COGNITION AND INSTRUCTION, 1991, 8(1), 29-67

Meanings of Skill: Effects of Abacus Expertise on Number Representation

Kevin F. Miller
University of Illinois at Urbana-Champaign

James W. Stigler
University of Chicago

Research on expertise has focused on the amount and organization of domain-relevant knowledge as the key feature distinguishing experts from novices. The representations of experts are described as being more functionally organized as well as more detailed than those of novices. There are two different senses in which knowledge could become more functionally organized with expertise. The knowledge of experts could become more functionally organized with respect to the execution of skill, or their knowledge could become more functionally oriented with respect to contexts broader than their particular skill. Abacus skill provides a domain in which these two alternatives can be evaluated. The first alternative, conceptual determination by skill, implies that abacus experts' representations of numbers should emphasize the features that are functional in their calculation strategies. The second alternative, conceptual transparency, implies that experts should not emphasize abacus-specific features when judging similarity of numbers, giving less weight than novices to abacus features when judging abacus representations and no more weight than novices to abacus features when judging numerals.

Conceptual transparency occurs when those features important in performing a skill are judged to be of little significance in understanding the skill. Number similarity judgments of abacus experts and novices from Taiwan and American children of comparable ages indicted that abacus skill was associated with a deemphasis of those features that are unique to the abacus. Implications of the phenomenon of conceptual transparency for models of skill development are discussed.

Acquiring a complex skill typically requires one to master a substantial amount of detailed knowledge, much of which is only relevant to a particu-

Requests for reprints should be sent to Kevin F. Miller, Department of Psychology, University of Illinois, 603 East Daniel, Champaign, IL 61820.

lar domain. This knowledge often takes the form of specialized mental representations. Previous studies of expertise have focused on describing the nature and functioning of the mental representations that make skilled performance possible. The research described here concerns a different but related question: What effect does expertise have on the perceived meaning of the specialized representations of knowledge that often underlie skilled performance?

Representational systems vary in the emphasis they place on different kinds of relations, and specific representations are often chosen to facilitate specific goals. For example, with binary numerals, it is simple to determine whether a number is divisible by 2 but difficult to determine whether it is divisible by 10. With Hindu–Arabic numerals, the situation is reversed. This relative accessibility of the modulo-10 value of a Hindu–Arabic numeral, however, is not necessarily conceptual but only procedural or functional: *Representations* form part of a system that functions in a goal-directed fashion, and their meaning arises out of their relation to the other elements of the system. *Conceptual meaning,* by contrast, refers to the relationship between the specialized representation and the wider domain of knowledge over which the skill operates.

For example, consider the paper-and-pencil algorithms school children learn for carrying out basic computations. These algorithms obviously have meaning in the sense that they work. Yet the degree to which they are imbued with conceptual meaning will vary greatly across individuals. Some individuals may not draw any connections at all between a particular algorithm and the mathematical principles on which it is constructed. Others may utilize their conceptual knowledge to understand why the algorithm works, utilize the fact that the algorithm works to enrich their conceptual knowledge, or do both.

Studying the conceptual consequences of skill has not been the focus of recent research, in part because of the domains in which expertise has been studied. For example, experts in chess and other games acquire and use mental representations specific to their domain of expertise. One cannot talk easily, however, about the wider significance of these representations outside the context of skilled performance, because games such as chess are not intended to relate to some larger domain. For other skills, such as mastering an instrument or learning a calculating technique, the question of how the skill relates to a larger domain (e.g., music or mathematics) is substantially more salient.

The current study was organized around two questions. The first concerns what relation there is between the features that organize performance of a skill and those that are salient in thinking about the domain over which skill operates. The second deals with how these relations change as a function of expertise. This study investigated these issues within the context of skill at abacus calculation. Experts at abacus calculation are able to perform rapid mental calculation by using a "mental abacus," a mental image that functions like an abacus (Hatano, Miyake, & Binks, 1977; Stigler, 1984). The representation is concrete, visual, and highly specialized. The question that motivates the study is the degree to which connections are made, with expertise, between the abacus and conceptual knowledge about numbers. After reviewing research on the kinds of knowledge that experts possess, two very different predictions concerning the conceptual significance of skill are discussed. These predictions are then applied to explaining the effects of mastering mental abacus calculation on children's representation of numbers.

DOMAIN-SPECIFIC KNOWLEDGE AS THE DISTINCTIVE FEATURE OF EXPERTISE

A remarkable degree of consensus has emerged around the view that acquisition of domain-specific knowledge is the distinctive feature of expertise. The list of domains to which this generalization applies is now a substantial one. In chess, for example, de Groot (1965) found that chess masters could retrieve meaningful positions involving 20 or more pieces after only 5 sec of exposure. Novices, by contrast, recalled an average of fewer than five pieces under the same conditions. This impressive memory ability did not, however, extend to random chess configurations. Although novices and experts performed equally well on random chess patterns, it seems likely that expert performance is affected by configurations that are impossible, whereas novices view these configurations as simply another pattern. Nonetheless, the basic finding that experts possess an impressive memory for meaningful configurations in the domain of skill has since been demonstrated with domains as diverse as electronic circuit diagrams (Egan & Schwartz, 1979); computer programming (Adelson, 1981; McKeithen, Reitman, Rueter, & Hirtle, 1981); and the games of *Go* (Eisenstadt & Kareev, 1975; Reitman, 1976), bridge (Charness, 1979; Engle & Bukstel, 1978), and baseball (Chiesi, Spilich, & Voss, 1979). Experts have more domain-relevant knowledge, and one estimate of the magnitude of this data base suggests that skill in any complex domain is built on a substantial amount of domain-specific information. For the case of chess, Simon and Chase (1973) estimated that chess experts can recognize on the order of 50,000 different configurations.

In addition to the sheer quantity of information experts have mastered, research in several domains suggests that the knowledge of experts is organized differently from that of novices. Chi, Feltovich, and Glaser (1981) reported that physics experts encoded mechanics problems according to the underlying physical principles (e.g., conservation-of-energy problems), whereas novices categorized the same problems in terms of the problems' surface structure (e.g., rotating things or blocks on inclined planes). Allard and Burnett (1985) found a similar difference in the ways that novice and expert basketball players categorized pictures of games, with novices focusing on the number of players present and experts focusing on the strategic implications of the scenes. Chi has shown that expert children's knowledge of familiar dinosaurs is largely organized in terms of implicit features such as dietary preferences, whereas these same children's knowledge of unfamiliar dinosaurs (Chi & Koeske, 1983) or the representations of nonexperts (Gobbo & Chi, 1986) do not seem to rely on these implicit features. Thus, in a variety of domains, the organization of expert knowledge appears to reflect an appropriate "deep structure," whereas novices are often captivated by superficial "surface structures."

CONCEPTUAL IMPLICATIONS OF SKILL

Given that expertise involves the acquisition of a large knowledge base organized by relations functional for the specific skill, it would be odd, indeed, if the knowledge base described in studies of expert memory did not affect the underlying conceptual representation of the domain in which the skill is based. Yet it remains unclear how expertise ought to alter conceptual representation. Based on the research discussed earlier, one would expect that representations of experts would differ from those of novices in the amount and organization of relations represented, with expert knowledge reflecting the functional deep structure of the domain.

On closer examination, however, the assertion that expert knowledge is more functional than that of novices becomes ambiguous for many real skills. Consider the problem of technical definitions, that is, familiar terms given to idiosyncratic definitions within a domain. Such new uses for familiar words can constitute a real stumbling block in learning a new skill. For example, McKeithen et al. (1981) noted that learning a programming language involves deemphasizing common meanings of terms (e.g., the similarity in meaning between *true* and *real*) that do not carry over into their use in programming. Where those relations that are functional for some skill contrast with relations that are meaningful in some broader context, how does the expert's representation differ from the novice's?

Even if expertise is associated with an increasing relation between skill and knowledge representations, one can still make two different predictions about the effects of expertise on representation, corresponding to two different views of the meaning of expertise. According to one view, expertise is a matter of accommodating to or getting deeply into some domain, whereas, in the other view, experts have worked their way out of the constraints that may be peculiar to a particular skill. Both views are consistent with the notion that expert knowledge has a functional form, but they vary in the nature of the function in which that knowledge is organized to serve. It is important to emphasize that either result is consistent with the view that expert knowledge utilizes a deep-structure representation relative to that of novices. The critical issue is how domain specific such a deep structure is. Conceptual determination by skill implies that the relevant deep structure is that which functions within a domain. Conceptual transparency of skill implies that the relevant deep structure is one that generalizes beyond the domain of expertise. Each view has a distinguished history. The phenomenon of expert abacus calculation, explored here, provides a natural domain for evaluating these very different views of the conceptual consequences of acquiring expertise.

Conceptual Determination by Skill

The view that expertise involves the extension of those relations functional in performing the skill was most eloquently described by Bryan and Harter (1899). Discussing the process of acquiring skills such as telegraphy, Bryan and Harter wrote: "In the measure that he has mastered the occupation, it has mastered him. Body and soul from head to foot, he has—or one may say he is—the array of habits which constitutes proficiency in that sort" (p. 348).

According to this view, a mastered skill should color one's perception of all domains it touches. Those relations relevant to performance of the skill should be perceived as being of general importance. Thus, a skilled pianist should view music through the filter of those features relevant to piano playing. A skilled abacus calculator should represent numbers by those features that are salient for abacus calculation, seeing the abacus in every number. Given that acquisition of skill in such domains as chess or mental abacus calculation requires hundreds of hours of practice, it should not be surprising to find that experts extend those concepts and relations that are functional within the context of their skill to those where it is not directly relevant.

Specifically, the notion of conceptual determination by skill predicts that those components that are important for performing a skill will also be important in one's conceptual understanding of the domain over which skill operates.

Conceptual Transparency of Skill

Yet one might also expect just the opposite effect of expertise on conceptual representation. Mastering a skill may make it possible to move beyond the constraints of those features that are uniquely relevant to the skill. If the knowledge of experts were truly functional, experts might deemphasize those features relevant only to their domain of specific skill with other relevant knowledge, transcending the limitations imposed by their skill. Hatano (1982) discussed a similar issue in distinguishing between two kinds of experts, termed *routine* and *adaptive,* based on the conceptual basis for skill. Although both kinds of experts may perform at high levels in familiar situations, only the adaptive expert understands the basis for a skill and is able to adapt familiar procedures to new circumstances. To the extent that one develops adaptive expertise within a domain, skill should be associated with a transcendence of those organizations peculiar to a given skill.

This view that expertise leads one to move beyond reliance on features uniquely relevant to a skill also has a distinguished history. In his study of blindfold chess, Binet (1893/1966) asserted that developing expert knowledge of chess involves moving beyond an emphasis on the concrete features of chess pieces and their spatial relations. Rather than possessing a great deal of detailed knowledge about the meanings of particular chess configurations, Binet argued that it is insight into the strategic structure of games that makes chess skill possible. Describing the expert's skill, Binet wrote:

> He remembers not that he moved his king to a certain square, but that, at a given moment, he had a particular plan of attack and defense which required the movement of his king. The move itself is only the conclusion of an act of thinking; that act must first be recaptured; the recall of its manifest result—the particular move—follows from it. (p. 149)

According to this view, the musical representation of the expert pianist might deemphasize those features unique to piano playing. The abacus calculator would represent numbers according to those features relevant to mathematics in general, in effect seeing numbers in every abacus instead of an abacus in every number. This possibility can be described by the term *conceptual transparency.* Conceptual transparency implies that features, information, and knowledge utilized in performing a skill are nonetheless deemphasized or ignored in thinking about the same domain. It does not imply that such knowledge becomes inaccessible (although it might become so) but, rather, that those features demonstrated to be important in organizing skilled performance are not important in thinking about the domain over which skill operates. Indeed, one way in which domain-specific knowledge might become relatively less prominent over time is through the addition of domain-general knowledge to the set of facts one knows about a domain.

