Understanding and Evaluating Measures of Computer Ability: Making a Case for an Alternative Metric

Robin H. Kay
University of Toronto

Abstract

Most researchers and educators evaluating computer ability or literacy have used a construct approach. A prototypical measure assesses programming ability, application software skill, and computer awareness. This kind of measure is often viewed as a final product rather than as a pedagogical tool. In this article, it is argued that the fundamental basis of computer ability measures should be altered in order to address a rapidly evolving computer software market and take advantage of significant advances made in instruction and cognitive science. A more fundamentally useful metric of computer ability can be developed using a process-oriented methodology. This new metric, comprised of actual learning activities, helps provide a comprehensive and coherent understanding of how a person interacts with a computer. Metaphors and constraints based on an extensive core of intelligence assessment research are used to illustrate how theorists have looked at computer ability and to provide guidelines for developing more useful ability measures. A microgenetic approach is offered as one promising example of a process-oriented method that could produce a richer metric of computer ability for researchers and educators. (Keywords: computer literacy, computer utilization, measurement.)

A need exists to critically evaluate current philosophy and methods used to investigate computer ability and to develop an alternative metric that could better serve the needs of educators and researchers. Previous researchers have developed computer ability measures that produce final test scores intended to reflect a student’s skill level (Kay, 1989, 1992a). These test scores offer little toward helping researchers and educators understand how individuals learn: the mechanisms of conceptual change, the role of context in learning, and factors that bring about “good” learning (Hambleton, 1982). A shift from categorical to process-focused methods is needed to develop a comprehensive and coherent understanding of how people learn with computers, and this information could be used to produce an alternative metric that will help researchers and educators explore and improve computer knowledge acquisition.

This article will first provide an overview of the kinds of definitions and methodology that currently dominate computer ability research. Then it will explain the principal reasons why current methodology is insufficient to keep up with rapidly changing technology and instruction research. Next, the kinds of metaphors used in general ability measures and the tension between universal versus diverse doctrines of intelligence will be examined in order to pro-
vide useful constraints for developing a new kind of computer ability measure. Finally, one promising example of the kind of methodology that could advance current thought on computer ability will be presented.

CURRENT DEFINITIONS OF COMPUTER ABILITY

Defining the nature of computer ability, or, as it is more commonly referred to, computer literacy, has proven to be an elusive endeavor. Since its relatively modest beginning in 1972 (Moont, 1987), the term computer literacy has taken on many definitions, giving it a rather broad meaning (Kay, 1989, 1992b). The extraordinarily rapid growth of technology since 1972 has significantly influenced the design of software, further complicating what it means to “know about computers” (Forrester, 1989; Lepper, 1985).

To bring the evaluation of computer ability into perspective, it is informative to look at the broader picture of research on general intelligence. The assessment of intelligence has been largely governed by two principle figures—Galton, who supported a unitary, linear approach, and Binet, who advocated a dynamic, adaptive model (Keating, 1990a; Sternberg, 1990). In recent research, Galton’s views have come to represent a categorical, construct-based doctrine, while Binet’s represent a process-oriented, intervention model.

Developers of computer ability measures have followed Galton’s path, choosing to identify and organize computer ability into statistically justified factors (see Kay, 1989, 1992b, for a review). The use of these standardized, reliable tests, though, has far more to do with ease of application than with theoretical rigour (Keating, 1990a).

A comprehensive computer ability measure might include questions about an individual’s ability to do a range of tasks, range of experience on a number of software packages and programming languages, and level of formal training. However, most measures include only one or two of these components (Kay, 1986). Conspicuously absent from these measures are data on how the subject interacts with a computer, the context of computer use, and the goals and needs of the individual.

