hypothesized. In fact, all are more like power rather than like speed tests. The simplest explanation is that all require reasoning.

Sternberg (1986) came to a similar conclusion after reviewing a different literature on cognitive tasks:

> An interesting finding that emerges from the literature attempting to relate cognitive task performance to psychometrically measured intelligence is that the correlations of task performance and IQ seems to be a direct function of the amount of reasoning involved in a given task, independent of the paradigm or label given to the paradigm.... Thus, reasoning ability appears to be central to intelligence. (pp. 309-310)

Reasoning and working memory

But why should reasoning tests be so important? Examination of patterns of correlations among test scores can only take us so far. We learn that abilities are correlated, not independent, and that the pattern of correlations implies a hierarchy. But correlational analyses cannot tell us why reasoning abilities run from the base to the apex of the hierarchy. Cognitive psychology, however, can offer more direct evidence on this question. A first piece of the puzzle, I believe, is that reasoning tasks place extraordinary demands on the management of attentional resources in working memory. Many researchers have claimed that a major source of individual differences on reasoning tasks lies in how much information one must maintain in working memory, especially while effecting some transformation of that information (Holzman, Pellegrino, & Glaser, 1980). More recently, Kyllonen and Christal (1990) found that latent variables for reasoning ability and working memory correlated approximately $r = .8$ in four large studies. These studies surprised many because working memory was noticeably absent from models that posited a series of consecutively executed component processes (see, e.g., Sternberg, 1984). But as Kyllonen and Christal (1990) noted, most of the performance processes (such as encoding and inference) and executive processes (such as goal setting, goal management, and monitoring) in Sternberg's information processing models of reasoning are presumed to occur in working memory. Thus, even though, say, the inference process may be effective, it must be performed within the limits of the working memory system. Therefore, although many different processes may be executed in the solution of a task, individual differences in them may primarily reflect individual differences in working memory resources.

Newer theories of working memory differ importantly from the older concept of short-term memory. A major difference lies in the relative emphasis on passive storage in the older short-term memory model versus more effortful, controlled processing of information that is also being maintained in an active state in the newer working memory models. Some see this in terms of a tradeoff between processing capacity and storage capacity (Daneman & Carpenter, 1980), whereas others view it in terms of different memory systems. For example, Baddeley (1986) posits a working memory with a storage component and a separate executive (or supervisory attentional system) that attends selectively to one stimulus while inhibiting another, coordinates performance in tasks, and switches strategies (Baddeley, 1996). When working memory is interpreted in this way, studies that find high correlations between working memory and reasoning seem less astonishing. In fact, working memory tasks in the Kyllonen-Christal studies were specifically designed to reflect Baddeley's theory, and thus required both storage and transformation (although the former was presumed to be more difficult).

Therefore, reasoning tasks burden working memory because they require that one simultaneously remember and transform information. This not only taxes attentional capacity, but also requires that one monitor fading memory traces and refresh them strategically. But there is more. Effective storage, retrieval, and transformation of information in working memory in the service of solving a task requires that one have a plan of attack. This often means assembling a preliminary strategy on the basis of task directions and experience on the first few items, and then modifying this strategy as new items are encountered. Assembly and control processes are thus the second (and third) pieces of the puzzle.

Assembly processes are reflected in activities in which an individual must organize a series of overt acts or covert cognitive processes into a sequence. This is an especially important activity on novel or ill-structured tasks. In their detailed analysis of the Raven Progressive Matrices test, Carpenter, Just and Schell (1990) argued that crucial task demands were the ability to decompose a complex problem into simpler problems, and the ability to manage the hierarchy of goals and subgoals generated by this decomposition. A “hierarchy of goals” is just another way to say “systematic plan for solving problems.”

Control processes are more diverse, although all involve the ability to monitor the effects of one's cognitions and actions. This affords the possibility of adjusting strategy or behavior on the basis of feedback from the environment or one's body. But the ability to monitor and adapt thinking also burdens working memory, especially when several ideas must be considered simultaneously or when goal images differ from images activated by perceptions.

The last piece of the puzzle is learning. Randomly ordering items on ability tests can make them poorer measures of Gf. Several experiments now suggest that randomization may reduce the probability that one can learn rules on easy items that will later be needed on harder items. In part, then, ability tests are themselves inductive-learning tasks.

This is but a partial account. Elsewhere (Lohman, 2000) I have summarized the importance of other hypotheses that have been advanced to explain the nature of Gf. Most (such a mental speed or task complexity, for example) contribute importantly, but by themselves cannot explain the whole. Binet (as translated by Terman, 1916, p. 45) came closer to this view than any other early theorist when he defined intelligence as “the tendency to take and maintain a definite direction; the capacity to make adaptations for the purpose of attaining a desired end; (and) the power of autocriticism.”

