THE NATURE of EXPERTISE

EDITED BY
MICHELENE T. H. CHI
ROBERT GLASER
MARSHALL J. FARR

The Nature
of Expertise

The Nature of Expertise

MICHELENE T.H. CHI
University of Pittsburgh

ROBERT GLASER
University of Pittsburgh

M.J. FARR
Office of Naval Research

Psychology Press
Taylor & Francis Group
New York London

First Published by
Lawrence Erlbaum Associates, Inc., Publishers
365 Broadway
Hillsdale, New Jersey 07642

Transferred to Digital Printing 2009 by Psychology Press
270 Madison Avenue, New York NY 10016
27 Church Road, Hove, East Sussex BN3 2FA

Library of Congress Cataloging-in-Publication Data

The nature of expertise.

Bibliography: p.
Includes index.
1. Performance—Psychological aspects—Congresses.
2. Apperception—Congresses. 3. Human information
processing—Congresses. I. Chi, Michelene T. H.
II. Glaser, Robert, 1921ffl III. Farr, Marshall J.
BF323.E2N37 1988 153 87ffl33071
ISBN 0-89859-711-0 CB
ISBN 0-8058-0404-8 PB

In loving remembrance of
William G. Chase
1940–1983

Contents

PART I PRACTICAL SKILLS

PART II PROGRAMMING SKILLS

PART III ILL-DEFINED PROBLEMS

PART IV MEDICAL DIAGNOSIS

In Memoriam

This volume is a tribute to the late William G. Chase. He died suddenly and unexpectedly three weeks after the conference from which the majority of the chapters in this volume are drawn. Although his career was short (15 years), the impact of his work is profound. In that short span he made several major contributions to the field of psychology. His dissertation, *Parameters of Visual and Memory Search*, would have been his first major piece of work. It had earned recognition from the American Institute of Research Creative Talents Award Program, but Bill's high standard of excellence prevented him from publishing it. He felt that the dissertation work was not deserving of publication because it did not answer to his satisfaction the question he asked.

Bill's subsequent contributions are well known. They consist of his work with Herbert Clark on the mental operations involved in the comparison of sentences and pictures (Chase & Clark, 1972); the work with Herbert Simon on the skills underlying chess playing (Chase & Simon, 1973); and the work with Anders Ericsson on skilled memory (Chase & Ericsson, 1981, 1982). In fact, his research on the nature of expertise is a primary source which shapes our current understanding. Numerous citations across the chapters in this volume describe his work, and attest to the significance and lasting contribution of his research. At the time of his death, he was on the verge of completing the most comprehensive theory of skilled memory to date, based on the extensive evidence he collected on the digit-span task. He was also pursuing the application of his theory to other domains, such as mental calculation (see Staszewski, this volume) and the memory of waiters (see Ericsson & Polson, this volume). Bill also began working on the fascinating new topic of cognitive maps (see Chase & Chi, 1981). Preliminary findings on the spatial representation of taxi drivers have been published (see Chase, 1983), but unfortunately the voluminous amount of data in the form of video tapes remain unanalyzed.

Bill had a unique research style: He used methodologically creative and rigorous ways to uncover mental processes and then developed stringent theories to explain and test his findings. His knack for reducing a complex task to a simple and elegant empirical investigation inspired all of us who were his students and colleagues.

Bill also became a superb teacher. His first teaching experience in the fall of 1968, at Carnegie-Mellon University, a course on Human Learning and Complex Behavior, presented him as a shy, soft-spoken, and incomprehensible lecturer. He would often mutter to the blackboard rather than to the students, and would present what seemed to the students to be a perpetual sequence of experimental evidence that was too complicated to understand. A decade later, his teaching became so enlightening that students would applaud him in a standing ovation at the last lectures of his classes.

Bill was a man of integrity. He spoke of what he believed to be right, and ignored any consequences that might have been damaging to himself. He was unpretentious and down-to-earth. Yet he had a fierce but good-natured competitiveness that seemed contradictory to his personality. This was manifested both in the way he conducted his research and the way he liked to excel in other accomplishments. For example, he took up badminton in his late thirties. He pursued it with such tenacity that within five years he won the championship for Pittsburgh Men's Singles. Everyone—students, colleagues, and friends—loved Bill. They loved him for his modesty, his gentleness, his sense of humor, and his accessibility. He was always available to listen to people's research problems as well as their personal problems, and offered genuine help, understanding, and suggestions.

Those of us who had the privilege of knowing Bill know how fortunate we are in having met such a special person. He will be missed by his family, friends, and colleagues for many years to come.

Micki Chi

REFERENCES

Chase, W. G. (1983). Spatial representations of taxi drivers. In D. R. Rogers & J. A. Sloboda (Eds.), *The acquisition of symbolic skills* (pp. 391–411). New York: Plenum Press.

Chase, W. G., & Chi, M. T. H. (1981). Cognitive skill: Implications for spatial skill in large-scale environments. In J. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 111–136). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chase, W. G., & Clark, H. H. (1972). Mental operations in the comparison of sentences and pictures. In L. W. Gregg (Ed.), *Cognition in learning and memory* (pp. 205–252). New York: Wiley.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Vol. 16* (pp. 2–58). New York: Academic Press.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141–189). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Overview

Robert Glaser

Michelene T. H. Chi
University of Pittsburgh

Expertise became an intriguing subject for investigation as a result of work in the mid- to late sixties, largely due to developments in artificial intelligence (AI) and cognitive psychology. Research in AI and attempts to simulate human capabilities had failed to construct programs that could outperform humans, even though computers were by then equipped with powerful search heuristics and essentially limitless search capabilities. Even in programs using selective search, such as Greenblatt's (Greenblatt, Eastlake, & Crocker, 1967) chess program, the best "plausible" move was still selected on the basis of an extensive evaluation, whereas human experts do not engage in particularly extensive searches or elaborate analyses, as shown by findings in cognitive psychology. Investigations of chess playing, for example, the early work of deGroot (1966) and the later extended work of Chase and Simon (1973), demonstrated that what distinguishes strong from weak players are their abilities to correctly reproduce large patterns of chess positions after a few seconds of viewing, rather than their searching more deeply or broadly than weaker players. Clearly, specialized structures of knowledge were strongly implicated, but the nature of this knowledge and of its interactions with general heuristic processes required further analysis.