PROBLEMS WITH THE CONSTRUCT APPROACH

While a construct-driven approach to assessing computer ability may have been informative and useful in the early days of computer technology, two significant developments have made it necessary to consider alternative metrics. The first development involves the evolution of computer technology. It is reasonable to assume that the rapidly expanding and constantly evolving software market is swiftly decreasing the value of standardized, all-inclusive measures of computer ability. Consider Lepper’s (1985) comments on the evolution of hardware development:

Had improvements in efficiency and reductions in the cost of automobiles followed patterns similar to the computer industry.
each of us would be able to buy a Rolls-Royce today for roughly $2.75; it would get nearly 3,000,000 miles to the gallon and would deliver enough power to tow an aircraft carrier. (p. 1)

While computer software has not evolved as quickly as hardware, it has changed considerably in the past 10 years. A construct approach to assessing software skills has the distinct disadvantage of trying to keep pace with a constantly changing content. As a software package develops and changes format, the construct used to assess ability to use that software has to change accordingly. On the other hand, a computer ability metric that addresses process more than content is less vulnerable to transformations in software. Even if the content or type of software changes, the processes involved in learning that software should remain relatively constant. Because a process-focused metric is more flexible in terms of content, it should ultimately be less costly and more practical.

The second development making it problematic to apply a construct approach to computer ability measures is the progress being made by cognition and instruction researchers. Promising trends in cognitive science research that have the potential to influence the educational process include analysis of expertise (see Ericsson & Smith, 1991, for a review), contextual importance for learning and motivation (Resnick, 1989), developmental levels/qualitative difference in thinking (e.g., Carey, 1985; Cheng, Holyoak, Nisbett, & Oliver, 1986; Kuhn, 1989), self-regulatory skills (e.g., Bereiter & Scardamalia, 1989; Brown & Palincsar, 1989), knowledge organization and structure (e.g., Chi & Bassok, 1989; Chi, Hutchinson, & Robin, 1989), individual differences (Glaser, 1990), intuition and naive theory (e.g., Carey, 1985), the teacher’s role in learning (e.g., Brown & Palincsar, 1989), and attitudes toward the learning process (e.g., Dweck, 1986). These trends are founded on a methodology that focuses almost exclusively on process, that is, on how the learner acquires new knowledge. Therefore, a construct-based measure, which is based on content and performance, is an inappropriate assessment tool in the context of these new trends. Unless attitudes toward evaluation and assessment co-evolve along with the instructional emphasis on process, research in cognitive science will have only a marginal impact. A process-centred metric based on actual behaviours and responses is needed if researchers are to accurately examine the mechanisms of human-computer interaction in the context of recent developments in cognitive science research. From a practical perspective, a process-centered metric is also important for helping educators address specific biases, problems, confusions, and frustrations that students experience (Lesgold, Ivill-Friel, & Boanr, 1989).

GENERAL INTELLIGENCE RESEARCH

Extensive research on general intelligence models (see Sternberg, 1990, for a review), provides a solid background with which to develop an alternative metric of computer ability. This research will be examined from two perspec-
tives. First, a review of internal and external metaphors used to measure ability will be examined to identify useful constraints on future computer ability measures. Second, the inherent tension in intelligence research between researchers attempting to identify universal principles and those preferring to explore diversity or individual differences will be discussed. Since computer ability measures have taken a predominantly universal approach to assessment, the potential contribution of a more diverse perspective will be emphasised.

**METAPHORS OF THE MIND**

Sternberg (1990) provided a concise description of internal and external metaphors of intelligence. Internal metaphors include geographic, computational, epistemological, and biological perspectives. External metaphors include sociological and anthropological models. While these metaphors have not been explicitly applied to investigation of computer-related skills, a brief overview of each helps identify useful constraints on future methods and models. The main characteristics and principal questions asked for each metaphor are presented in Table 1.

<table>
<thead>
<tr>
<th>Metaphor</th>
<th>Characteristics</th>
<th>Principal Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographic</td>
<td>Internal</td>
<td>What factors make up intelligence? Are there different kinds of intelligence?</td>
</tr>
<tr>
<td></td>
<td>Identifying constructs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paper-and-pencil tests</td>
<td></td>
</tr>
<tr>
<td>Computational</td>
<td>Internal</td>
<td>In what ways do humans think like computers?</td>
</tr>
<tr>
<td></td>
<td>Formal logic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Task analysis</td>
<td></td>
</tr>
<tr>
<td>Epistemological</td>
<td>Internal</td>
<td>Do people acquire knowledge at a fixed rate? How can one explain both consistency and diversity?</td>
</tr>
<tr>
<td></td>
<td>Developmental</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fixed stages</td>
<td></td>
</tr>
<tr>
<td>Biological</td>
<td>Internal</td>
<td>What is the physiological basis for intelligence?</td>
</tr>
<tr>
<td></td>
<td>Brain functions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Physiological models</td>
<td></td>
</tr>
<tr>
<td>Anthropological</td>
<td>External</td>
<td>How does culture and educational environment influence learning?</td>
</tr>
<tr>
<td></td>
<td>Role of culture</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Developmental</td>
<td></td>
</tr>
<tr>
<td>Sociological</td>
<td>External</td>
<td>How does context influence learning?</td>
</tr>
<tr>
<td></td>
<td>Role of context</td>
<td></td>
</tr>
</tbody>
</table>
Internal Metaphors