Assessing these two different views on the conceptual consequences of skill requires a domain of skill that involves mastering an alternative way of representing a familiar domain of knowledge. To the extent that one can describe skill-related features in subjects' representations of a domain, it will be possible to evaluate whether expertise is associated with extension or deemphasis of the special relations associated with skill. It should not be assumed that the conceptual consequences of expertise will be the same in every domain, but relatively straightforward assessments of these two very different views of the meaning of skill can be made for domains of expertise that involve the mastery of special representations.

Previous studies of the conceptual consequences of expertise have not permitted one to make this distinction. Schoenfeld and Herrmann (1982), for example, looked at changes in the categorization of mathematical problems by subjects who took a course in either mathematical problem solving or in structured programming. Students who took the problem-solving course, which focused on Pólya's (1957) heuristics approach, showed a shift toward categorizing problems according to their mathematical structure. This effect was not found for students who took the programming course. The course did not focus on problem perception per se, but it is described by the authors as encouraging students to consider the underlying plausibility of results and to consider carefully the conditions of problems. Thus, it is reasonable that students who become more skilled in solving problems will learn to perceive problems according to those components (the underlying mathematical structures) that determine the solution. Although Schoenfeld and Herrmann's study suggests that mathematical problem-solving expertise has important conceptual consequences, it does not permit one to distinguish between the problem-solving principles that students became expert in and the underlying structure of mathematical problems.

The distinction between knowledge utilized in performing a skill and one's conceptual understanding of the domain over which the skill operates corresponds to part of Marr's (1982) model of levels of description in understanding cognition. Marr argued that one can distinguish among descriptions of computational theory, representation/algorithm, and implementation in describing complex information-processing systems. Computational theory concerns the goals, principles, and logic of the strategy by which processing occurs. The level of representation and algorithm concerns what specific information is to be represented and how it is processed, whereas the implementational level concerns the specific electronic or physiological means by which the representation and algorithm are made silicon or flesh in order to accomplish the ends of the computational theory. Marr's distinction between computational theory and the ways in which it is implemented also suggests that one might expect features important for performing a skill to be of little importance in thinking about the domain. The possibility that expertise might lead to the conceptual transparency of domain-specific knowledge is paradoxical, because experts are distinguished by the facility with which they use such information in performing their skill. On the level of computational theory, however, it seems clear that one's understanding of a domain need not be determined by the features that determine one's performance.

One skill that permits comparison of skill-related versus non-skill-related representations is mental abacus calculation. Abacus experts are expert at abacus calculation. They learn a system that is based on a set of rules that differ in identifiable ways from the base-ten structure of numbers. It is reasonable to suspect that experts utilize the deep structure of abacus calculation, but there are two plausible candidates for the deep structure of abacus calculation. It may be the formal features of the abacus or, instead, some set of numerical features that transcends the device. Both are deep structures. A novice first learning to use the abacus experiences difficulty both in seeing the numerical significance of abacus configurations and also in mastering the structure of the device itself. With practice, knowledge can become deeper in both senses. The structure of the abacus permits one to distinguish between two very different models of the conceptual consequences of expertise.

To the extent that expertise leads to an emphasis on those features that are functional within a given domain of skill, the abacus expert should represent numbers using the features emphasized in abacus procedures. Thus, the representation of numbers of the abacus expert should be more abacus-like than that of the nonexpert. If expertise instead leads to the conceptual transparency of the features that underlie skilled performance, the abacus expert should be less likely than someone at a novice or intermediate level to represent abacus relations in terms of abacus-relevant features and no more likely than novices to represent number relations by abacus-relevant features. The phenomenon of mental abacus calculation as practiced in Taiwan provides a natural instance for studying the consequences of skill for the relations between two alternative means of representing numbers.

THE NATURE OF ABACUS SKILL

Prior research has documented the impressive computational skills developed by adults and children who receive extended practice in abacus calculation (Hatano et al., 1977; Hatano & Osawa, 1983; Stigler, 1984). Perhaps the most intriguing aspect of this skill is the development of mental abacus calculation, in which subjects calculate with reference to an internal image of the abacus. Before describing data supporting these claims, a brief review of how the abacus works is in order.

Figure 1 shows how numbers are represented on the Japanese abacus used throughout Asia. Beads "count" as they are pushed toward the center (horizontal) bar by the thumb (lower beads) or forefinger (upper bead). The upper bead represents five times the column value; the lower beads represent one unit each. The value represented by a column is the sum of the top bead (0 or 5) and the lower bead (0 to 4), with the total multiplied by the column value (as with standard place-value notation). Within a column,

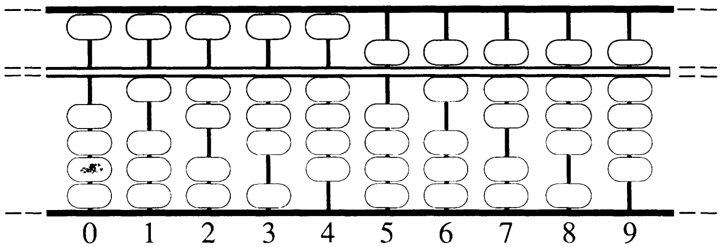


FIGURE 1 Representation of digits on the Japanese-style abacus used in this study. Beads "count" as they are moved toward the center (horizontal) bar. The top bead represents 5, the lower beads each represent 1. Value within a column is the sum of the top bead (0 or 5) plus the number of lower beads (0 to 4) pushed toward the center bar.

the abacus represents a modulo-5 number system but remains a base-ten system between columns. Thus, numbers such as 1 and 6 that differ by ±5 have similar abacus representations, differing only in the placement of the upper bead.

Persons who develop a high level of skill at abacus calculation report calculating with reference to a mental abacus, using an image of the abacus to perform mental arithmetic. Supporting this claim, studies of the mental calculation of abacus experts have found several ways in which the structure of the abacus is reflected in subjects' performance. Stigler (1984) reported that for abacus-trained children the number of steps involved in an abacus calculation was associated with reaction time for mental calculation. He also reported that these abacus-trained children could distinguish true intermediate states from foils. Perhaps the strongest evidence for the abacus-like nature of the mental calculation of these children came from analysis of their errors. Abacus calculators (but not American college students) were apt to make errors that could be accounted for by misrepresenting the location of one bead on the abacus. These included leaving out the value of one column and errors in which the answer was off by five in some column from the correct sum. Because the abacus represents numbers that differ by five in a similar way, the finding of an increased incidence of these modulo-5 errors provides a convincing demonstration of the abacus-like nature of the calculation of those who have become experts at the skill of mental abacus calculation.

The prevalence of modulo-5 errors in the mental calculation of abacus-trained children is important, because it provides clear evidence that the structure of the abacus has impressed itself on the organization of calculation by these children. It also provides an important marker for evaluating the conceptual consequences of abacus skill. The model postulating conceptual determination by skill predicts that abacus experts will find modulo-5 relations more important than do other groups in thinking about relations between numbers. The model postulating conceptual transparency of skill argues that abacus experts should (a) not view these relations as being important in judging similarity between numbers presented as numerals and (b) view modulo-5 relations as being less important than do other groups in judging similarity of numbers presented as abacus figures.

REPRESENTING THE MEANING OF SKILL

Data on the abacus-like nature of the mental arithmetic of abacus experts indicate that the structure of the abacus has a substantial impact on the numerical performance of these experts. The current study addressed whether the importance of the abacus to their calculation shows up in the features of numbers that these experts view as being important.

It is difficult to get individuals to describe their conceptual representation of a domain such as number and far from clear what the significance of such introspections might be (e.g., Ericsson & Simon, 1980; Nisbett & Wilson, 1977). Although subjects have considerable difficulty describing any such global representation, they have much less difficulty judging the similarity among particular numbers (Shepard, Kilpatric, & Cunningham, 1975). These judgments can then be analyzed by a variety of means (Carroll & Arabie, 1983; Miller, 1987), including nonmetric multidimensional scaling (MDS) procedures (Kruskal, 1964a, 1964b; Shepard, 1962a, 1962b) and clustering techniques (Johnson, 1967; Shepard & Arabie, 1979) that reduce matrices of such judgments into more manageable spatial representations or sets of discrete clusters. To the extent that such judgments reflect meaningful features of the domain, analyses of similarity judgments can provide a method for diagnosing those features judged to be central or salient at varying levels of expertise.

One of the first applications of Shepard's original MDS procedure concerned the effects of skill on representation of Morse code. Shepard (1963) analyzed results of prior studies on confusions among Morse-code stimuli by subjects of varying degrees of experience with the system. Although the various studies were not directly comparable, Shepard found that subjects unfamiliar with Morse code used length and the relative proportion of dots and dashes to discriminate signals. Errors by subjects early in the process of learning Morse code were based on length and heterogeneity of stimuli: It was more common to mistake stimuli composed entirely of dots as compared with those made entirely from dashes. Finally, expert errors were no longer affected by length, with items varying only in the length of runs of consecutive dots or dashes frequently confused with each other. These studies of the effects of skill on perception of Morse code suggest that MDS procedures are sensitive to at least some of the representational changes associated with skill. Because of the need to determine separate solutions for each individual or group, however, representational changes associated with skill are not directly measured by the MDS techniques used by Shepard (1963).

A more powerful means of describing changes in representation is provided by the individual differences scaling (INDSCAL) model of Carroll and Chang (1970). This mathematical model entails a psychological assumption about how individual differences affect representation of stimulus domains, namely, that individuals differ by differentially emphasizing common stimulus dimensions. According to this model, individuals differ by differentially weighting the dimensions of a common stimulus space. Because of the related way that INDSCAL treats subject and stimulus variation, programs to fit this model provide separate spatial representations of the overall stimulus relations and of the weights that subjects place on the dimensions of this stimulus space. High subject weights along a stimulus dimension indicate that the corresponding feature was emphasized by a particular subject or group of subjects.

An important distinction between the INDSCAL model and most other MDS models is that, because subject variation is represented by weighting dimensions, INDSCAL solutions are not invariant across rotation through an arbitrary angle, thus providing a unique orientation for the solution up to reflections or permutations of the axes. In other words, the INDSCAL model uses information on variation among *subjects* to uniquely orient a multidimensional *stimulus* space. The dimensions of an INDSCAL solution correspond to those dimensions that capture variation in subjects' judgments, and one might expect these dimensions of variation among subjects to correspond with important features of the stimuli that those subjects were judging.

The INDSCAL model has been used to describe cognitive representations in a variety of domains. Effects of expertise on representation have been explored in several studies of music comprehension. Krumhansl and Shepard (1979) found that increased musical experience was associated with a decreasing emphasis on pitch differences when judging the similarities of notes in scale contexts, with an increasing weight given to differences within octaves. Thus, musical expertise was associated with increased use of the conventional scale structure in judging similarities of notes. Pollard-Gott (1983) asked subjects of varying expertise to judge the similarity among passages from a Liszt sonata after listening to the piece a varying number of times. Repeated listening led to an increasing emphasis on the relation between each passage and the sonata's two themes as the basis for similarity. Pollard-Gott also collected judgments from two expert subjects, one of whom had performed the piece, the other who had written an analysis of it. A single dimension, reflecting the distinction between these two themes, accounted for 84% of the variance for these subjects.