Newell and Simon (1972) described the chess master's "perceptual" ability as follows:

> Clusters of related pieces in a position are recognized as familiar constellations; hence, each cluster is stored as a single symbol. Less skilled players have to describe the board as a larger number of simpler chunks—hence cannot

> hold all of the information required to reproduce the board in short-term memory. When the same number of pieces is arranged on the board at random, few of the resulting configurations are familiar even to grandmasters. They then need more symbols to describe the position that can be held simultaneously in short-term memory, hence, they perform as poorly as weaker players. (p. 781)

By using the concept of chunks to explain the chess master's pattern recognition superiority, it became necessary to identify experimentally the structure and size of chunks in the knowledge base, because a chunk appeared to be a defining unit of knowledge structure. Hence, early in its history, the study of expertise provided evidence of a knowledge-competence dimension as a primary focus.

In AI research, it became widely acknowledged that the creation of intelligent programs did not simply require the identification of domain-independent heuristics to guide search through a problem space; rather, that the search processes must engage a highly organized structure of specific knowledge for problem solving in complex knowledge domains. This shift in AI was characterized by Minsky and Papert (1974) as a change from a power-based strategy for achieving intelligence to a knowledge-based one. They wrote:

> The *Power* [italics added] strategy seeks a generalized increase in computational power. It may look toward new kinds of computers ("parallel" or "fuzzy" or "associative" or whatever) or it may look toward extensions of deductive generality, or information retrieval, or search algorithms. . . . In each case, the improvement sought is intended to be "uniform" – independent of the particular data base. The *Knowledge* strategy sees progress as coming from better ways to express, recognize, and use diverse and particular forms of knowledge. (p. 59)

This point of view has since been reiterated in the textbooks and handbooks on building expert systems (e.g., Hayes-Roth, Waterman, & Lenat, 1983). These texts point out that the principal developments in AI fostered the current emphasis on knowledge-based expert systems and the related field of knowledge engineering. Machines that lack knowledge can perform only intellectually trivial tasks. Those that embody knowledge and apply it can approximate the performance of human experts. As a consequence, expert-system building has concentrated on the knowledge that underlies human expertise and given less emphasis to the significance of domain-independent problem-solving heuristics.

Thus, the seeds of the study of the characteristics of highly competent expert performance were sown in the fertile ground of Newell and Simon's 1972 book, *Human Problem Solving*, although the topic *expertise* was not listed. In the ensuing years, the need for research in expertise has been recognized, and much research in cognitive psychology has been devoted to this topic.[1] In the following pages, we briefly summarize some key characteristics of experts' performances that this research has uncovered. These findings are robust and generalizable across the various domains that have been studied (Glaser, 1988). We also highlight other relevant findings, and speculate briefly on the nature of the organization of the knowledge base that generates each characteristic.

1. Experts Excel Mainly in Their Own Domains. There is little evidence that a person highly skilled in one domain can transfer the skill to another. As Minsky and Papert (1974) noted: "A very intelligent person might be that way because of specific local features of his knowledge-organizing knowledge rather than because of global qualities of his 'thinking' "(p. 59). Evidence for such a conclusion can be drawn from the work of Voss and Post (this volume) on problem solving in political science. In that work, nondomain experts (chemists) solved political science problems much like novices, describing the causes for the problem at a very concrete and specific level, whereas domain experts described more abstract causal categories.

The obvious reason for the excellence of experts is that they have a good deal of domain knowledge. This is easily demonstrated; for example, in medical diagnosis, expert physicians have more differentiations of common diseases into disease variants (Johnson et al., 1981). Likewise, in examining taxi drivers' knowledge of routes, Chase (1983) found that expert drivers can generate a far greater number of secondary routes (i.e., lesser known streets) than novice drivers.

2. Experts Perceive Large Meaningful Patterns in Their Domain. As mentioned, this is apparent in chess, where it is well known that chess masters excel in their recall of the clusters of pieces that they see. This perceptual superiority has been replicated in several other domains, such as in the game of GO (Reitman, 1976), in reading circuit diagrams (Egan & Schwartz, 1979), in reading architectural plans (Akin, 1980), and in interpreting x-ray plates (Lesgold et al., this volume). It should be pointed out, however, that this ability to see meaningful patterns does not reflect a generally superior perceptual ability; rather, it reflects an organization of the knowledge base. Programmers, for example, can recall key programming language words in meaningful clusters (McKeithen, Reitman, Reuter, & Hirtle, 1981), and expert programmers can also recognize and recall familiar subroutines (see Soloway, Adelson, & Ehrlich, this volume).

---

[1] The topic of expertise first appears in major textbooks in cognitive psychology in 1985, in John Anderson's second edition of *Cognitive Psychology and Its Implications.*

3. Experts are Fast; They Are Faster than Novices at Performing the Skills of Their Domain, and They Quickly Solve Problems with Little Error. An easy way to observe the skill of master chess players is to watch them play "lightning chess," where they have only a few seconds to decide on a move. Although studies in the literature actually find experts slower than novices in the initial phases of problem solving, experts solve problems faster overall.

There are at least two ways to explain experts' speed. For simple tasks, such as typing, the speed that experts have acquired comes with many hours of practice, which makes the skill more automatic and frees up memory capacity for processing other aspects of the task (see Gentner, this volume). Thus, they can be fast because they are actually faster at the skill itself or because they have more capacity to perform the total task. The expert typists in Gentner's study were fast because their fingers moved quickly (there were more overlapping movements), as well as because they could free up resources to perform related tasks such as typing degraded pseudowords, whereas novices had few resources available for attending to pseudowords.

A further possible explanation for experts' speed in solving problems rests on the idea emphasized earlier that experts can often arrive at a solution without conducting extensive search. The patterns that chess experts see on the board suggest reasonable moves directly, presumably because, through many hours of playing, they have stored straightforward condition-action rules in which a specific pattern (the condition) will trigger a stereotypic sequence of moves. Cab drivers, for instance, will recognize a shorter route while traveling to their destination, even though they may not have generated this shorter route in the laboratory (Chase, 1983).

4. Experts Have Superior Short-Term and Long-Term Memory. With recently presented materials, experts' recall seems to exceed the limits of short-term memory. This is not because their short-term memory is larger than other humans', but because the automaticity of many portions of their skills frees up resources for greater storage. Experts seem to excel in long-term recall as well. For example, in chess, it is not uncommon for chess masters to recognize plays from certain well-known games.

Chase and Ericsson's (1982) study demonstrated experts' superiority in both short-term and long-term recall. They found that their trained memory expert could remember more than 80 digits in a short-term memory serial recall task. They also found, however, that he could recognize over 80–90% of the digit groups that had been presented to him for recall a week earlier.