Geographic Metaphor. The most prolific model of intelligence is based on a geographic metaphor, which attempts to map or identify different intelligence factors on which people differ (Sternberg, 1990). The prototypical computer ability test is almost exclusively grounded on the geographic perspective. While tests based on the geographic metaphor form the metric by which other metaphors are measured, these tests offer precious little information about the learning pathways to computer proficiency.

Computational Metaphor. The computational or information-processing metaphor views the human mind as being analogous to the operation of a computer (Sternberg, 1990). A strong emphasis is placed on identifying mechanisms as researchers attempt to trace the flow of information through the mind (Keating, 1990b). However, a number of cognitive researchers have questioned the validity of this approach. Formal logic and algorithms do not provide a good description of how we actually reason (Dreyfus & Dreyfus, 1989; Kuhn, 1990; Voss, 1989). What is easiest to simulate with computers (carrying out extended chains of reasoning) is most difficult for humans, and what is easiest for humans (pattern recognition and associative memory) has eluded computer programmers (Bereiter, 1991; Bobrow, 1985; Johnson-Laird, 1983; Nisbett & Ross, 1980; Tversky & Kahneman, 1974). Apart from ecological validity, this model has been criticized for its lack of attention to actual information structures and its inadequate attention to the complexity and sophistication of human skills (Keating, 1990b).

Researchers of computer ability have, somewhat ironically, sidestepped the computational metaphor. In a partially restricted and readily modifiable environment, the computational metaphor could provide some clues about how people interact with computers. For example, a simple comparison between menu (icon) versus command (text) might reveal information about the role of knowledge structure in the learning process. Computationalists might examine keystroke intervals in order to evaluate reaction time, memory access, and attention level. The computer environment is well suited to the computational perspective.

Epistemological Metaphor. This perspective is largely based on Piaget’s work on children (Keating, 1990b; Sternberg, 1990). Piaget maintained that coherent logical structures underlie children’s development. Through a process of assimilation (incorporating environment into one’s internal structure) and accommodation (incorporating one’s internal structure to fit the environment), the child proceeds through a fixed sequence of qualitatively different stages.

Again, researchers have not used this metaphor to explore human-computer interaction. Questions that could be raised include the following: Do people acquire computer-related concepts at a fixed and ordered rate? What is the role of previously learned software in learning new concepts? What is the nature of the external-to-internal and internal-to-external fit in the learning process? These are interesting and potentially fruitful lines of investigation.

Unfortunately, the epistemological model may be fundamentally flawed. The problem of how one coordinates a fixed rate of development within an open,
changing system must be resolved (Keating, 1990b). In essence, this problem represents the same tension experienced when one is trying to explain both the consistencies and variations observed in human behaviour. Since Piaget did not examine the strategies and processes that children use to solve various tasks (Sternberg, 1990), it remains a mystery how “open” and “closed” systems can coexist.

Informative research on human-computer interaction must address the ways in which consistency and diversity co-develop in a learning environment. Another way of putting this is that theorists have to look at consistencies as well as individual differences in learning, and at the role that these differences play in conceptual change.

Biological Metaphor. The biological model seeks to understand intelligence in terms of the functioning of the brain (Sternberg, 1990). Much of the data are derived from brain-damaged patients, EEG analysis, and monitoring of blood flow. Parallel Distributive Processing (PDP) theorists have borrowed from this model to develop a provocative model on the construction and modification of knowledge structures. The obvious reductionist bent to the biological method has both advantages and disadvantages. Reductionism is a pleasing notion in the sense that it is generally agreed that human thought is ultimately derived from biological building blocks (Churchland, 1988; Searle, 1984). Yet this approach provides only one kind of explanation, one that would not be practically appealing to researchers of computer abilities.