These two studies provide ambiguous support for the two models of expertise described earlier. To the extent that expertise is associated with assimilating the conventional scale structure of music, this development is consistent with Bryan and Harter's view that features of skill are incorporated into conceptual representations of the domain over which skill operates. Pollard-Gott's finding that musical expertise is associated with a shift away from physical features (e.g., loudness) to higher order relations (e.g., that between a segment and the theme) is consistent with Binet's view that expertise involves moving away from details to a higher order structure. The case of abacus calculation provides a clear instance for evaluating these alternative models because of the presence of abacus-specific features critical to the performance of the skill.

THE DEVELOPMENT OF NUMBER REPRESENTATIONS

In the absence of abacus skill, what does the development of children's representations of numbers look like? A baseline against which to evaluate the impact of abacus skill on number representation is provided by Miller and Gelman (1983). In this study, data on perceived similarity of numbers were collected from adults and children aged 5, 8, and 12 years.

Miller and Gelman determined separate MDS solutions for the resulting similarity matrices from each age group. Both the 5- and 8-year-old subjects produced essentially one-dimensional configurations reflecting numerical magnitude, whereas solutions for 12-year-old and adult subjects incorporated additional multiplicative features such as a division into odd and even numbers. Miller and Gelman described the development of children's number representation as a matter of overlaying new numerical relations on an initial representation of numerical magnitude. An individual differences reanalysis of these data using the SINDSCAL (Pruzansky, 1975) program is presented in Figure 2. The overall correlation between predicted values and obtained judgment is $r = .881$. The two dimensions of the stimulus solution shown in the top panel of Figure 2 clearly correspond to magnitude (Dimension 1, approximate variance accounted for [VAF] = .633) and odd/even (Dimension 2, approximate VAF = .144). The lower panel of Figure 2 shows how individual groups of subjects weighted these dimensions. The length of the vector between the origin and the subject's location is proportional to the amount of variance accounted for by the sum of these dimensions. The angle of this vector shows the relative significance of the two dimensions. Thus, the kindergarten and third-grade subjects showed a small weighting for the odd/even dimension and a very large weighting on the magnitude dimensions corresponding to their nearly complete insensitivity to multiplicative relations such as odd versus even.

It is often useful to supplement the continuous spatial model provided by MDS with analyses that attempt to describe similarity as the products of discrete clusters. If children are incorporating into their understanding of

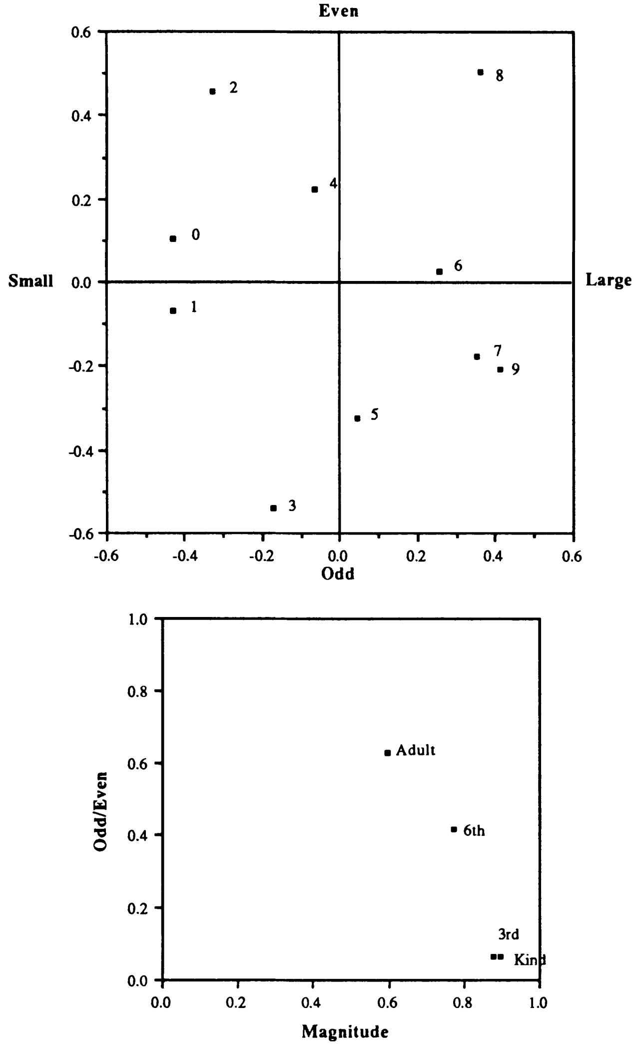


FIGURE 2 Weighted MDS analysis of children's number-similarity judgments (data from Miller & Gelman, 1983). Top panel shows stimulus space. Lower panel shows subject or weights space, which indicates the salience of each dimension for a particular age group.

numbers multiplicative features such as divisibility by two, then sets of numbers that share such features should be judged to be similar to each other. Figure 3 shows results of a second individual differences analysis of the Miller and Gelman data, describing the data as a set of overlapping clusters rather than as continuous dimensions. The individual differences clustering model used, Carroll and Arabie's (1983) INDCLUS, is an individual differences generalization of the overlapping clustering model (AD-

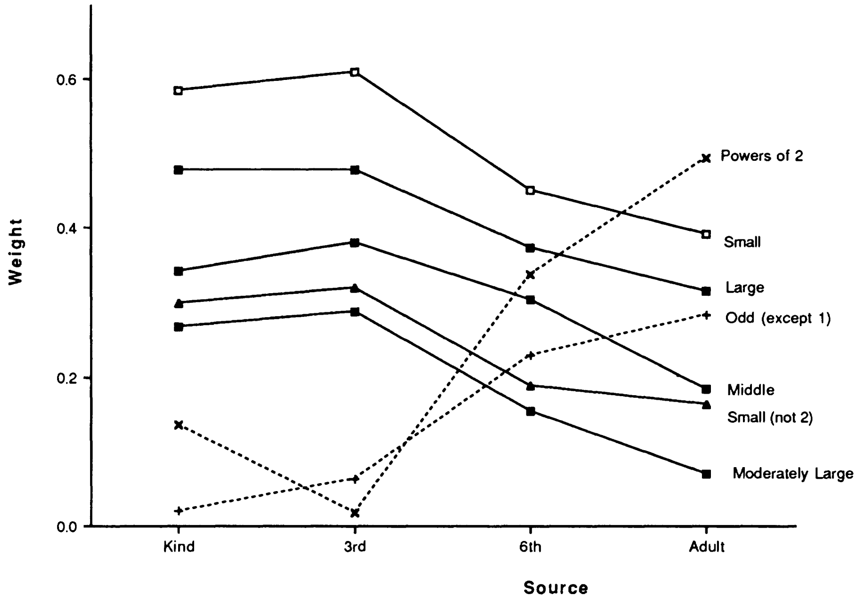


FIGURE 3 Cluster weights for individual differences clustering (INDCLUS) analysis of Miller and Gelman data on children's number-similarity judgments. The INDCLUS algorithm determines separate weights for individual subjects (or groups) of a common set of overlapping clusters. Cluster weights provide a measure of the salience of particular clusters for a given age group. With development, there is a declining emphasis on magnitude-based features (moderately large, small, medium-size, large, and small numbers) and an increase in salience for clusters related to multiplicative features (powers of 2, and odd numbers).

CLUS) developed by Shepard and Arabie (1979). INDCLUS is analogous to INDSCAL in that it assumes that individuals differ in their weighting of a common set of clusters. The INDCLUS model describes similarity as a function of sets of overlapping clusters, with individuals differing in the extent to which they weight particular clusters. The INDCLUS model can be fitted by a program of the same name, which is a generalization of the Arabie and Carroll (1980) MAPCLUS algorithm for fitting the ADCLUS model. As with MDS techniques, the INDCLUS program determines clusters algorithmically from the data rather than requiring them to be specified by the user in advance (although the user specifies the number of clusters to be found, analogous to specifying the number of dimensions to be used in MDS).

Results from the clustering analysis are quite similar to the SINDSCAL results in showing a gradual developmental deemphasis on clusters based on numerical magnitude, as various multiplicative features (e.g., odd numbers or powers of two) come into prominence.

The results shown in Figures 2 and 3 indicate that learning new numerical operations can result in changes in one's representation of numbers. Development involves a gradual shift away from an initial magnitude-based representation of numbers to an emphasis on multiplicative relations.

Miller and Gelman (1983) described their results in a manner consistent with Bryan and Harter's (1899) view of skill, namely, that mastering new numerical operations leads to the incorporation of those features into children's representation of number. To the extent that this process is a general byproduct of mastering new numerical relations, one would predict that abacus skill would result in the addition of abacus features to the number representations of experts. This prediction is far from self-evident, however, because the numerical features found by Miller and Gelman are both relevant within the context of calculation and relevant to mathematics in general. The structure of the abacus permits one to distinguish features that are primarily relevant to calculation from other numerical features. Although the abacus and numeral system share most relations, learning the abacus should increase the set of features available to be incorporated in one's conception of number. The extent to which these features do become a part of the number representations of abacus experts provides a means for evaluating the conceptual consequences of skill.

METHODS

To evaluate the effect of acquiring abacus skill on children's representation of number, we collected judgments of the similarity of numbers presented either as Hindu–Arabic numerals or as abacus figures to subjects from two cultures with three levels of abacus expertise. Subjects were Americans with no familiarity with the abacus and children in Taiwan with either brief exposure to abacus calculation (novices) or extensive training in mental abacus techniques (experts).

Settings

The study was conducted at Dongyuan Elementary School in Taipei, Taiwan, and at the Verlinden Elementary School in Lansing, Michigan. Both are urban public elementary schools, although the school in Taiwan was chosen because it has an unusually good abacus program (according to the Commissioner of Taipei schools). In both countries, children performed the task in groups during the regular school day.

Abacus training and certification. As is the case in schools throughout Taiwan, children in Dongyuan are first introduced to the abacus as part of the fourth-grade elementary school mathematics curriculum. Those children who wish to become experts may undertake more intensive training by enrolling in programs that meet after school. These after-school programs (called *buxiban*) are usually associated with an elementary school, and the participants engage in interschool competitions involving both abacus and mental calculation.

Complementing the abacus program at Dongyuan Elementary School is Dongyuan Buxiban, an after-school program directed by the abacus teacher from the elementary school. Children enrolled in the *buxiban* attend classes three afternoons per week for 1 to 1½ hr on each afternoon. Of the approximately 750 students at each of Grades 4 through 6, about 100 from each grade elect to attend these abacus classes after school. A few students in first, second, and third grades also attend classes at the *buxiban*. Members of the abacus team that will represent Dongyuan Elementary School in competition are chosen from among those children who attend classes at the *buxiban*. Anyone is welcome to join the *buxiban* program, and a large cross-section of children attend.

Children who participate in the after-school program also take certifying exams administered by the government-run Chinese Abacus Association. These exams provide the basis for rating abacus operators on their level of expertise and include sections on both abacus and mental computation. The highest rating attainable is that of *duan wei,* followed by Grades 1 through 6 in decreasing order of skill. Even a rating of Grade 6 requires a fair amount of proficiency, and passing a test is far from simple. During the second testing session of 1980, for example, 148,927 people took one of the seven qualifying exams throughout Taiwan. Of these, only 29,979 passed, thus attaining the rating they sought (Zhusuan Xuehui, 1980).