5. Experts See and Represent a Problem in Their Domain at a Deeper (More Principled) Level than Novices; Novices Tend to Represent a Problem at a Superficial Level. An easy and robust way to demonstrate this is to ask experts and novices to sort problems and analyze the nature of their groupings. Using physics problems, Chi, Feltovich, and Glaser (1981) found that experts used principles of mechanics to organize categories, whereas novices built their problem categories around literal objects stated in the problem description. Similar results have been found in the domain of programming (Weiser & Shertz, 1983); when expert and novice programmers were asked to sort programming problems, the experts sorted them according to solution algorithms, whereas the novices sorted them according to areas of application (e.g., whether the program was supposed to create a list of employees' salaries or whether it was supposed to keep a file of current user identifications). These results indicate that both experts and novices have conceptual categories, but that the experts' categories are semantically or principle-based, whereas the categories of the novices are syntactically or surface-feature oriented.

6. Experts Spend a Great Deal of Time Analyzing a Problem Qualitatively. Protocols show that, at the beginning of a problem-solving episode, experts typically try to "understand" a problem, whereas novices plunge immediately into attempting to apply equations and to solve for an unknown. What do the experts do when they qualitatively analyze a problem? Basically they build a mental representation from which they can infer relations that can define the situation, and they add constraints to the problem. Paige and Simon's (1966) well-known example illustrates this by asking students to solve simple algebra word problems, such as: A board was sawed into two pieces. One piece was two thirds as long as the whole board and was exceeded in length by the second piece by four feet. How long was the board before it was cut? Paige and Simon found that some students immediately applied equations, which then resulted in their coming up with a negative length; others, however, remarked that the problem was meaningless because one cannot have a board with a negative length. One can conclude that those students who paused had formed a mental model of the situation and made some inferences about the relation between the boards.

The utility of qualitative analysis for adding constraints to a problem can be seen most clearly in ill-defined problems. Voss and Post (this volume) presented economic problems, such as: Imagine you are the Minister of Agriculture for the Soviet Union. Crop productivity has been too low for the past several years. What would you do to increase crop production? About 24% of the experts' solution protocols (those of political scientists specializing in the Soviet Union) were elaborations on the initial state of the problem, as opposed to 1% of the novices' protocols. By elaborating the initial state, the experts identified possible constraints, such as Soviet ideology and the amount of arable land. (Adding constraints, in effect, reduced the search space. For example, introducing the constraint of the amount of arable land eliminated the solution of increasing planting, and considering the constraint of the Soviet ideology precluded the solution of fostering private competition—a capitalistic solution.) Other examples of adding constraints can be seen in the work of Lawrence (this volume) on magistrates' decision-making processes.

7. Experts Have Strong Self-Monitoring Skills. Experts seem to be more aware than novices of when they make errors, why they fail to comprehend, and when they need to check their solutions. For example, the expert physics-problem solver in Simon and Simon's study (1978) would often check his answer, and the expert physics-problem solver in Larkin's study (1983) would often abandon solution attempts before carrying out the mathematical details. Experts' self-knowledge is also manifested in their being more accurate than novices in judging the difficulty of a physics problem (Chi, Glaser, & Rees, 1982). Expert chess players are more accurate than novice players at predicting how many times they will need to see a given board position before they can reproduce it correctly (Chi, 1978). Experts ask more questions, particularly when the texts from which they have to learn are difficult (Miyake & Norman, 1979). Novice learners, on the other hand, ask more questions on the easier materials.

We can argue that, in each of the above examples, the superior monitoring skills and self-knowledge of experts reflect their greater domain knowledge as well as a different representation of that knowledge. We illustrate this dependence on domain knowledge with an example from our own work on physics. As stated, we found that expert physicists were more accurate than novices in predicting which physics problems will prove more difficult to solve. If we probe further and look at the bases on which they made such judgments, we see that they relied on the same knowledge of principles in this task as they used to sort problems into categories. Although about a third of both experts' and novices' decisions about problem difficulties were based on the problems' characteristics (such as "the problem is simplified because it is frictionless"), another third of the experts' judgments were based on the underlying physics principle governing the solution (such as "it's a straightforward application of Newton's second Law"). Only 9% of the novices' judgments were based on the underlying principle. In addition, novices used nonproblem related characteristics (such as, "I've never done problems like this before") about 18% of the time as compared to 7% for the experts (Chi, 1987). The ability of experts to predict accurately which problems were difficult and which were easy enabled them to monitor accurately how they should allocate their time for solving problems. Thus, the monitoring skills of experts appear to reflect their greater underlying knowledge of the domain, which allowed them to predict problem difficulty on the basis of the physics principles rather than less relevant surface features.

Summary

The short history of research on expertise might be written as follows: Information-processing studies of problem solving in the 1960s and 1970s and early work in AI and expert systems accepted a tradition of concentrating primarily on basic information-processing capabilities that humans employ when they behave more and less intelligently in situations in which they lack any specialized knowledge and skill. The pioneering work of Newell and Simon and others richly described these general heuristic processes, but they also offered crucial beginning insight on the learning and thinking of experts, processes that require a rich structure of domain-specific knowledge. In recent years, research has examined knowledge-rich tasks—tasks that require hundreds and thousands of hours of learning and experience. These studies of expertise, together with theories of competent performance and attempts at the design of expert systems, have sharpened this focus by contrasting novice and expert performances. These investigations into knowledge-rich domains show strong interactions between structures of knowledge and processes of reasoning and problem solving. The results force us to think about high levels of competence in terms of the interplay between knowledge structure and processing abilities. They illuminate the set of critical differences highlighted in this overview between individuals who display more and less ability in particular domains of knowledge and skill. We interpret these differences as primarily reflecting the expert's possession of an organized body of conceptual and procedural knowledge that can be readily accessed and used with superior monitoring and self-regulation skills.

Now research needs to go beyond this stage of analysis. We must better understand the properties of domain structure and integrated knowledge; the mechanisms of problem-space definition with minimal search through rapid pattern recognition; and the processes involved in redefining the space of ill-structured and difficult problems. To do so, we should investigate the forms of reasoning and problem-solving strategies that structured knowledge facilitates. We also need to understand how expertise is acquired, how it can be taught, and how beginning learners can be presented with appropriate experience. The papers in this volume consider these themes and represent the type of research that is presently being carried out that investigates both human and artificial expertise.

The Contents of This Volume

The majority of the chapters in this volume were presented at a conference held at the Learning Research and Development Center at the University of Pittsburgh, sponsored by the Personnel and Training Research Program, Office of Naval Research. The conference focused on four areas: practical skills, programming skills, medical diagnosis, and ill-defined problems. In each domain, we selected work that is representative and we sought a diversity of approaches.