Nonetheless, a biological approach could be applied to research on human-computer interaction. As stated previously, the computer environment offers a unique combination of structure and flexibility. One could imagine manipulating the structure of software and assessing the effect that these simple manipulations had on physiological measures. One could also assess physiological correlates with performance while using the computer, particularly in times of stress and conceptual change. The biological model, used in a computer learning environment, might help address the nebulous relationship between emotions and cognition.

However, one should again be wary, because the biological metaphor, like other metaphors, reveals only a piece of the puzzle and should be used in conjunction with other models.

External Metaphors

Anthropological Metaphor. This model emphasizes the role of culture on intelligence (Sternberg, 1990). Cultural factors prescribe the sequence and age at which certain abilities develop. While the details of how culture affects intelligence have not been worked out, this metaphor adds an important constraint to any serious investigation of ability. An analysis of the education milieu in which computer-related tasks and concepts are learned is critical to understanding the broader picture of conceptual change. One could evaluate differences in motivation based on the social web of goals and needs in various social situations; including work, school, casual use, and enforced use.
Sociological Metaphor: The sociological perspective, perhaps best illustrated in the work of Vygotsky (1978), involves a theory of internal construction of external operations, in which knowledge is internalized from the outside world (Sternberg, 1990). Researchers using this approach are interested in the socialization of knowledge within a particular context. The emphasis on context is noticeably absent in computer ability research. Yet current cognitive theory stresses three interrelated aspects of learning: knowledge construction, knowledge dependence, and situation dependence (Resnick, 1989). The geographic perspective taken by most computer ability researchers glosses over all three of these aspects. It is important, then, for research designed to uncover the mechanisms of conceptual change in computer-related tasks to be sensitive to the role of context.

Summary

Sternberg (1990) pointed out that “we often see theories as competing when in fact they are not: They are different answers to different questions, not different answers to the same question” (p. 284). For example, the biological metaphor does not provide educators with the answers required to enhance conceptual change in their students. Correspondingly, the sociological perspective is not particularly sensitive to the issue of neurological deficits. How people come to view intelligence, then, is determined by their methods. Any one metaphor cannot serve as a complete explanation for something as complex as human intelligence (Sternberg, 1990).

It is clear that most computer ability researchers have relied on the geographic metaphor. It is equally clear that other metaphors can contribute significantly to improving computer ability measures and to increasing our understanding of how people acquire computer knowledge. The microgenetic approach (Keating, 1990b), discussed later, incorporates a combination of Sternberg's metaphors, including geographic, epistemological, and sociological. It does not include either biological or anthropological models.

UNIVERSAL VERSUS DIVERSE DOCTRINES OF INTELLIGENCE

Keating (1990b) noted an important polarity in the way scientists explore and explain human behaviour. At one end of the continuum, scientists look at behaviour in terms of universal laws (the Athenian perspective). At the other end, behaviour is seen as much more contextual—how something works in a particular situation (the Mancunian perspective). Traditionally, universality has been seen as inherently more meaningful than diversity, but, as Keating noted, “[evidence] of cognitive diversity is pervasive; understanding its sources and trajectories ... is fundamental to understanding cognition” (p. 39). Because universal principles have already been emphasized in computer ability literature (Kay, 1989; 1992b), the opportunities that the Mancunian perspective offers will be presented here.
The Role of Context

The Mancunian philosophy emphasizes the role of context in determining behaviour. Implied in this perspective is the notion of domain dependence, a notion that is inextricably linked with current cognitive thought (Resnick, 1989). The debate between domain-independent (general) and domain-dependent (factor) models has been largely conceived on factor analytic grounds (Keating, 1990b). Unfortunately, this statistical technique is misused and misapplied. It is often used in an a priori fashion to support tentative and shaky theories. Clearly, both consistency and diversity is observed in human behaviour, so it would be unwise to take an either-or position in this debate. We need to adopt a methodological openness to investigating both consistencies and variations (Keating, 1990b).