Summary

It is probably no coincidence that the chief complaint of the cognitive scientist about Gardner's theory (i.e., its rejection of a central working memory) turns out to be the basis for the differential psychologists complaint, at least with respect to the need for a higher-order G or Gf factor. An extreme modular view of both working memory and intelligence fails to explain the empirical observations in both domains.

Surely general inductive reasoning or fluid analytic abilities are not the whole of human cognitive abilities, and cognitive abilities are not the whole of human character. But there is a large and important difference between being tone deaf or physically uncoordinated or socially inept, and being mentally retarded. We can value different excellencies, but we need not equate them to do so. Therefore, inductive reasoning abilities will be pivotal in the identification and nurture of giftedness long after the waves of modularity have receded.

References

Baddeley, A. D. (1986). Working memory. Oxford: Clarendon Press.

Baddeley, A. (1996). Exploring the central executive. The Quarterly Journal of Experimental Psychology, 49A(1), 5-28.

Carpenter, P. A., Just, M. A., & Schell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices test. Psychological Review, 97, 404-431.

Carroll, J. B. (1993). Human cognitive abilities. A survey of factor-analytic studies. Cambridge, UK: Cambridge University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Mathematical Behavior, 19, 450-466.

Degler, C. N. (1991). In search of human nature: The decline and revival of Darwinism in American thought. New York: Oxford University Press.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Gardner, H. (1993). Multiple intelligences: The theory in practice. NY: Basic Books.

Guilford, J.P. (1985). The structure-of-intellect model. In B.B. Wolman (Ed.), Handbook of intelligence (pp. 225-266). New York: Wiley.

Gustafsson, J. -E. (1988). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 4, pp. 35-71). Hillsdale, NJ: Erlbaum.

Gustafsson, J. -E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner & R. C. Calfee (Eds.), Handbook of educational psychology (pp. 186-242.) New York: Simon & Schuster Macmillan.

Holzman, T. G., Pellegrino, J. W., & Glaser, R. (1982). Cognitive dimensions of numerical rule induction. Journal of Educational Psychology, 74, 360-373.

Kelley, T.L. (1928). Crossroads in the mind of man. Stanford, CA: Stanford University Press.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! Intelligence, 14, 389-433.

Lohman, D. F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Ed.), Handbook of Intelligence. New York: Cambridge University Press.

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. Intelligence, 7, 107-128.

Messick, S. (1992). Multiple intelligences or multilevel intelligence? Selective emphasis on distinctive properties of hierarchy: On Gardner's Frames of Mind and Sternberg's Beyond IQ in the context of theory and research on the structure of human abilities. Psychological Inquiry, 3, 365-384.

Terman, L. M. (1916). The measurement of intelligence. Boston: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1937). Measuring intelligence. New York: Houghton.

Thorndike, E. L. (1914/1970). Educational psychology. Volume III. Mental work and fatigue and individual differences and their causes. Westport, CN: Greenwood Press.

Shepard, R. N. (1962). Analysis of proximities: Multidimensional scaling with an unknown distance function. II. Pyschometrika, 27, 219- .

Snow, R. E. & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. Linn (Ed.) Educational Measurement (3rd ed., pp. 263-332). New York: American Council on Education/Macmillan Publishing Company.

Spearman, C. E. (1927). The abilities of man. London: Macmillan.

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

Sternberg, R. J. (1986). Toward a unified theory of human reasoning. Intelligence, 10, 281-314.

Thurstone, L. L. (1938). Primary mental abilities. Psychometric Monographs, 1.

Figures

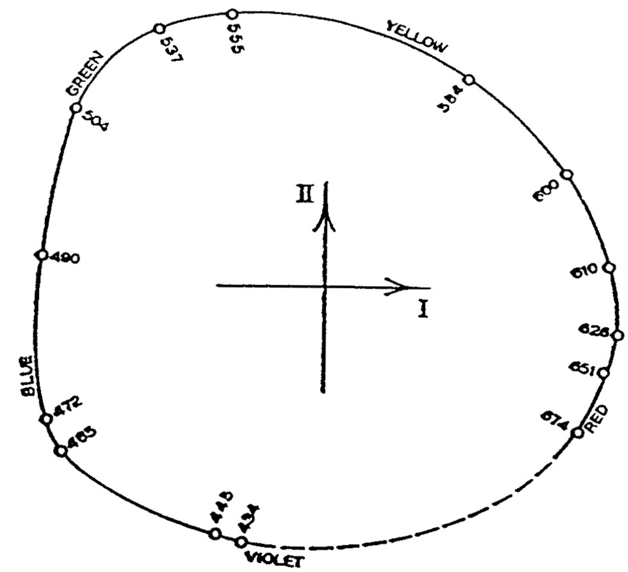

Figure 1. A two-dimensional scaling of similarity ratings for fourteen colors varying in hue (from R. N. Shepard, “Analysis of proximities: Multidimensional scaling with an unknown distance function. II,” Pyschometrika, 1962, p. 236.)