Michael I. Posner, in his introduction to this volume, briefly reviews some key readings on expertise. He indicates that the impressive coding and chunking feats of experts are also present more generally in people who have been exposed to a sufficiently large number of experiences to allow performance to become truly automated, and he emphasizes the importance of memory representation for understanding expert performance. He speculates on the role of individual differences in learning abilities that could influence the development of expertise and suggests that the problem of producing an expert may be, to a large extent, that of creating and maintaining the motivation for the long training that is necessary.

Practical Skills

The chapters in the section on practical skills discuss expertise in three areas: typing, memorizing restaurant orders, and mental calculation. Gentner (Expertise in Typewriting) is impressed with the resiliency of expertise in motor skills, and minutely examines the details of the typist's skill. He considers an overlapped processing model of skilled performance that suggests critical roles for parallel mental processes that underlie the typing of successive letters, and the importance of the substantial amount of unused cognitive resources that the automated performance of experts makes available for planning, handling texts that are difficult to read, concurrent phone conversations, and easy response to varied contextual and task demands.

As another example of a practical skill, Ericsson and Polson (A Cognitive Analysis of Exceptional Memory for Restaurant Orders) analyzed the exceptional memory of a headwaiter who was able to remember dinner orders from over 20 people at different tables without extensive aids. Their theoretical framework is the model of skilled memory proposed by Chase and Ericsson (1982). According to this framework, skilled memory requires efficient encoding of presented information using existing semantic knowledge and patterns; the stored information is then rapidly accessed through retrieval cues associated with the encoding during initial storage. In the work reported in this chapter, naive subjects used very different encoding processes that could be described by standard models for free recall, developed to describe memory for unrelated material in laboratory tasks. Of significant interest was the ability of skilled subjects to generalize their skills to other kinds of information. This latter unusual finding suggests the possibility of the existence of transferrable acquired general cognitive processes that can improve memory in a range of situations.

Staszewski's chapter (Skilled Memory and Expert Calculation) examined the extent to which skills in mental calculation, as exercised by people who are proficient at it, are trainable to an average person. Through tracking the learning acquired by two subjects through training, Staszewski found that the principles of skilled memory adequately characterize the way in which mental calculation experts manage the heavy memory demands that arise in mental arithmetic. However, expert-level performance in mental calculation also requires that experts devise strategies to use content information from long-term memory efficiently. Thus, although a mental calculation task does not have the explicit goal of information retention for its own sake, successful performance in that task does require access to more information than short-term memory can hold. The memory skill that underlies a critical component of mental calculation is the proficiency with which individuals can learn to represent and maintain large amounts of task-related information in an easily accessible state—in effect, expand their working memory capacity.

Programming

Three different tasks are examined in the section on programming: understanding, learning, and software design. Soloway, Adelson, and Ehrlich (Knowledge and Processes in the Comprehension of Computer Programs) report on efforts to investigate the knowledge and processing strategies programmers employ in attempting to understand computer programs. They ask: What is it that expert programmers know that novice programmers do not? They focus on two types of knowledge: The first type, programming plans, consists of program fragments that represent stereotypic action sequences; and the second type, rules of programming discourse, consists of rules that specify conventions in programming. These two types of knowledge correspond to the notion of schema and to chunks or patterns that represent functional units in a domain of knowledge. Modeling an experiment after the Chase and Simon chess study, Soloway and his colleagues presented both plan-like programs that conformed to programming conventions and runnable unplan-like programs to experts and novices. The data replicated the chess experiments in that the performance of advanced programmers was reduced to that of novices on the unplan-like material. This finding and others are considered with respect to the development of a measure of program complexity, a model of the processes used in reading and writing programs, and program design.

Anderson, Pirolli, and Farrell (Learning to Program Recursive Functions) discuss work that investigates learning to write recursive functions in LISP. As a framework for their discussion, they describe a model of the programming behavior of an expert. They explain why recursive programming is difficult and propose how it is learned. It appears that recursive programming is difficult because it is a highly unfamiliar mental activity and because it depends on acquiring a great deal of knowledge about specific program patterns. In the instructional context provided, students first studied worked-out examples and then solved similar problems. It was found that students solved problems by mapping analogically the solution of examples to their current problem. They generalized their solutions to a problem by developing new problem-solving operators that could be applied to another problem. Protocols and simulations are presented to give evidence for the conclusions that were made about learning.

Adelson and Soloway (A Model of Software Design) report an analysis of the problem-solving skills of expert software designers and present a model of the process of software design. Their model unites a number of recurrent behaviors found in protocol analysis. These behaviors include the use of mental models that begin as abstract versions of the task and become more concrete as the design progresses, as well as "balanced development" in which the modules or elements of design are defined at the same level of detail. In addition, the model accommodates the finding that expert designers make *memory notes* of constraints, partial solutions, or potential inconsistencies, which eventually they will need to handle. The experts also repeatedly conducted mental simulation runs of partially completed designs. Adelson and Soloway comment on the findings by raising such questions as: Why is balanced development found so frequently in expert behavior? What role does it play and how is it acquired? They see as an unresolved major issue the specification of mechanisms that facilitate the interactions between the domain-independent design model they propose and domain-specific knowledge in particular applications.

Ill-Defined Problems

Three chapters in this section focus on ill-defined problems from very different perspectives. Johnson (Expertise and Decision Under Uncertainty: Performance and Process) considers expertise in an ill-defined problem from the point of view of research in behavioral decision theory. In contrast to the task domains studied in the previous chapters, the tasks he considers require decision under uncertainty, such as when some uncontrolled intervening event occurs between the choice and the outcome. In these tasks, experts are not consistently better than novices, and linear regression models are more accurate than experts most of the time. Two different domains are studied here — the evaluation of applicants for medical internships and the prediction of changes in stock prices. In these tasks, no single or optimally correct procedure exists, only rules that are relatively more accurate under varying circumstances. In general, in these tasks, experts focus on fewer cues than novices, and they use different information and different patterns of search that take advantage of usual, opportunistic information. In his discussion, Johnson considers why expert performance may be inferior to the predictions made by simple linear models.