Subjects

Chinese subjects. A total of 64 sixth-grade students from Dongyuan Elementary School participated in the Taiwan portion of the study. Half were classified as abacus experts, and half as novices. The novices had all studied abacus calculation as part of the fourth- and fifth-grade mathematics curriculum, but none had ever attended the *buxiban* or attained a very high degree of proficiency in abacus calculation. Their competence with the abacus is roughly comparable with that which an American child in a good elementary school might have with Roman numerals: the ability to recognize numbers presented in this format and to do simple arithmetic with them but without any real facility. All the abacus expert subjects regularly attended the Dongyuan Buxiban after-school program. Experts had all received ratings of abacus skill from the Chinese Abacus Association. Twenty of our expert subjects had attained the highest rating (*duan wei*). Eight subjects were ranked at the next level, Grade 1, and the remaining 4 were ranked at Grade 2. Of the experts, 17 were boys and 15 were girls. Of the novices, 18 were boys and 14 were girls. Experts ranged in age from 11 years, 4 months, to 12 years, 3 months, with a mean age of 11 years, 10 months. Novices ranged in age from 11 years, 0 months, to 13 years, 2 months, with a mean age of 11 years, 8 months.

American subjects. A group of 22 American subjects (14 boys, 8 girls) was recruited from a middle-class public elementary school in Lansing, Michigan. The American subjects ranged in age from 11 years, 3 months, to 12 years, 9 months, with a mean age of 12 years, 1 month. None of the American subjects was familiar with the Japanese abacus used in the present research.

Materials

To include a meaningfully large set of abacus and numeral relations, stimuli for this experiment were extended from the 10 digits used in Miller and Gelman (1983) to include all numbers from 0 to 20. One consequence of using a larger set of stimuli is that it greatly increases the number of comparisons to be made. With 21 stimuli, there are 210 pairwise comparisons and 1,330 triads. Instead of the triadic judgment procedure used by Miller and Gelman, subjects in the present study judged similarity of pairs of numbers. To reduce the length of the task, the 210 pairs of stimuli from this set were randomly broken into two sets, and parallel booklets were prepared containing stimuli represented as either numbers or abacus forms. Each subject saw only one type of stimulus. Subjects rated the similarity of pairs of numbers or abacus representations along a 5-point scale ranging from *very similar* to *very different*. These data were then coded so that large values corresponded to judgments that two numbers were very different from each other.

Because subjects judged only half the pairs of stimuli, there should be some increase in error variance in the results over the situation where subjects judge the entire set. Judged similarities within each Task × Expertise group include some between-subjects as well as between-numbers variance. This should increase the overall error variance in the results. Because number pairs were randomly assigned to the two stimulus sets, there should not be any other systematic change in results, just greater variability in judgments than might otherwise be the case.

Procedure

Chinese data collection. Chinese subjects were tested in two separate groups, one composed of experts and the other of novices. Half the subjects in each group received booklets with arabic numerals, and half received booklets with abacus representations. A female examiner, native to Taiwan, instructed each group in the nature of the task and the use of the

5-point rating scale. Subjects were first asked to rate the similarity of pairs of geometric shapes, and responses were checked by the examiner. When the examiner was confident that all subjects in the group were correctly using the rating scale, the subjects were allowed to begin rating the number pairs in their booklets. The examiner again checked to make sure all subjects understand the task.

Subjects were instructed to judge how similar they felt each pair of numbers to be. They were told that there were no right or wrong answers and that we were interested in their opinions. They were not constrained as to the criteria they used in judging similarity of abacus or numerical stimuli. Thus, the data collected bear on those features subjects deemed important in judging similarity, rather than on whether abacus expertise affected one's ability to perceive or attend to specific features if instructed to do so.

American data collection. The procedure used with the American subjects differed only slightly from that employed in Taiwan. The American children had no systematic prior exposure to the abacus and were, thus, all complete novices with respect to the organization and use of the abacus. They were told that the figures in some of the books showed abacus representations and that the abacus was a way of representing numbers. They were not told anything about how numbers are represented on the abacus until after the study was over. The same instructions, materials, and preliminary tasks were used for the American subjects as for the Chinese children.

RESULTS

Data Aggregation

Data from 5 subjects (1 each from the novice abacus, novice numbers, expert abacus, U.S. numbers, and U.S. abacus groups) were discarded because of misalignment between the number of stimuli and number of answers. Analyses were, therefore, based on a varying number of subjects: 16 (expert number judgments), 15 (novice abacus, novice numbers, and expert abacus), or 10 (U.S. numbers and U.S. abacus) for the different groups. Data were aggregated by determining the average similarity rating given to each pair of numbers in a given Expertise × Stimulus group.

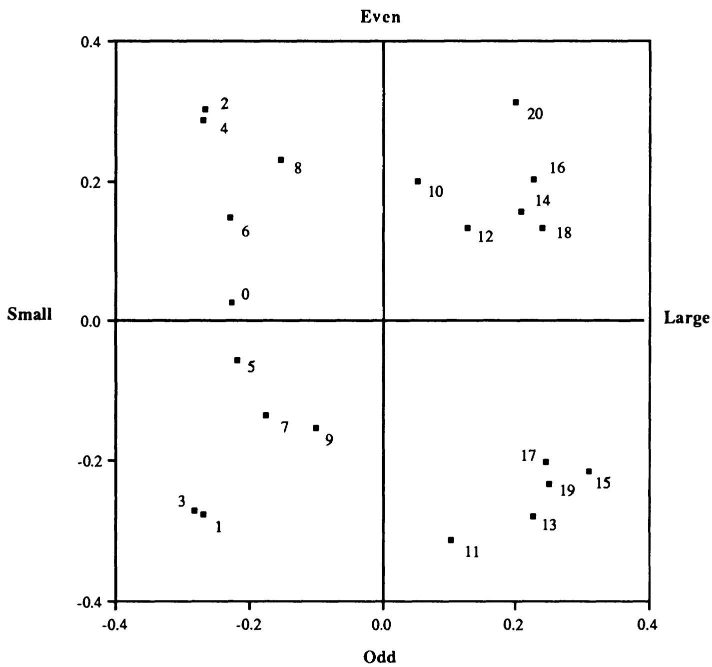
Overview of the Analysis

To obtain a general picture of the features subjects used in judging similarity between numbers, an individual differences MDS solution (using SINDSCAL) and an individual differences nonhierarchical clustering solution (using INDCLUS) were obtained. These provide a general sense of the extent to which different groups of subjects place different emphasis on abacus versus number features in judging numbers. Features that emerged from these analyses (numerical magnitude, odd/even parity, and number of beads used to represent a number) were then used as predictors in separate regression equations calculated for the original similarity judgments from each Task × Expertise group. Hypotheses concerning differences between groups were tested by comparing regression weights for these predictors across different Task × Expertise groups.

MDS Results

Results from a SINDSCAL analysis of the aggregated similarity judgments indicate that three dimensions were the fewest that accounted for a reasonable amount of variance and produced an interpretable set of dimensions. The overall correlation between predicted values and scalar products based on obtained judgments for three dimensions was $r = .638$. Additional dimensions accounted for additional variance, of course, but these dimensions appeared to correspond to local features, for example, groupings such as 2, 4, 8. As noted previously, SINDSCAL assumes that individual differences are the result of subjects' varying emphases on the dimensions of a common underlying stimulus space. Therefore, SINDSCAL solutions are not in general rotationally invariant, and one can expect that dimensions accounting for variation in subjects' judgments will also correspond to meaningful stimulus dimensions. Figures 4, 5, and 6 present the three panels of the SINDSCAL solution, with the stimulus spaces on top and the subjects' weightings below. These dimensions correspond to magnitude (Dimension 1, approximate VAF = .188), odd versus even (Dimension 2, approximate VAF = .113), and a beads dimension (Dimension 3, approximate VAF = .104) corresponding to the number of beads used to represent a number on the abacus. The top panels of Figures 4, 5, and 6 present the stimulus spaces produced by combining each pair of these dimensions.

Figure 4 shows the first two dimensions of the three-dimensional SINDSCAL solution. The first dimension (small vs. large) appears to correspond to the numerical magnitude of the numbers. This apparent interpretation of the first dimension is supported statistically, with a Pearson product–moment correlation between magnitude of the stimulus numbers and their values on this dimension of $r(19) = .94$. By contrast, correlations between magnitude and the other dimensions were substantially smaller: $r(19) = .01$ for the second dimension; $r(19) = -.27$ for the third dimension. The second dimension (odd vs. even) distinguishes odd from even numbers, with a Pearson product–moment between odd/even status and values on the second dimension of $r(19) = .93$. Correlations between odd/even parity and the other two dimensions were negligible: $r(19) = -.09$ with the first dimension; $r(19) = .08$ with the third dimension. The correlation analysis



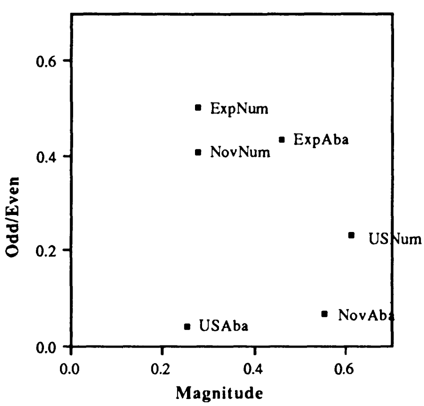


FIGURE 4 Weighted MDS analysis of abacus-numeral similarity judgments. This figure pairs the first two dimensions, which roughly correspond to magnitude and odd versus even. Odd numbers are shown in italic type. The top panel shows weighting of stimuli in a common space derived from all subjects' judgments. The lower panel presents the subject or weight spaces corresponding to the dimensions presented in the corresponding upper panel. Thus, for example, U.S. subjects judging the abacus placed very little weight on the odd/even dimension relative to magnitude.