As stated previously, researchers are beginning to re-address the complexities of intelligence (Churchland, 1988; Keating, 1990a; Neisser, 1987; Resnick, 1989; Sternberg, 1990). The rational approach to thinking that dates back to Aristotle is too neat and structured to account for the kinds of diversity observed in everyday life. An accurate model of human-computer interaction cannot be developed without understanding the social forces that shape behaviour. Factors like emotion and motivation relating to human needs and goals, while not easily quantified, are particularly important in the field of computers, where considerable debate reigns regarding the use of these machines (Kay, 1986).

Individual Differences

The Mancunian philosophy opens the door to examining individual differences in intelligence. Instead of identifying global constructs, one can look at the styles and habits of individuals within a specific context. Considerable research has been done on the strategies people use to solve various problems (Globerson & Zelniker, 1989; Schmeck, 1988; Weinstein, Goetz, & Alexander, 1988). Some research emphasizes general styles that are active in a number of domains (Mayer, 1988; Schmeck, 1988). This approach is similar to Gardner’s (1983) exposition on multiple intelligences. Other authors (in Zelniker, 1989) have dissected the individual information processing involved in doing a task.

Few computer ability researchers have looked at individual differences within specific contexts, but this kind of analysis is important for understanding diversity in conceptual change. Again, the critical task is to develop a model that accounts for consistency and flexibility in learning styles.

Summary

The statistical and theoretical construction of universal, elaborate structures does not secure the presence of specific structuring principles (Bereiter, 1991; Keating, 1990b). As Keating suggested, ignoring diversity is theoretically risky—key elements could be missed. Accenting the role of context and individual differences in learning style and strategies would help advance research
on computer abilities. New methods should look at the characteristics of the person and the situation during the development of models that reflect the process of learning computer-related concepts. For example, the microgenetic approach is designed to analyze both general patterns and individual differences by focusing on learning activities in a particular context over time.

AN ALTERNATIVE APPROACH

Many kinds of constraints can be considered when doing process-oriented research on computer ability. Developing an all-inclusive study that addresses each of the potential constraints is unrealistic, but research must lean toward reconstruction of the pathways of learning computer-related tasks. As Keating (1990a) commented, “If cognitive activity is generally integrated, then observing how that integration occurs over the course of development offers the chief route to understanding it” (p. 16). Keating (1990b) suggested one promising technique that could be used to investigate conceptual change during the learning of software tasks. He used the term microgenetic to describe this method of research. Applied to the area of computer abilities, it would require researchers to observe and record a series of short-term observations (approximately one hour long) while a subject is learning new concepts on the computer. Simply stated, its emphasis would be on activity, not on ability.

A Process-Oriented Approach

In the first part of a study conducted by the present author (Kay, 1992c), 10 adults (4 female and 6 male), consisting of graduate students and professionals, filled out a brief questionnaire on their intentions to use computers, affective and cognitive attitudes, sense of control over the computer, and learning style. This was followed by a detailed interview that identified the kinds of software the subjects had used and their understanding of this software. Finally, an open-ended interview was conducted, which investigated the subjects' perceptions of how they approached different learning tasks. The two interviews were recorded on tape.

The purpose of administering the questionnaire and conducting the interviews was to provide some broad background information on the subjects' experiences with computers and learning in general. It might seem excessive, and perhaps too "geographic," to have recorded a detailed history of the subjects' experience using computers, but it proved important for the analysis and interpretation of the process data collected in the second part of the study.

For example, in the first part of the study subjects could simply be asked for how many years they had used computers and which software packages they knew. But answers to these questions would have offered no insight into how often they had used computers, what they used computers for, or how extensive their use of the software was. This kind of information was helpful in trying to evaluate the nature of the subjects' guesses and errors when they learned a new software package.
The purpose of the second part of the study was to observe the subjects’ developmental process of learning a new software package. In particular, episodes of conceptual change were examined in detail. In this part of the study, all subjects were videotaped for 60 minutes (the camera was focused on the computer screen) while they were learning the Lotus 1-2-3 (Version 2.2) spreadsheet package on an PC 286 clone. Each subject was asked to do the following tasks: move the cursor, enter rows and columns of numbers, enter labels, insert blank rows and columns, move and/or copy rows or columns of data, and compose formulas to add numbers.