Lawrence (Expertise on the Bench: Modeling Magistrates' Judicial Decision-Making) describes studies of the judicial decisions of magistrates on the bench. She points out that legal judging is a problem-solving domain where problems are always ill-structured, solutions are inconclusive, and important features of the problem space become apparent from different sources at different times only after initial processing has begun. She describes a technique for analyzing verbal protocols by identifying information selection propositions connected to consequent inferences and decisions in an *if–then condition–action* form. Her model of the judging process involves the magistrates' (a) frames of reference, which include penal philosophies, sentencing objectives, and views of the severity of particular crimes; and (b) external, environmental constraints, such as interpretation of statutory forces, laws of evidence, parliamentary ranges of penalties, and case load. According to her findings, the experts' performance was different from the novices' in terms of the amount and kind of information and goals that influenced the inferences made, based on case details. Although novices knew and responded to ritualized evidence-gathering procedures, they seemed to work with single details, as compared with the more patterned approach of experts. These patterns enabled experts to reduce their work loads.

Voss and Post (On the Solving of Ill-Structured Problems) extend the analysis of previous accounts of the nature of ill-structured problems and comment on Johnson's and Lawrence's work and their own research on problems in social science. Ill-structured problems are described as problems in which there is little consensus regarding the appropriate solutions; they include open constraints that are resolved in the course of solution; and as solution proceeds they may become at some point well-defined. As experts proceed, structure is obtained by decomposing the ill-structured problem into a set of well-structured problems which are then solved. To be able to do this, it is asserted here, the expert must have a relatively larger amount of information in memory so that they can utilize appropriate components of knowledge to organize the problem solution. In their own work on political science problems, such as domestic policy in a foreign country, Voss and Post contrast specialists and novices. Experts develop a problem representation by using the general strategy of decomposition to delineate major factors causing the problem; these factors then are used to convert the problem into one that can be solved. In utilizing this general strategy, experts draw on their knowledge to state a history of previous attempts at solution, and to build a case by enumerating reasons why their solution might work.

Medical Diagnosis

Finally, the chapters on medical diagnosis introduce three very distinct analyses and approaches. Groen and Patel (Relationship Between Comprehension and The Reasoning in Medical Expertise) approach the study of diagnostic expertise in the context of research on comprehension. Theories of comprehension, they claim, have been primarily concerned with structural issues and propositional analyses, whereas theories in the area of problem-solving have been concerned with the explication of processes. The connection between the two needs to be considered in the study of expertise. Groen and Patel conducted a series of studies that used propositional analysis of the recall of textually presented clinical cases to assess differences between experts and novices. Their results indicate that experts make inferences from relevant information, whereas novices infer from less relevant material. With texts scrambled on the basis of propositional structure, the differences between experts and novices disappeared; experts recalled as much irrelevant material as novices. A major conclusion is that the selectivity of experts with respect to relevant information can be explained by the development of a problem representation that filters out irrelevant information.

Lesgold, Rubinson, Feltovich, Glaser, Klopfer, and Wang (Expertise in a Complex Skill: Diagnosing X-ray Pictures) report on expertise in radiological diagnosis, the interpretation of x-ray pictures, which is a skill that involves the integration of knowledge from physiology, anatomy, medical theories of disease, and the projective geometry of radiography. In a series of studies, they observed radiologists in their offices and then moved to more controlled experiments using radiologists with 10 years of experience and residents with 1 to 4 years of training. Their quantitative analysis involved both *findings* – in particular, identification of specific properties of the film or patient, and *relationships* – especially the reasoning paths between findings. As contrasted with the residents, the experienced radiologists showed a greater number of findings, more clustered findings, and larger reasoning chains. Qualitative analysis of protocols led to a general account of the behavior of an expert radiologist. During the initial phase of building a mental representation, a schema entails a set of criteria it must satisfy before it can control viewing and diagnosis. The expert works efficiently toward the stage where an appropriate schema is in control. This schema then contains a set of procedures that allows a diagnosis to be made and confirmed. Novice performance involves incompleteness in each of these three aspects, and novices are less able to modify a schema in response to new information. Lesgold and his colleagues conclude with discussion of the course of acquisition of this complex skill.

Clancey (Acquiring, Representing, and Evaluating a Competence Model of Diagnostic Strategy) discusses NEOMYCIN, a computer program that models a physician's diagnostic reasoning in a limited area of medicine. The diagnostic procedure is represented in a well-structured way, separately from the domain knowledge it operates on. His general objectives are to articulate a design that will enable an expert system to acquire knowledge interactively from human experts, to explain reasoning to people seeking advice, and to teach students. His premise is that these applications, particularly explanation and teaching, necessitate closer adherence to human problem-solving methods and more explicit knowledge representation than systems that are not required to be comprehensible to people. The major section of the chapter considers (a) how the model is acquired—what a representation methodology is for replicating what people know and what they do; (b) a description of the flow of information in the diagnostic model—how, in the course of reasoning, knowledge is activated, problems are formulated, and hypotheses confirmed; and (c) evaluating the model—how well the model matches expert performance and reasoning and the detail of explanation to students. Clancey's chapter clearly describes his search for a design of a knowledge representation that can be used to model human diagnostic reasoning and human explanation capability. In general, he sees such a model as requiring relatively stereotypical patterns that encompass richly structured knowledge about possible solutions and problem features that greatly facilitate search and classification. In addition, the model requires *metacognition* (knowledge for organizing knowledge) that orients the problem solver toward constructing and refining an appropriate problem space.

The chapters in the book display the variety of domains and human performances to which the study of expertise has been carried. The different approaches employed show the influence of methodologies from cognitive psychology, artificial intelligence, and cognitive science in general. The chapters also make a case for increased attention to learning—to how expertise is acquired and to the conditions that enhance and limit the development of high levels of cognitive skill.

ACKNOWLEDGMENTS

Funding for the publication of this volume and the organization of the conference was provided by the Office of Naval Research, Psychological Sciences Division, Personnel and Training Research Program on Cognitive Processes. We greatly appreciate the help of Joyce Holl in typing and organizing all the manuscripts, Shoulamit Milch-Reich for preparing the subject index, and Anne Robin for editing and proofreading.

REFERENCES

Akin, O. (1980). *Models of architectural knowledge.* London: Pion.

Chase, W. G. (1983). Spatial representations of taxi drivers. In D. R. Rogers & J. H. Slobada (Eds.), *Acquisition of symbolic skills* (pp. 391-405). New York: Plenum.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 2-58). New York: Academic Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.

Chi, M. T. H. (1987). Representing knowledge and metaknowledge: Implications for interpreting metamemory research. In F. E. Weinert & R. H. Kluwe (Eds.). *Metacognition, motivation and understanding* (pp. 239-266). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73-96). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121-125.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 17-76). Hillsdale, NJ: Lawrence Erlbaum Associates.

deGroot, A. (1966). Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.), *Problem solving* (pp. 19-50). New York: Wiley.

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

Glaser, R. (1987). Thoughts on expertise. In C. Schooler & W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 81-94). Norwood, NJ: Ablex.