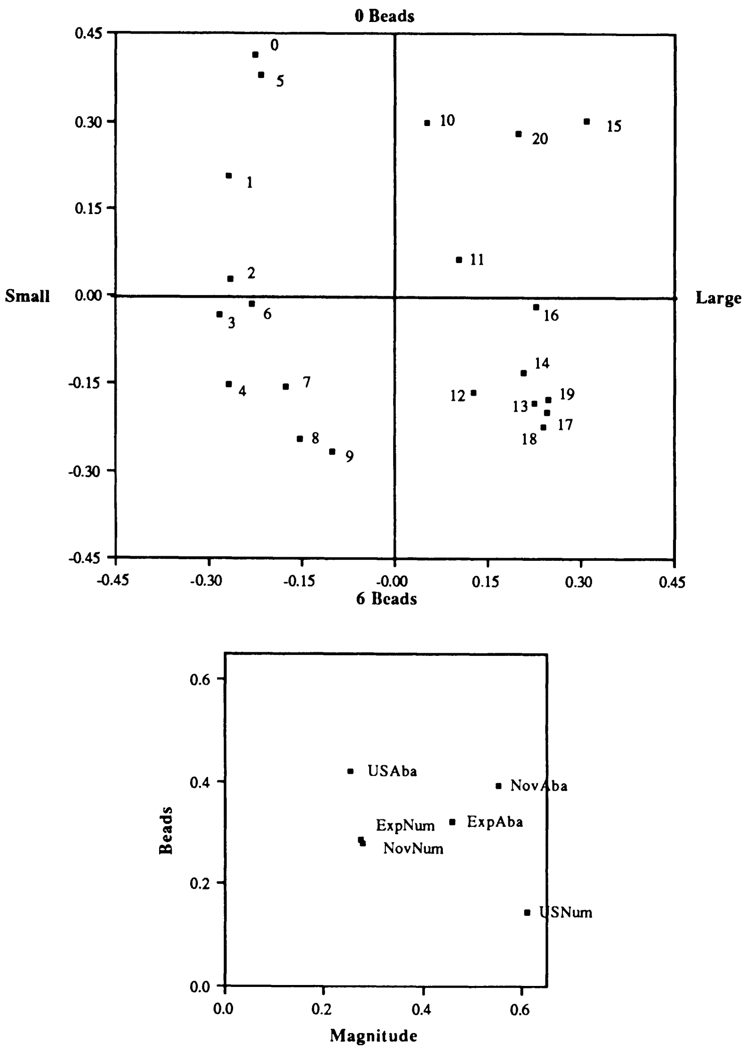


FIGURE 5 Weighted MDS analysis of abacus-numeral similarity judgments (continued). This figure pairs the first and third dimensions, which roughly correspond to magnitude and the number of beads used to represent a number on the abacus. This roughly corresponds to the value of the number modulo-5. Thus, 0, 5, 10, 15, and 20, equal 0 modulo-5, and 4, 9, 14, and 19, equal 4 modulo-5. The lower panel presents the subject or weight spaces corresponding to the dimensions presented in the corresponding upper panel. U.S. subjects viewing the abacus relied almost entirely on beads value, relative to magnitude.

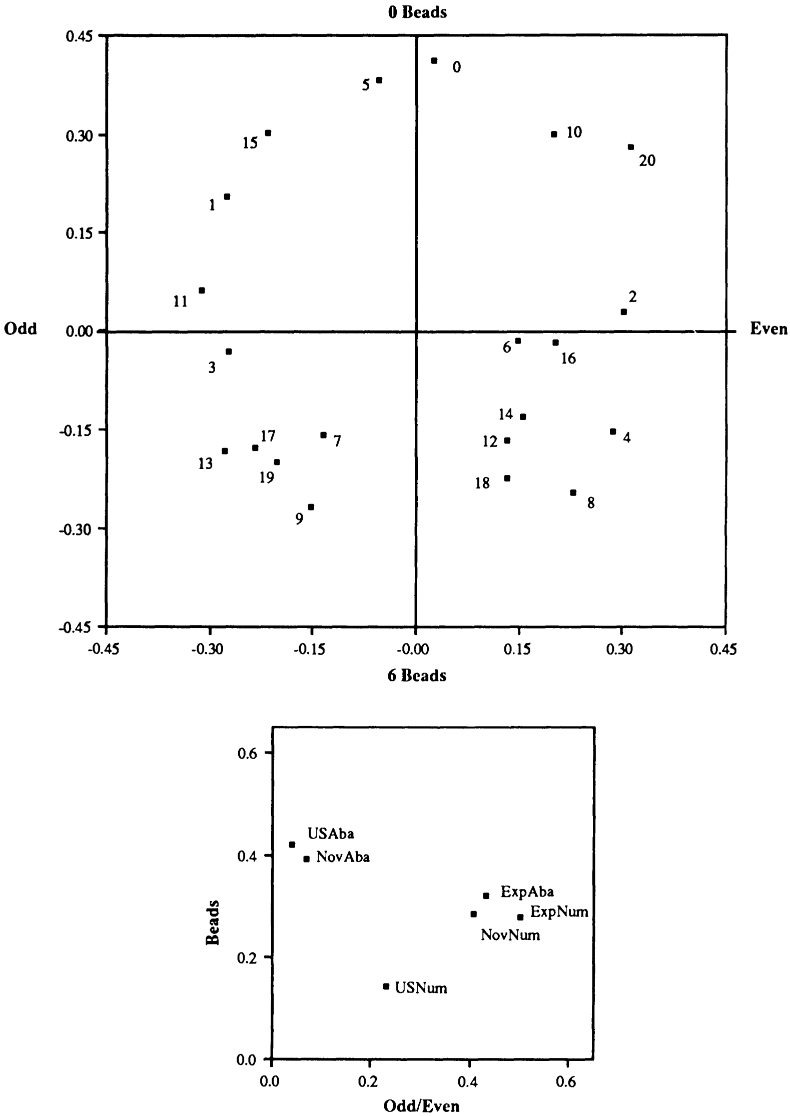


FIGURE 6 Weighted MDS analysis of abacus-numeral similarity judgments (continued). This figure pairs the second and third dimensions, which roughly correspond to odd versus even and the number of beads used to represent numbers on the abacus. The lower panel presents the subject or weight spaces corresponding to the dimensions presented in the corresponding upper panel. U.S. subjects viewing numerals relied almost entirely on odd/even parity relative to the beads dimension.

supports the interpretation of the first dimension as reflecting numerical magnitude and the second as reflecting odd/even status of stimulus numbers. These magnitude and odd/even dimensions correspond to those found by Miller and Gelman (1983).

Figure 5 shows the third dimension, beads value, plotted against the magnitude dimension. The beads-value dimension, marked 0 beads to 6 beads, deserves some elaboration. The beads value of a number is the number of beads used in representing it on the abacus. Looking at Figure 1, this is the number of beads that are pushed toward the center bar; for example, 8 is represented by four beads. Because of the modulo-5 structure of the abacus (i.e., that numbers differing from each other by exactly five use the same number of lower beads), beads value also largely corresponds to the modulo-5 value of a number.

Looking at the upper panel of Figure 5, one can count from 0 to 4 moving consistently downward before jumping back up nearly to the top for 5. This process is repeated for 6 through 9, with a jump up to 10, and to some extent for 11 through 15 and 16 through 20, although the modulo-5 interpretation of this dimension is less clear for these numbers. The correlation between this dimension and the number of beads used to represent a number is $r(19) = .87$, and the correlation with each number's value modulo 5 of stimuli is $r(19) = .86$. By contrast, correlations between the bead representation and the other dimensions were substantially smaller: $r(19) = .04$ for the first dimension; $r(19) = -.06$ for the second dimension. Thus, the third dimension is clearly most highly correlated with the bead representation of the stimulus number. The third dimension provides support for the view that the modulo-5 feature of abacus representations of numbers finds its way into subjects' similarity judgments. Figure 6 maps the beads-value dimension against the odd/even dimensions.

Consideration of subject spaces is necessary to determine the role that these various dimensions play in judgments by subjects in the different groups. The lower panels of Figures 4, 5, and 6 show the subject spaces for the dimensions represented in the corresponding stimulus spaces above them. For each group of subjects, the weighting for the pair of dimensions shown directly above is plotted on the three lower graphs. The distance along each axis is proportional to the total variance accounted for by those dimensions. For example, the location of the symbol "ExpNum" in the lower panel of Figure 4 shows that experts judging numeral stimuli placed more emphasis on odd/even than on magnitude relations.

The American subjects were forced to judge the abacus using abacus features, because they were not familiar with the numerical significance of the abacus representations. It is, therefore, instructive to consider how these children viewed the abacus (the "USAba" condition). Americans placed relatively little emphasis on either the odd/even dimensions or on magnitude in judging abacus similarity (shown by their low loadings on both dimensions in the lower panel of Figure 4). This result is reasonable, because the abacus represents neither magnitude nor odd/even status straightforwardly. American subjects placed heavy emphasis on the beads value of abacus stimuli, which provides an abacus-specific measure of how many beads would have to be moved to convert one figure into the other. The American subjects judged the abacus by its appearance, thereby providing a baseline for evaluating the performance of the other groups.

When Chinese novices judged the abacus, they placed little stress on odd/even features that were not represented on the abacus. Novices differed from Americans in emphasizing magnitude, presumably because, unlike the Americans, they knew which numbers were represented and were affected by this knowledge in judging similarity of abacus forms. Unlike the experts, however, novice judgments of abacus figures often differed from their judgments of numerals. An emphasis on the odd/even dimension was shown in novice numeral judgments, although this was given little stress in their abacus judgments. The beads dimension received much greater emphasis for the abacus than it did for numerals (shown in the lower panel of Figure 5), but the magnitude dimension was more salient for the abacus than for numbers. The novices "enrich" the physical features of the abacus with their understanding of the numbers represented; however, they have still not entirely integrated the two representations. Using the developmental progression from magnitude to multiplicative relations described by Miller and Gelman (1983), novice judgments of abacus similarity are less advanced than are their judgments of numerals.

Although novice and American subjects were greatly affected by whether they were shown abacus or numeral stimuli, this was not the case for experts. Looking across the three subject-weights panels, it is clear that the expert subjects were the least affected by the mode in which the stimuli were shown. In each weight panel of the SINDSCAL solution, the two expert conditions (ExpNum and ExpAba) are closer together than the corresponding loadings for the novices or the American subjects. Thus, experts treated the abacus and numerals as though they share the same important features. The next question is whether this common representation more closely resembled the abacus or the numeral judgments of the other groups.

In general, expert judgments were more like the numeral judgments of the novices and Americans than they were like the abacus judgments of these groups. As shown in the lower panel of Figure 4, experts judging the abacus resembled other sophisticated groups (experts and novices judging number and American adults) in emphasizing the odd/even dimension at the expense of magnitude. This finding is noteworthy, because the abacus does not distinguish between odd and even numbers in any fashion. In the weights space of Figure 5, the experts judging the abacus again resemble novices and experts judging numbers in placing relatively small emphasis on magnitude and the beads dimension. Experts viewing the abacus were midway between the two novice conditions in their emphasis on magnitude, suggesting that there are still some modality effects even for experts. Finally, the lower panel of Figure 6 shows both expert conditions together with the novices judging number in their relative loadings on the odd/even and beads dimensions. The persistence of the expert subjects in using such features as odd/even, even when shown the abacus stimuli, suggests that they were more willing than novices to bring nonabacus features to the task of judging abacus similarity.

A final result concerns the differences in number similarity judgments between the Chinese and American subjects. When judging numerals, American subjects placed greater emphasis on magnitude relative to the odd/even dimension than did the Chinese children. Using the developmental progression observed by Miller and Gelman from magnitude to reliance on multiplicative relations, the American subjects were less advanced than were their Chinese peers.

The SINDSCAL analysis of number and abacus similarity judgments produces a set of three meaningful dimensions, adding on abacus space to the two-dimensional representation found in a study of Americans judging numerals. Abacus expertise was associated with an emphasis on the numeral rather than on the abacus aspects of this space, as well as with greater consistency across the two modes of stimulus presentation.