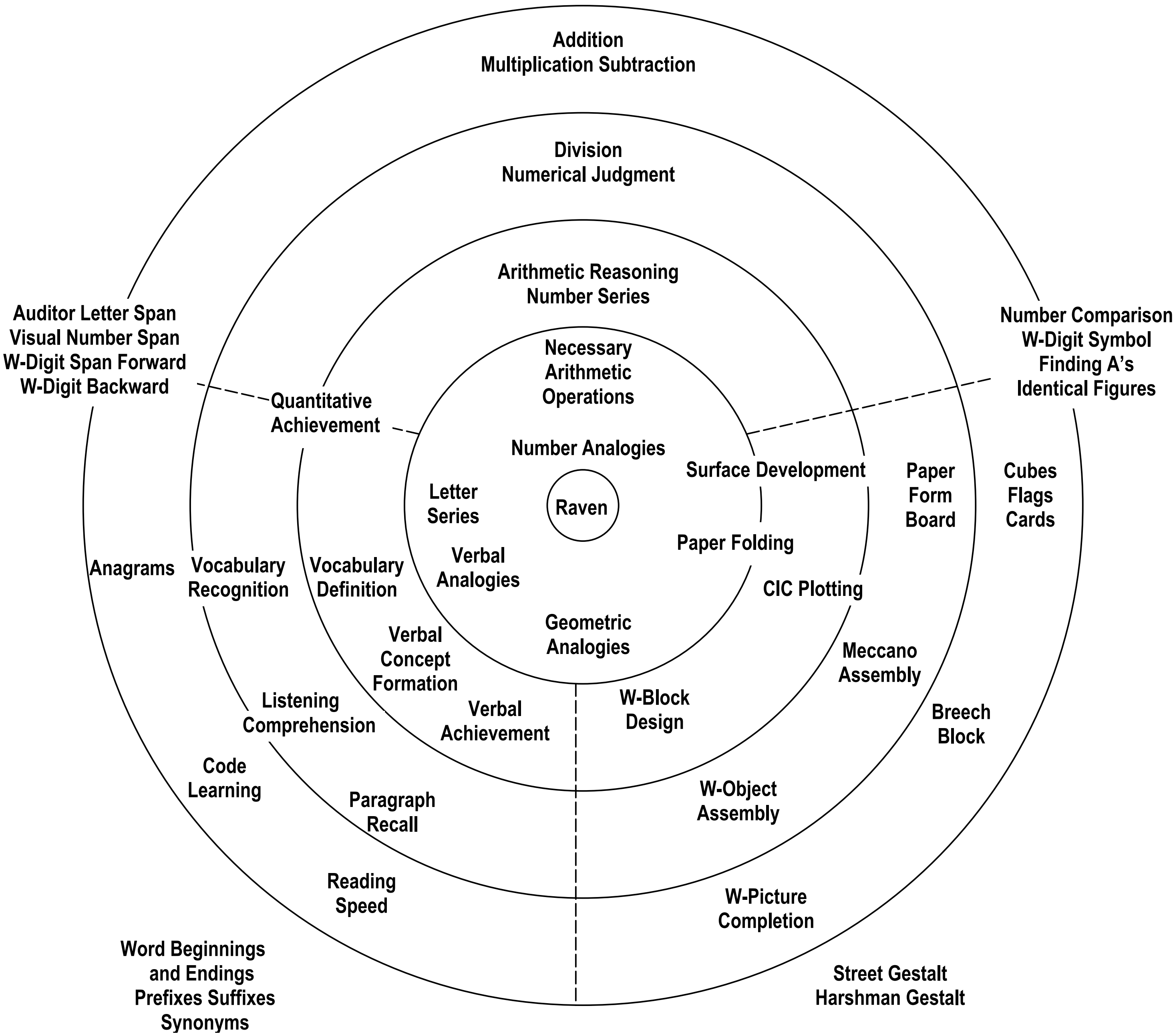

Figure 2. A summary dimensional scaling of various ability tests and learning tasks. “W” identifies a subtest of the Wechsler Adult Intelligence Scale (from R. E. Snow and D. F. Lohman, “Implications of cognitive psychology for educational measurement.” In R. Linn (Ed.), Educational Measurement (3rd ed.), 1989, p. 318)