Greenblatt, R. D., Eastlake, D. E., & Crocker, S. D. (1967). The Greenblatt chess program. *Proceedings of the Fall Joint Computer Conference*, *31*, 801-810.

Hayes-Roth, F., Waterman, D. A., & Lenat, D. P. (Eds.). (1983). *Building expert systems*. Reading, MA: Addison-Wesley.

Johnson, P. E., Duran, A. S., Hassebrock, F., Moller, J. H., Prietula, M., Feltovich, P. J., & Swanson, D. B. (1981). Expertise and error in diagnostic reasoning. *Cognitive Science*, *5*, 135-283.

Larkin, J. H. (1983). The role of problem representation in physics. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 75-100). Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

Minsky, M., & Papert, S. (1974). *Artificial intelligence*. Condensed lectures, Oregon State System of Higher Education, Eugene.

Miyake, N., & Norman, D. A. (1979). *To ask a question one must know enough to know what is not known. Journal of Verbal Learning and Verbal Behavior*, *18*, 357-364.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Paige, J. M., & Simon, H. A. (1966). Cognition processes in solving algebra word problems. In B. Kleinmuntz (Ed.), *Problem solving* (pp. 119-151). New York: Wiley.

Reitman, J. S. (1976). Skilled perception in GO: Deducing memory structures from interresponse times. *Cognitive Psychology*, *8*, 336-356.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 325-348). Hillsdale, NJ: Lawrence Erlbaum Associates.

Weiser, M., & Shertz, J. (1983). Programming problem representation in novice and expert programmers. *Instructional Journal of Man-Machine Studies*, *14*, 391-396.

Introduction:
What Is It to Be an Expert?

Michael I. Posner
University of Oregon

How do we identify a person as exceptional or gifted? One aspect is truly expert performance in some domain. An adult or child who composes exceptional music, runs extremely fast, or receives particularly high scores on academic achievement tests, may be said to be gifted or exceptional. Only in the last dozen years or so has experimental research in cognitive psychology and related disciplines begun to discover what is required to be expert in some domain.

EXPERT PERFORMANCE

How did we arrive at this understanding of expertise, and what implications might it have for understanding the nature of giftedness or exceptionality in children?

One of the most striking examples of experimental research into exceptional performance attempts to explain the ability of people to perform exceptional feats of memory. A very simple traditional memory test is to repeat, as accurately as possible, a series of digits that you have just heard. The average college student is capable of repeating about eight of these digits. Memory experts, however, often repeat twenty or more. What is the basis of this exceptional memory? Several years ago, William Chase at Carnegie–Mellon University (Ericsson & Chase, 1982) trained two normal people to remember a sequence of random numbers, so they could repeat it back immediately after presentation. The best subject, after 250 days of practice, could repeat random digit strings as long as 80 items. He did so by a relatively simple technique. He grouped the digits into chunks of three or four digits based on what he thought was a codable running time. Or, as Chase put it, "What S.F. did was begin mentally to encode three- and four-digit groups as running times for various races. For example, he'd remember 3, 4, 9, 2, as 3 minutes, 49.2 seconds, near the world record time for running a mile." These chunks, then, could be grouped into higher level chunks; and finally, when asked to recall, he was able to organize a recall of a very large number of digits where each chunk had only 3 to 4 digits in it. Chase was able to train other expert memorizers in very much the same way. These experts showed normal spans if the items were shifted from digits to letters. In other words, Chase found that what appears to be a very exceptional, perhaps a photographic memory, could be obtained presumably by any normal person whose practice was sustained over many days and who applied a systematic method of coding information in memory.

There is some evidence that even the increase in memory span between the ages of five and adulthood, from about 2 or 3 items to 8 items, also depends on specific learning. Most people believe that the memory span increases because the capability of storing information changes from childhood to adulthood. Chi (1976) reasoned that the change in memory span might result from the child simply learning more about how to code and recognize letters and digits. Compare a brief exposure to a list of 8 arabic digits with a list of 8 roman numerals. The roman numerals are certainly familiar but they take time to name, and thus the number coded in a fixed exposure time is reduced. The studies of Chase and Chi indicate how important specific long continued experience is for expert skill.

Striking evidence on the nature of expertise arose from studies of chess which began nearly 25 years ago, when a famous Dutch chess master, Adrian deGroot (1966), began to study the intellectual capabilities and coding processes of chess masters compared to less expert chess players. He began his study of expertise by the use of protocol analysis. That is, he attempted to have chess masters, experts, and novices speak aloud as they selected a chess move. He then analyzed in detail the depth to which they searched the board, their use of various heuristics (e.g., to control the center of the board), and other aspects of their thought processes. He found relatively little difference between people who were only fair chess players and people who were chess masters. A striking thing was that, although the chess masters picked the right moves, there was nothing in the protocol of their thought processes that seemed to indicate why it was that they were so much better than lesser experts.

Next de Groot did an interesting thing. He asked each of the players to reproduce a chess position after a 5-second exposure to a slide of the 20th move of a chess master's game. These were games with which all the players were equally unfamiliar, but they were the type of game that might be played by masters. He found striking differences in memory. Chess masters could reproduce nearly all the pieces on the board with few errors, while chess experts and novice players had much poorer performance. This work has been replicated many times with chess players and has also been found in other forms of expert performance.

Chase and Simon (1973) studied in detail some of the mental processes involved in this memory performance, and, unsurprisingly, they found much the same result as in the study of expert memory. The reproduction consists of chunks of information which represent units on the chess board. When either the chess masters or the expert players were required to produce a meaningless chess position—that is, one where the pieces had been scrambled—their performances were very much reduced and about equivalent to each other's. Chess masters, like the memory experts, were good in the specific domain of meaningful chess positions. They did not show a greater memory in general, but only in this specific domain.

Cognitive psychologists suppose there are some domains where nearly everyone becomes an expert much like the chess master. Consider reading English words. We all have many hundreds of hours' experience with reading. We can do so effortlessly, essentially automatically. Yet reading is a formidable achievement. A very brief exposure to a set of letters produces a representation of the particular word which was seen. In the expert reader, words are handled so well that letters seen within a word are often more visible than the letter is by itself. Reicher (1969) exposed subjects to either four-letter words (e.g., WORD) or individual letters (D). These were followed by a mask so it became very difficult to see the words or letters. He then gave each subject a forced choice between two letters (D or K) that both made perfectly acceptable English words. Thus, they could not guess based on their knowledge of the English language, but they still did better when the letter was in a word than in isolation. Thus, some of the same impressive coding and chunking feats that are features of expert chess players are also present in those of us who may be less generally gifted, when we have been exposed to a sufficiently large number of trials to allow the performance to become truly automated.