Clustering Analyses

The abacus and numeral features can also be considered as sets of clusters that share common features. INDCLUS was also used in analyzing the effects of abacus expertise on number similarity judgments.[1] The abacus figures shown in Figure 7 report the members of the resulting clusters. In general, they can be clearly classified as numeral or physical abacus properties on the basis of inspection alone. The even numbers, for example, share

[1]Notes on the Clustering analysis. To use INDCLUS, one needs to determine the number (but not the content) of clusters to be fit. Carroll and Arabie (1982) suggested beginning with about half the number of stimuli, then gradually reducing the number of clusters. Solutions for 6–9 clusters were obtained in this analysis. A seven-cluster solution produced a good balance between interpretability and variance accounted for, with clusters representing a diverse set of abacus and numeral features. Four of the set of 42 cluster weights are (slightly) negative; negative clusters are intrinsically uninterpretable, but, given the small absolute value of their weights, they can be interpreted as representing random variation around an essentially zero weighting of that cluster by that group of subjects. If the solution is constrained to avoid negative weightings, the resulting solution resulted in minor changes to three of the seven clusters. The cluster labeled "lower bead 0 or 1" (containing 0,1,5,6,11,15,16,20 in the reported solution) added 3 and deleted 20. The cluster labeled "5-bead up" (1,2,3,4,11,12,13,14) added 9 and 10. The cluster labeled "5-bead down" (5,6,7,8,9,15,16,17,18,19) deleted 5 and added 12 and 13. These changes lead to a slight increase in overall variance accounted for (VAF = .372 vs. VAF = .344 for the model reported), albeit at the expense of some loss in interpretability of the clusters.

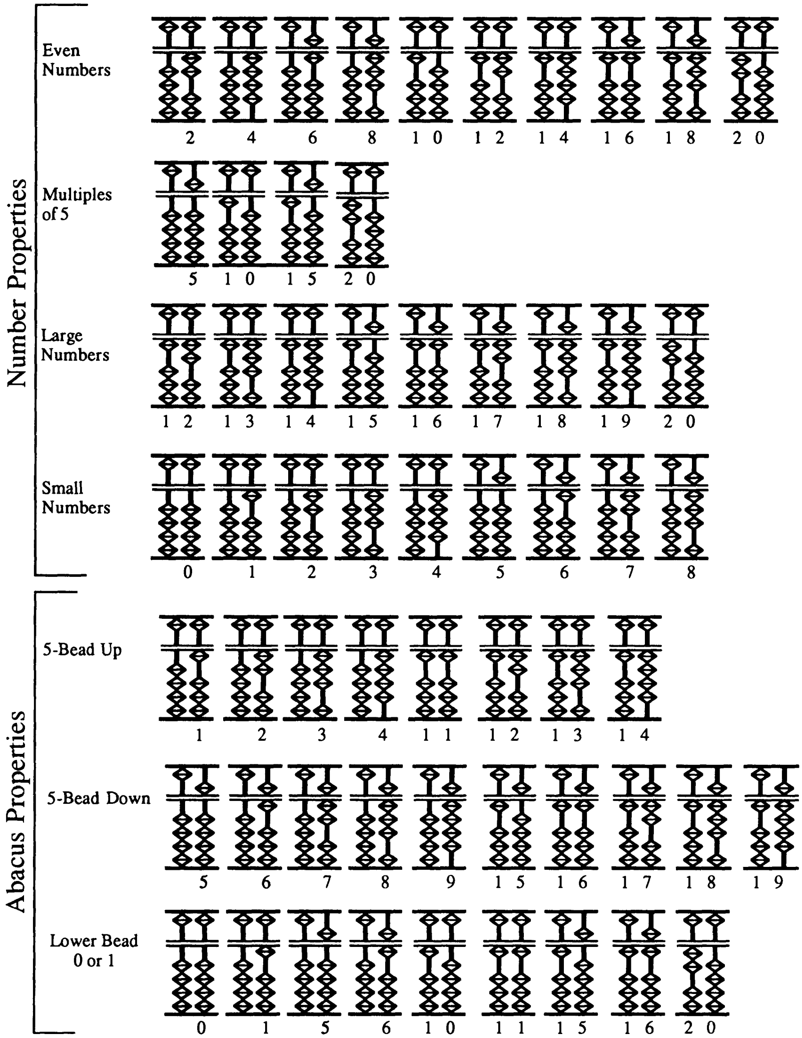


FIGURE 7 INDCLUS nonhierarchical clustering analysis of abacus-numeral similarity judgments. The Japanese abacus shown here is a modulo-5 system within a column. Beads "count" as they are moved toward the center (horizontal) bar. The top bead represents 5, whereas the lower beads represent 1.

no common feature in their representation on the abacus. The reader might wish to look at the first row of Figure 7 and attempt to find some common abacus feature that this set of numbers shares. The physical abacus properties presented in the lower rows of Figure 7, such as "5-bead up," "5-bead down," or having zero or one lower bead up, are clearly evident in the abacus figures.

There are several differences between the INDCLUS and SINDSCAL results that illustrate the differences between the continuous dimensional model produced by MDS techniques such as SINDSCAL and the discrete clusters found by INDCLUS. First, it may be of some interest that, although a cluster of even numbers was obtained, there was no corresponding cluster of odd numbers. The even numbers share an important feature: They are all evenly divisible by two. The odd numbers, however, are not all divisible by any common number. As the name *odd* number might suggest, their common feature is that they are not divisible by two. Second, it may seem remarkable that the set of numbers, 5, 10, 15, 20, is marked as being a number rather than as an abacus feature in the INDCLUS analysis, whereas the modulo-5 dimension was described as an abacus property in the SINDSCAL solution. These numbers also differ by modulo-5, but the set, 5, 10, 15, 20, also reflects the important multiplicative feature that all these numbers are evenly divisible by 5. Comparing results from scaling and clustering analyses provides a way of looking at specific features that may underlie dimensional judgments. The INDCLUS results suggest that the odd/even dimension from SINDSCAL results from the similarity that even numbers share and that it is useful to distinguish from the overall modulo-5 relation represented on the abacus as a particular set of numbers (multiples of 5) whose importance extends beyond the abacus.

The general modulo-5 relation represented on the abacus and reflected in the third dimension of the SINDSCAL solution is a characteristic of limited importance beyond its role in performing abacus calculation. Nonetheless, the particular modulo-5 set 5, 10, 15, 20 is useful in a variety of situations, such as measuring time and counting money, as well as being a common set size for ad hoc groupings of items. Thus, it should not be surprising that our most mathematically sophisticated subjects showed a general deemphasis on the overall modulo-5 relation at the same time they showed an increased emphasis on the set consisting of multiples of 5.

In general, the INDCLUS results correspond to those obtained with SINDSCAL in documenting an overall decline in the importance of physical abacus properties with the acquisition of expertise. This result is particularly clear in the top panel of Figure 8, which shows the consistent decline in weight for abacus features (lower bead 0 or 1, 5-bead up, and 5 bead) across expertise groups, but there is an overall increase in the weights of numeral-related clusters. Ironically, those subjects who spent the most time mastering the details of the abacus seemed least affected by its structure.

Regression Analyses

The scaling and clustering analyses produced a description of the stimulus variables that subjects varying in abacus expertise used in judging the similarity of numerals and abacus figures. Testing hypotheses concerning the effects of abacus expertise on number representation can be done most

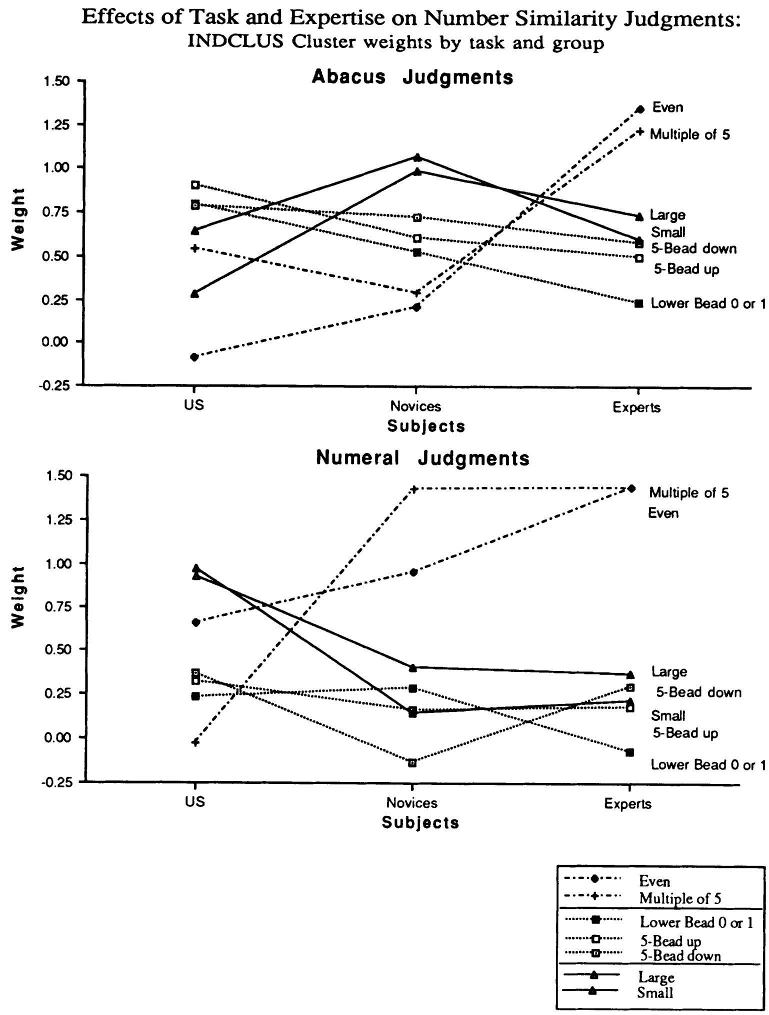


FIGURE 8 Results of an INDCLUS analysis of abacus and numeral similarity judgments. Results for abacus judgments are shown in the top panel, whereas numeral judgments are shown in the bottom panel. Features of the physical abacus are shown by dashed lines; other number features are connected with solid lines.

conveniently by using the variables that emerged from the scaling and clustering analyses as the basis for traditional regression analyses.

Relations between abacus and numeral judgments. One question that can be answered by looking directly at ratings is whether expertise is associated with increasing differentiation or integration of judgments across the two representational formats subjects judged. Table 1 presents correlations between ratings by subjects in the various Expertise × Task conditions, which were first normalized so that within-group means and standard deviations were identical. Evaluating the significance of these correlations is problematic, because judgments of the various pairs of stimuli are not independent of each other. Hubert (1978, 1979) described a heuristic procedure for evaluating the likelihood of obtaining a similar pattern of entries across two matrices. Hubert's approach involves comparing an obtained concordance statistic with a test distribution of the statistic generated from repeated random permutations of the rows and columns of one of the matrices being compared.

To evaluate the significance of the obtained correlations, we generated 100 random permutations (with replacement) of the rows and columns of each Task × Expertise similarity matrix (the same permutation was applied to both rows and columns). For each permutation, Pearson product-moment correlations were calculated for each predictor (the other Task × Expertise groups plus the numerical predictors listed in Table 1), and extreme values were estimated from this distribution for two-tailed tests ($ps = .05$ and $.01$). These were used to evaluate the obtained correlations. Each correlation in Table 1 was evaluated against the empirical distribution of the correlation between the matrix listed in the left column (American number for the first entry) and each of 100 random permutations of the matrix listed in the top row (American abacus for the first entry). The inter-

TABLE 1
Correlations Between Number-Similarity Judgments and Predictors

	American		*Novice*		*Expert*		*Predictors*	
	Abacus	*Number*	*Abacus*	*Number*	*Abacus*	*Number*	*Magnitude*	*Beads*
American								
Abacus								
Number	.138							
Novice								
Abacus	.432**	.517**						
Number	.060	.299**	.127					
Expert								
Abacus	.163**	.601**	.429**	.419**				
Number	−.034	.360**	.124	.521**	.641**			
Predictors								
Magnitude	.154*	.590**	.544**	.069	.448**	.187**		
Beads	.545**	.216**	.598**	.159*	.321**	.079	.270**	
Odd/even	−.114	.075	−.095	.305**	.325**	.378**	−.073	−.078

Note. Correlations are based on raw similarity judgments, normalized within each Task × Expertise group. Critical values for correlations involving similarity judgments are based on Hubert's (1979) matrix-conditional heuristic procedure as described in the text. Critical values for the predictor–predictor correlations are based on tabled values for the Pearson product-moment correlation coefficient.