Subjects were given instructions similar to those given in a typical protocol analysis (Oatley, 1988). The standard procedure was to introduce the subject to a task and ask the subject to think aloud while engaged in self-directed learning. It was critical to give subjects at least some tasks that they had not performed before, because individuals who are too familiar with a task have difficulty articulating what they are doing and why they are doing it. The task has become too automatic (Oatley, 1988). Unlike typical protocol analysis (Ericsson & Simon, 1984), subjects were given calculated “hints” when they were unable to proceed. This question-asking protocol was used with some success by Kato (1986). Every effort was made to encourage subjects to solve problems on their own. Attempts were made to limit intervention to situations where subjects did not know the syntax of a particular command.

**Possible Biases**

The biases of the microgenetic approach in this study should be noted. Integration of the geographic, cultural, and sociological methods is stressed at the expense of the neurological and computational metaphors. Furthermore, the method is designed to reveal individual differences, if any exist. Finally, the context of the study is limited to a tutorial kind of interaction and may not apply to other learning situations, such as small-group or classroom settings.

It should also be noted that collecting data, even in a developmental, constructivist manner, is only half of the experimental procedure. Analyzing the data, particularly protocol analysis, requires clear choices to constrain the numerous options available to the researcher.

**ANALYSIS AND AN ALTERNATIVE METRIC**

The first 50 minutes of each of the 10 subjects’ videotaped session were transcribed. Verbal expressions and sounds, as well as critical keystrokes, were included in the transcriptions. Each subject’s transcription was used to identify and categorize learning activities. These activities formed the core structure for the alternative computer ability metric.

Figure 1 provides an outline of activities that were observed in and coded from the transcripts. Four main categories of learning behaviours are listed: previous knowledge, task interpretation, actual learning or problem-solving activities, and errors made. An external factor, pace, is influenced by task interpreta-

[...]

[...]

[...]

[...]

[...]
Figure 1. Learning behaviour metric for assessing computer ability.

tion and is a main influence on specific learning activities. Furthermore, help activities, while not delineated yet, are affected by errors made and influence future learning activities.

While the subclassifications of learning behaviours are neither complete nor discussed in detail, they are presented to give an idea of the variety of activities that occur during learning (see Kay, 1992a, for a detailed presentation).

Each set of learning behaviours in Figure 1 can be compared to a progressive problem-solving score. This score is a measure of whether a subject has made progress toward solving a problem after displaying a specific behaviour.
By correlating this score with specific learning activities, effective and ineffective strategies can be identified. Subjects can then receive feedback to continue or alter certain behaviours. Also, educators and researchers can determine which patterns of learning need to be encouraged or corrected.

The learning behaviours given in Figure 1 can be applied to any situation where a particular piece of software is being learned, thereby establishing the stability of this measure. In other words, the metric used adapts to a rapidly changing software market. Furthermore, the metric's emphasis on learning activities allows researchers and educators to evaluate learning on a more progressive level, an idea that has been suggested by considerable research in cognitive science. Finally, feedback from the learning behaviour metric is conducive to being used as a pedagogical tool as opposed to a final test of a student's learning level. Emphasising actual behaviours makes it easier to act on particular problems than if general construct or content area is being evaluated.

CONCLUSION

If researchers are to advance current theory in the area of computer abilities, they will have to alter their methods and explore the construction of knowledge during the learning of new concepts. Current geographic methods, while concise and easily understood, provide almost no information about the dynamics of human-computer interaction. A microgenetic approach is a promising example of the kind of developmental, process-oriented method that will help researchers understand how people acquire and use computer-related concepts.

Contributor

Robin H. Kay obtained his B.Sc. and M.A. at the University of Toronto, where he specialized in psychology and computer applications in education. He is currently working on his Ph.D. in educational psychology at the Ontario Institute for Studies in Education. His research interests include the way in which people learn to use computers, gender differences in computer use, and the development of scales to assess computer behaviours. (Address: 118 Gloucester Ave., Oakville, Ontario, L6J 3W4, Canada.)

References