SKILL ACQUISITION

Herbert Simon (Chase & Simon, 1973) has reasoned that masters level chess players have spent 10,000 to 20,000 hours staring at chess positions. To put this in perspective, the student who spends 40 hours a week for 33 weeks spends 1320 hours a year studying. Imagine spending more than 10 years in college studying one subject, chess, and you get some appreciation of the time commitment of master level players. As a result of such extensive study, they are believed to store from 10,000 to 100,000 different chess patterns. Simon concluded that it is reasonable to assume a chess master can recognize 50,000 different configurations of chess, not too far different from the number of different words an English reader may be able to recognize.

Fortunately, we now know something about how information is represented in the memory system during learning (Anderson, 1983). Consider learning a set of statements such as "a doctor is in the bank," "a fireman is in the park," "a lawyer is in the church," "a lawyer is in the park." This set of sentences can be represented in terms of a propositional network. What is important is that the concept "lawyer" is related both to its location in the church and its location in the park. There are two relationships stemming from the same concept.

How do we know the structure of the memory? One clue is that the more ideas associated, the longer the time it takes to retrieve any particular item. Thus, experimental studies having subjects learn relationships of this sort, and asking them then to answer as quickly as possible such questions as "is there a doctor in the park?" show that the time needed to answer any question is a function of the number of relationships associated to the idea—in this case, "*doctor.*"

This finding suggests that propositions about lawyers are tied to a single concept in memory. Clearly there is a paradox—the more items attached to any one concept, the longer the retrieval time, and yet experts do not necessarily take longer to retrieve information. In part, this paradox is overcome by the long practice which is associated with becoming an expert. The reaction time for retrieving information improves with practice, and comes to be independent of the number of propositions attached to a node. The expert must also unify the stored information into a meaningful whole that allows it to be retrieved more rapidly. Usually we are not asked to retrieve exact information, but instead are asked questions which might be inferred or thought to be consistent with what we know.

Experimental studies suggest that when subjects are required only to say whether information is consistent with what has been learned, the more information they have, the more quickly such information can be retrieved. This may also be a way in which the expert can quickly answer questions which would require a study of the individually stored propositions on the part of nonexperts. These studies of memory representation after long training may begin to give us methods to assess and guide the training of expertise.

EXPERT SYSTEMS

Our understanding of the specific nature of human expertness has progressed far enough so that we are beginning to see the development of artificial computer-based systems able to achieve some of the same performance as the expert. These systems take advantage of the fact that digital computers are general-purpose processors of symbols (Duda & Shortliffe, 1983). Some of these expert systems are being put to use in tasks like deciding which antibiotic to prescribe, testing pulmonary functions, or guiding geological exploration.

These findings about the importance of representing information in memory for understanding expert performance have been embodied in these computer systems. In the article just cited, Duda and Shortliffe say:

> The early hope that a relatively small number of powerful general mechanisms would be sufficient to generate intelligent behavior gradually waned. When significant problems were addressed, it was often discovered that problem independent, heuristic methods alone were incapable of handling the sheer, combinatorial complexity that was encountered. Similarly, general problem solving techniques confronted in precisely stated "problems," uncertain "facts," and unreliable "axioms" were found to be inadequate to the task.
>
> When it was asked how people were able to devise solutions to these problems a frequent answer was that people possess knowledge of which the programs were wholly innocent. This knowledge is employed in a variety of ways. . . in clarifying the problem, suggesting the kinds of procedure to use, judging the reliability of facts and deciding whether a solution is reasonable. (p. 261)

The growing recognition of the many kinds of knowledge required for high performance reasoning systems changed the shape of AI research. In the words of Goldstein and Pappert, as quoted by Duda and Shortliffe:

> Today there has been a shift in paradigm. The fundamental problem of understanding intelligent behavior is not the identification of a few powerful techniques, but rather the question of how to represent large amounts of knowledge in a fashion that permits their effective use and interaction. The current point of view is that the problem solver, whether man or machine, must know explicitly how to use its knowledge with general techniques supplemented by a domain specific pragmatic knowhow. Thus, we see AI as having shifted from a power based strategy for achieving intelligence to a knowledge-based approach. (p. 262)

INDIVIDUAL ABILITY

The burden of this work on expertness is a hopeful one. Ordinary people seem to have within them the potentiality for expertise, should they be able to acquire the large technical vocabulary and make the long commitment of study that such expertise requires. But this can hardly be the whole story. When we have the insight that children are gifted or exceptional we suppose not just that they are experts in some domain, but that they have the capacity to acquire expertise in many domains more quickly than others who may not be able to acquire it at all. The studies that I have presented so far do not support this insight, but there are other studies which do.

Lyon (1977) studied the very same memory span tests I referred to in the work of Chi and Chase. Lyon studied normal adults who differ widely in their memory span. Memory span is a skill which is correlated with the overall scores that people obtain in intelligence tests. Lyon wished to know what is the basis of these individual differences. The burden of the studies of Chi and Chase would suppose that it must lie in either overall experience or some specific strategy such as chunking or grouping material. Lyon induced several specific strategies in his subjects, and their performance greatly increased. However, the individual differences remained pretty much the same. The strategies, although important, did not erase the seemingly more basic differences among individuals.

Most of us maintain an intuition that there are important underlying differences among people in how well they could develop expertise. Even within individuals, it's hard to believe that there is not more potential for the development of some capabilities than of others. Some people seem to learn music easily, but have trouble with languages, while others acquire mathematics effortlessly, but do poorly in mechanical ability. A hundred years of testing of intelligence by psychometricians is based on an effort to build a measurement technology based upon this intuition.

Although the psychometric approach to intelligence testing was not related to any theory of intelligence, more recent efforts by cognitive psychologists (Hunt, 1983) are based upon the theory of symbolic representations that I have been discussing. They seek to measure cognitive processes in various domains. It seems clear that some people are systematically faster than others in retrieving the names of letters or words, in performing mental operations on visual images, and in comprehending written and oral words. Correlations are often found between processing speed and psychometrically measured intelligence (Jensen, 1979). This has led some to argue that we now have clear measures of fundamental mental operations that are culture-free and that give us a more direct approach to the efficiency of the neural systems underlying cognition. Indeed, there are some impressive results tieing studies of speeded processing in normal and brain injured humans that suggest deep ties between the performance of elementary mental operations and brain function, particularly in reading (Coltheart, 1981).