*$p < .05$. **$p < .01$.

correlations among the three numerical predictors of similarity judgments (magnitude, beads, and odd/even parity correlated with each other) were evaluated against ordinary tabled values for the Pearson product-moment correlation, because these qualified as independent correlations.

The relation between judgment of numbers presented as numerals or abacus figures varied between the different expertise groups. The correlation between ratings of the same pair of numbers presented as numerals or as abacus figures is small for both the Americans ($r = .138$, ns) and the novices ($r = .127$, ns). For experts, on the other hand, there was a substantial correlation between judgments across the two tasks ($r = .641$, $p < .01$). The correlation analysis quantifies what the scaling results suggested: With increasing expertise, the mode in which numbers are presented is less important. The correlational results also indicate that expert abacus judgments are based on numerical features rather than the reverse. Expert abacus judgments correlate significantly with the numeral as well as with the abacus judgments of the other expertise groups, whereas expert number judgments are significantly correlated only with the numeral judgments of the other expertise groups.

The correlation results do not indicate which mathematical features are emphasized and which are neglected as experts develop a common view of abacus and numerical stimuli. To answer these questions, the three dimensions (magnitude, beads value, and odd/even parity) derived from the SINDSCAL analysis were used to predict similarity judgments for the various Task × Expertise groups. Specifically, magnitude (1 to 20) was coded as the numerical difference between a pair of stimuli. The beads predictor (1 to 6), coded as the number of beads that differed between the abacus representations of the two numbers, can be thought of as the number of beads that would have to be moved to change one abacus figure into the other. Finally, the odd/even parity (1 to 2) of a pair of numbers was coded as 1 if the numbers were both odd or both even and as 2 in all other cases. Figure 9 shows simple regression beta weights for the three predictors (beads, magnitude, and odd/even parity) for each Task × Expertise combination.

Differences between these predictors were evaluated using a procedure described by Hubert and Golledge (1981). To evaluate the direction and significance of differences between expertise groups in the weight given to these numerical features, the permutational method described before was applied to matrices of the differences between the normalized judgments of different groups of subjects judging the same stimuli. Thus, the significant and positive correlation found between magnitude and the difference between American and novice abacus judgments means that Americans placed greater weight on this feature than did novices; a significant negative correlation would imply that novices placed greater weight on this attribute than did Americans when judging abacus figures.

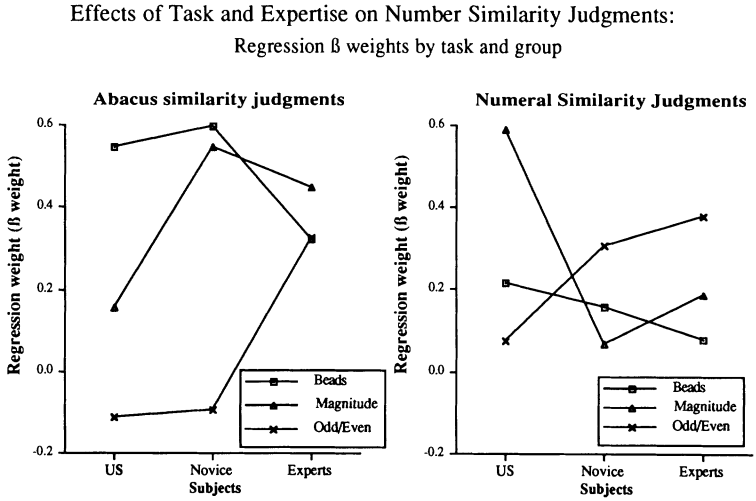


FIGURE 9 Regression of number similarity judgments by task and expertise groups. Simple regressions for three predictors (beads difference, odd/even parity, and magnitude difference) were determined for each Task × Expertise combination.

Abacus judgments. Inspection of Figure 9 suggests that, for abacus judgments, expertise is associated with a decrease in emphasis on abacus-specific features. The beads predictor is represented most clearly on the abacus, and both U.S. and novice groups placed heavy emphasis on this feature in their judgments of abacus figures. There was no significant difference between U.S. and novice subjects on beads weighting ($r = -.05$, ns), but each placed significantly more emphasis on beads than did experts ($r = .173$, $p < .05$ for U.S. experts; $r = .259$, $p < .01$ for novices vs. experts). Odd/even parity is not represented on the abacus, and the relation between expertise and this dimension was the mirror image of that obtained for the beads predictor. U.S. and novice subjects were similar in not weighting odd/even parity heavily ($r = -.017$, ns), but experts placed significantly greater emphasis on this feature than did either of the other two groups ($r = -.339$, $p < .01$ for U.S. experts; $r = -.393$, $p < .01$ for novices vs. experts). A different pattern was found for magnitude. Americans differed from novices and experts in placing less emphasis on magnitude ($r = -.366$, $p < .01$ for U.S. novices; $r = -.227$, $p < .05$ for U.S. experts). Novices and experts did not differ from each other in their emphasis on magnitude ($r = .090$, ns). As noted before, one of the main ways that novices differed from U.S. subjects was that novices knew which number an abacus figure corresponded to. This knowledge appears to be enough to make magnitude a significant determinant of judgments, although the knowledge of novices was not sufficient to make odd/even parity relevant to their judgments of abacus figures.

Numeral judgments. A very different set of relations between expertise and similarity judgments emerged for the numerical stimuli. There were no significant differences among expertise groups in subjects' emphasis on beads in judging numerical stimuli ($r = .049$, ns, for the U.S. novice comparison; $r = .081$, ns, for novices vs. experts; and $r = .121$, ns, for U.S. experts). The three groups differed greatly in their familiarity with the abacus, but there was no evidence that any group was more likely than the others to apply the features of the beads representation to their judgments of stimuli presented as numerals.

For the other two predictors, magnitude and odd/even, there were significant differences between the American subjects and the two groups of Chinese subjects, but there were no differences between the two Chinese groups. American subjects placed greater emphasis on magnitude than did either of the other groups ($r = .440$, $p < .01$ for U.S. novices; $r = .356$, $p < .01$ for U.S. experts), which did not differ from each other ($r = -.120$, ns). Americans also placed less weight on odd/even parity than either of the other groups ($r = -.194$, $p < .01$ for U.S. novices; $r = -.268$, $p < .01$ for U.S. experts), which did not differ ($r = -.075$, ns).

The regression analyses quantify the pattern suggested by the scaling and clustering results. With expertise in using the abacus, there was a decrease in the emphasis of abacus-specific features in judging similarity, as they are replaced by more general number-related characteristics. In judging similarity of abacus figures, American subjects were largely limited to abacus features, as represented by the beads predictor. Abacus novices, familiar with which numbers are represented by specific abacus figures, added an emphasis on numerical magnitude to the beads representation in judging abacus figures. Experts, despite their facility with using the abacus, showed the least emphasis on abacus-specific features in judging abacus stimuli. They emphasized other number properties instead, showing a greater emphasis than other groups on odd/even parity, which is not represented on the abacus.

In judging numeral stimuli, no group showed much of an emphasis on specific abacus features. U.S. subjects differed from the two Chinese groups in their heavy emphasis on magnitude and their comparatively lower emphasis on odd/even parity. In light of Miller and Gelman's (1983) finding of a developmental shift in relative emphasis between magnitude and multiplicative features in judging number similarity, these results suggest that the U.S. subjects were less sophisticated in their number representations than were their Chinese peers.

Regression analysis of number similarity judgments supports the hypothesis of conceptual transparency. Experts at using a particular representational system for calculation showed a conceptual deemphasis, relative to less expert groups, on those features unique to the representational system they had mastered.

Although expert judgments of number and abacus similarity showed a relative deemphasis on abacus-specific features, it is important to note that expert judgments of these two representations of numbers were not identical. Rather experts showed a deemphasis, relative to other groups, on features specific to the representational modality shown.

DISCUSSION

Analyses of numeral and abacus-similarity judgments across the different levels of abacus expertise support three main findings: (a) Acquisition of skill at abacus representation was associated with greater consistency of judgments across modes of representation; (b) expert judgments of abacus similarity deemphasized abacus-specific features, relative to judgments by subjects of less expertise; and (c) abacus expertise was associated with a general increase in the sophistication of the numerical features represented in judging abacus stimuli, corresponding to the previously described shifts related to mathematical expertise.

Conceptual Integration as a Consequence of Skill

The finding that expertise was associated with increasing consistency of judgments across modes indicates that experts relate these two physical representations when thinking about number. The issue of how two number representations are related is generally analogous to the distinction between coordinate and compound representation of language in bilinguals (Ervin & Osgood, 1954; Weinreich, 1953/1974). In Weinreich's view, languages can be represented either as largely separate systems (coordinate bilingualism) or as a single-fused system (compound bilingualism). Although the linguistic evidence is quite ambiguous (Hakuta, 1986; Paradis, 1977), the present results indicate that, in the case of the abacus, skill is associated with development of conceptual relations between these two representations. Note that these results do not answer whether abacus expertise involves learning relations between numeral and abacus systems or both to access a single underlying representation of number. What they do suggest is that the features specific to a particular mode of representing number become of less conceptual significance as that mode is mastered.

Conceptual Transparency and the Abacus

The finding that abacus-specific features become less important with increasing expertise when judging abacus stimuli supports the hypothesis of conceptual transparency. This finding is somewhat counterintuitive. Experts have spent the most time working at learning the abacus and, presumably, are the most familiar with it. Furthermore, analyses of their calculation performance show that the physical features of the abacus play a major role in determining the sorts of mistakes they make in calculation. Although the actual characteristics of the abacus are important in performing abacus calculations, our results indicate that with expertise they become less important when thinking about number. It appears that abacus experts relate their representations of the abacus and numerals by downplaying those features of the abacus that do not have significance beyond the performance of calculation. The fact that numbers that are exactly five apart have similar representations on the abacus is quite important for operating the abacus and for understanding the mistakes of abacus calculators. In the context of a general shift from magnitude-related to multiplicative features in number representation, however, the fact that 8 and 3 are exactly five apart becomes of decreasing significance with increasing numerical sophistication.

Expertise at using the abacus as an alternative representation for number is associated with a general increase in numerical sophistication. This conclusion has to remain very tentative, because significant differences between novices and experts were limited to numbers presented as abacus figures. American subjects were more likely than either Chinese group to judge number similarity based on magnitude rather than on multiplicative relations. Although general mathematics achievement differences (Stevenson, Lee, & Stigler, 1986) would lead one to conclude that Americans might be less sophisticated about number than their Chinese peers, there are also other reasons to predict that expertise with an alternative representation for number would lead to increasing sophistication of number concepts. In a study comparing abacus experts and novices in Taiwan, Stigler, Chalip, and Miller (1986) found that abacus expertise was less associated with increased calculation skill per se than with increased performance on the number concept and with applications portions of standardized achievement tests. The bilingualism literature may also be relevant here, although the analogy is somewhat tenuous. Saxe, Becker, Sadeghpour, & Sicilian (1989) found that French–English bilinguals showed earlier understanding than their monolingual peers of the formal features of counting. Having an alternative representation for number available, children may become more sensitive to those features of number that are representation dependent versus those that are general across particular representations.

In discussing the effects of abacus expertise on number representation, we have been assuming that numerical sophistication is a consequence of abacus expertise rather than its cause. This issue is not often addressed in the expertise literature, and the present study does not permit one to determine a direction of causality concerning the relation between numerical sophistication and abacus skill. A previous study (Stigler et al., 1986) suggested it is unlikely that initial numerical facility accounts for much of the variation in acquiring abacus skill. Looking at who chose to take abacus training, Stigler et al. found that a model incorporating first-grade mathematics grades and sex accounted for a small amount of variance ($R^2 = .035$, $p < .001$) but did not permit one to classify correctly any subject in the abacus-trained or abacus-untrained group. Eventual skill at the abacus was predicted primarily by the number of hours of organized practice, with a smaller contribution from general intelligence and no significant prediction based on mathematics grades. Although abacus expertise may still be predicted by selection variables (e.g., whatever motivates one to spend long hours learning a skill), there is no strong relation between initial mathematical aptitude and eventual abacus skill. Abacus experts do seem more sophisticated in their judgments of numbers, whether presented as abacus figures or numerals, and this finding is more likely a consequence than a cause of their expertise with abacus calculation.

Taken together, the results of these analyses suggest that it is useful to maintain a distinction between the features that are meaningful in using a system such as the abacus and those that are meaningful in reflecting on it. Successful performance of a complex skill requires mastery of the relevant details. Yet integrating the skill with the rest of one's knowledge may, as in the case of the abacus, require deemphasizing those same details. The fact that 3 and 8 appear similar on the abacus is relevant to understanding the performance of abacus experts. This pair of numbers has very little else in common, and constructing an integrated representation of number involves deemphasizing such abacus-specific features.

Binet's model of expertise as a matter of transcending domain-specific details thus provides a better description of the conceptual consequences of abacus skill than does Bryan and Harter's belief that a mastered skill masters its practitioner. This result is consistent with Anderson's (1982; also see Neves & Anderson, 1981) description of levels of processing in the development of skill. In this view, mastery of a skill involves a transition from central processes accessible to consciousness to a final proceduralized state in which the procedures become less accessible. Although the present study did not address issues of accessibility, it is clearly the case that abacus features of the abacus become less salient conceptually with abacus expertise. As an object for thought, the mental abacus presents a very different appearance from the features shown in studies of expert abacus calculation.

Implications for Other Skills

The phenomenon of conceptual transparency, which posits that features important in the implementation of a skill are deemphasized conceptually, appears to be applicable beyond the limits of abacus skill. Gentner's work (Forbus & Gentner, 1986; Gentner, 1988; Gentner, Landers, & Ratterman, 1986) on analogy is a particularly good example of this phenomenon.

Gentner et al. (1986) presented subjects with sets of stories that varied in surface matches (e.g., whether or not they involved eagles) and in underlying relational structure. Although subjects rated stories having the same underlying relational structure as being the best analogies, surface matches turned out to be more effective at reminding subjects of previously read stories. The surface features of an analogy appear to follow the same pattern of conceptual transparency shown in the abacus representations of experts. Although such features are important in performing the task of retrieving previous stories, they are deemphasized as bases for analogies.

There is no reason to expect, however, that what is true of one skill will be true for every domain of expertise. In physics (Chi et al., 1981) and in mathematical problem solving (Schoenfeld & Herrmann, 1982), acquiring expertise is a matter of learning very general principles that are used to organize both problem solving and thinking about the domain. Many domains, such as learning a musical instrument, involve a heavy emphasis on domain-specific features (e.g., the relation between notes on a trombone) reminiscent of abacus skill. To the extent that results from this study can be generalized, they suggest that such domain-specific features become less important with expertise in thinking about the domain over which skill operates.

Studies of expertise have focused on the skills and processes that define expertise, with the result that much is known about how experts function. Even experts exist outside the constraints of their expertise, however, and becoming expert at a procedural skill can pose conceptual problems. The pianist hears music from other instruments, and the abacus calculator sees other representations of numbers. Understanding the meaning of expertise in semantically rich domains requires one to move beyond the constraints of the skill. As in the case of abacus skill, obtaining a conceptual understanding of the skill may result in a diminished significance for those features that organize performance of the skill.

ACKNOWLEDGMENTS

The research reported in this article was supported by National Science Foundation Grants BNS-85-10546 and BNS-89-09566 to Kevin F. Miller and Spencer Fellowships from the National Academy of Education held by both authors. We thank Richard Houang and Shin-ying Lee for their assistance in conducting the research. We also acknowledge helpful comments by Franca Agnoli, Maureen Callanan, Earl Hunt, Jamie Campbell, Howard Nusbaum, Herbert Pick, and Robert Sternberg. We are also grateful to the students and staff of Verlinden School in Lansing, Michigan, and of Dongyuan Elementary School in Taipei, Taiwan, for their assistance with the research reported here.

REFERENCES

Adelson, B. (1981). Problem solving and the development of abstract categories in programming languages. *Memory and Cognition, 9,* 422–433.

Allard, F., & Burnett, N. (1985). Skill in sport. *Canadian Journal of Psychology, 39,* 294–312.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369–406.

Arabie, P., & Carroll, J. D. (1980). MAPCLUS: A mathematical programming approach to fitting the ADCLUS model. *Psychometrika, 45,* 211–235.

Binet, A. (1966). Mnemonic virtuosity: A study of chess players. *Genetic Psychology Monographs, 74,* 127–162. (Original work published 1893)

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language. The acquisition of a hierarchy of habits. *Psychological Review, 6,* 345–375.

Carroll, J. D., & Arabie, P. (1982). *How to use INDCLUS, a computer program for fitting the individual differences generalization of the ADCLUS model.* Murray Hills, NJ: AT&T Bell Laboratories.

Carroll, J. D., & Arabie, P. (1983). INDCLUS: An individual differences generalization of the ADCLUS model and the MAPCLUS algorithm. *Psychometrika, 48,* 157–169.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart–Young decomposition. *Psychometrika, 35,* 283–319.

Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology, 33,* 1–50.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Chi, M. T. H., & Koeske, R. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology, 19,* 29–39.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18,* 257–273.

de Groot, A. (1965). *Thought and choice in chess.* The Hague: Mouton.

Egan, D., & Schwartz, B. (1979). Chunking in recall of symbolic drawings. *Memory and Cognition, 7,* 149–158.

Eisenstadt, M., & Kareev, Y. (1975). Aspects of human problem solving: The use of internal representations. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition.* San Francisco: Freeman.

Engle, R. W., & Bukstel, L. (1978). Memory processes among bridge players of differing expertise. *American Journal of Psychology, 91,* 673–690.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87,* 215–251.

Ervin, S., & Osgood, C. (1954). Second language learning and bilingualism. *Journal of Abnormal and Social Psychology, 49*(Suppl.), 139–146.

Forbus, K. D., & Gentner, D. (1986). Learning physical domains: Toward a theoretical framework. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 2, pp. 311–348). Los Altos, CA: Morgan Kaufmann.

Gentner, D. (1988). Analogical interference and analogical access. In A. Preiditis (Ed.), *Analogica: Proceedings of the First Workshop on Analogical Reasoning* (pp. 63–88). London: Pitman.

Gentner, D., Landers, R., & Ratterman, M. J. (1986, November). *Analogical access: A good match is hard to find, Part 2.* Paper presented at the annual meeting of the Psychonomic Society, New Orleans.

Gobbo, C., & Chi, M. T. H. (1986). How knowledge is structured and used by expert and novice children. *Cognitive Development, 1,* 221–237.

Hakuta, K. (1986). *Mirror of language*. New York: Basic Books.

Hatano, G. (1982). Cognitive consequences of practice in culture specific procedural skills. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 4,* 15–18.

Hatano, G., Miyake, Y., & Binks, M. G. (1977). Performance of expert abacus operators. *Cognition, 5,* 47–55.

Hatano, G., & Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. *Cognition, 5,* 95–110.

Hubert, L. J. (1978). Generalized proximity function comparisons. *British Journal of Mathematical and Statistical Psychology, 31,* 179–192.

Hubert, L. J. (1979). Generalized concordance. *Psychometrika, 44,* 135–142.

Hubert, L. J., & Golledge, R. G. (1981). A heuristic method for the comparison of related structures. *Journal of Mathematical Psychology, 23,* 214–226.

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika, 29,* 1–27.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579–594.

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika, 29,* 1–27.

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika, 29,* 115–129.

Marr, D. (1982). *Vision*. New York: Freeman.

McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology, 13,* 307–325.

Miller, K. F. (1987). Geometric methods in developmental research. In J. Bisanz, C. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology* (pp. 216–262). New York: Springer-Verlag.

Miller, K. F., & Gelman, R. (1983). The child's representation of number: A multidimensional scaling analysis. *Child Development, 54,* 1470–1479.

Neves, D. N., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 57–84). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Paradis, M. (1977). Bilingualism and aphasia. In H. A. Whitaker & H. Whitaker (Eds.), *Studies in neurolinguistics* (Vol. 3). New York: Academic.

Pollard-Gott, L. (1983). Emergence of thematic concepts in repeated listening to music. *Cognitive Psychology, 15,* 66–94.

Pólya, G. (1957). *How to solve it* (2nd ed.). Garden City, NY: Doubleday.

Pruzansky, S. (1975). *How to use SINDSCAL: A computer program for individual differences in multidimensional scaling*. Murray Hills, NJ: AT&T Bell Laboratories.

Reitman, J. (1976). Skilled perception in go: Deducing memory structures from inter-response times. *Cognitive Psychology, 8,* 336–356.

Saxe, G. B., Becker, J., Sadeghpour, M., & Sicilian, S. (1989). Developmental differences in children's understanding of number word conventions. *Journal for Research in Mathematics Education, 20,* 468–488.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 484–494.

Shepard, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function: I. *Psychometrika, 27,* 125–140.

Shepard, R. N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function: II. *Psychometrika, 27,* 219–246.

Shepard, R. N. (1963). Analysis of proximities as a technique for the study of information processing in man. *Human Factors, 5,* 33–48.

Shepard, R. N., & Arabie, P. (1979). Additive clustering: Representation of similarities as combinations of discrete overlapping properties. *Psychological Review, 86,* 87–123.

Shepard, R. N., Kilpatric, D., & Cunningham, J. (1975). The internal representation of numbers. *Cognitive Psychology, 7,* 82–138.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61,* 394–403.

Stevenson, H. W., Lee, S. Y., & Stigler, J. W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science, 233,* 693–699.

Stigler, J. W. (1984). "Mental abacus": The effect of abacus training on Chinese children's mental calculation. *Cognitive Psychology, 16,* 145–176.

Stigler, J. W., Chalip, L., & Miller, K. F. (1986). Consequences of skill: Abacus training in Taiwan. *American Journal of Education, 94,* 447–479.

Weinreich, U. (1974). *Languages in contact: Findings and problems.* The Hague: Mouton. (Original work published 1953)

Zhusuan Xuehui. (1980). *Chinese abacus association examination results record.* Taipai, Taiwan: Author.