Yet I remain skeptical that the measurement so far achieved can be thought to be free of past learning in any fundamental sense. Remember that Chi found the improvements over 15 years in the speed of recognizing digits could influence even the basic memory span. It can be argued that many of the tests employed to study elementary operations similarly reflect past experience with symbols, training in maintenance of alertness, and conforming to the experimenter's goals.

The problem of producing an expert may be not so much in selecting someone who has special capability, but to create and maintain the motivation needed for long-continued training. Whether someone will work hard is itself a possible basis of individual differences. Only motivated people are likely to undergo the long training necessary to become a chess master. Perhaps we should be using these new cognitive tests of mental operations to determine which domains of material seem to be the most promising for a given individual—understanding that these may be properties not only of innate abilities, but of interest in the field. From this perspective, it will be important to assess motivation, as well as capability of learning, within any domain. Tests of speed of mental operations may also help us assess the structure obtained at any level of practice, and guide the exercises needed to aid in the development of expertise.

SUMMARY

I begin this chapter by noting that there is now a cognitive science related to the representation and execution of expert performance. This science has developed a technology in the form of programs for performing tasks formerly done only by experts. Although this technology is still primitive, it represents an important contribution of fundamental research on the nature of representation of information in memory. Behind this technology is a better understanding of what it means to be an expert. Expertness lies more in an elaborated semantic memory than in a general reasoning process. Such knowledge is present not only in the performance of unusual people, but in a skill like reading which is widely distributed in most of us. We are beginning to understand the nature of the propositional network underlying such representation. The expert has available access to a complex network without any conscious representation of the search processes that go on in its retrieval.

Despite the overwhelming emphasis in the recent cognitive literature on the ability of any person to achieve expert performance with practice, there is still considerable evidence in the literature that individuals may differ in overall ability or particular abilities. Our new knowledge about the nature of expertness should suggest ways in which these basic capabilities and interests may interact with the acquisition of information in the production of expert performance. Perhaps they will also give us a method for sustaining the necessary motivation to achieve truly expert performance.

ACKNOWLEDGMENTS

This chapter was written for a conference of teachers of gifted children held at the University of Oregon in the summer of 1983. The general conference stressed methods of locating children with extraordinary abilities. In this paper I tried to show that psychologists were also trying to understand how expertise might be acquired by people who were otherwise average in their ability. It is not based on original research of my own but was based upon relevant literature.

I am grateful to Dr. Mary K. Rothbart for her help with the manuscript and to Dr. M. T. H. Chi for suggesting that it might be of interest to a wider audience.

REFERENCES

Anderson, J. R. (April 1983). Retrieval of information from long-term memory. *Science, 220,* 25–30.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81.

Chi, M. T. H. (1976). Short-term memory limitations in children: Capacity or processing deficits? *Memory & Cognition, 4,* 559–572.

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language, 15,* 245–286.

deGroot, A. D. (1966). Perception and memory versus thought. In B. Kleinmuntz (Ed.), *Problem Solving Research, Methods and Theory* (pp. 19–50). New York: Wiley.

Duda, R. O., & Shortliffe, E. H. (1983). Expert systems research. *Science, 220,* 261–268.

Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory. *American Scientist, 6,* 607–612.

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

Jensen, A. R. (1979). *Bias in Mental Testing.* Riverside, NJ: Free Press.

Lyon, D. R. (1977). Individual differences in immediate serial recall: A matter of mnemonics. *Cognitive Psychology,* 9, 403–411.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 275–280.

1 Expertise in Typewriting

Donald R. Gentner
University of California, San Diego

INTRODUCTION

The other chapters in this book are concerned primarily with expertise in mental tasks. Even though an expert waiter or radiologist may use motor skills, such as speech and handwriting, the motor skills themselves are not of direct interest to most investigators. This chapter, on the other hand, is concerned with the acquisition and performance of the motor skill of typewriting. Motor skills provide a unique psychological insight, because they are the direct, concrete product of the large amount of mental processing required for the planning, coordination, and control of actions. From a practical standpoint, motor skills offer a unique advantage to the scientist studying expertise. Most of the interesting events in mental skills go on inside the head, and are hidden from our view. The scientist must make indirect inferences about these mental events from such data as reaction times and verbal protocols. In contrast, the normal performance of a motor skill produces an externally observable sequence of events that are directly related to the task.

It is clear from anatomical studies of the brain, and observation of patients with brain injury, that even in humans a large portion of the brain is involved in the performance of motor skills. Some motor skills, such as walking and speech, develop in childhood as the motor system itself develops, and are normally acquired without special effort. Other motor skills, such as juggling, playing piano, or flying an airplane, although based on existing perceptual and motor skills, require special instruction to acquire and gain expertise. Expertise in typewriting belongs in the latter class.

Prospective typists spend hundreds of hours in classes and at practice before they are expert enough to be employed. In fact, when typewriters were first manufactured, they were operated by the hunt-and-peck method. It took at least another 20 years before it was generally realized that it was even possible to type using all ten fingers and without looking at the keyboard.

A typical professional typist has accumulated an incredible amount of practice. A conservative assumption would be that a typist averages 50 words per minute (wpm) for 20 hours per week. Over the course of 10 years, that would amount to 150 million keystrokes or 25 million words. In those 10 years, this hypothetical typist would have typed the word *the* 2 million times, and typed a common word like *system* 10,000 times. The speed of professional typists is also quite remarkable. A typing rate of 60 wpm corresponds to an average of five keystrokes per second. The fastest typists I have studied maintain an average of more than nine keystrokes per second over the period of an hour.

ACQUISITION OF TYPEWRITING

In common with the other tasks described in this book, it takes people a surprisingly long time to become expert typists. The performance norms listed by West (1983, p. 346) give the following median typing speeds for students: 38 wpm for students completing the first year of high school typing, 44 wpm for students completing the second year of high school typing, and 56 wpm for students at the end of business school training. (These scores are gross words per minute, with no correction for errors.) The surprising finding is that after 3 years of practice, the median graduate of business school is just barely meeting minimal employment standards. Estimating 5 hours of practice per week and 40 weeks per year, in 3 years a student would have accumulated about 600 hours of practice on the typewriter.

It's instructive to contrast the time required to become an expert typist with the time required to learn to fly an airplane, which is generally acknowledged to be a reasonably difficult motor skill. A private pilot's license requires only 35 hours of flight time. Even combat pilots in the U.S. Air Force have only 300 to 350 hours flying time plus another 75 hours of simulator training when they report to their operational squadron (D. Lyon, personal communication, August, 1983). Of course there are probably motivational and aptitude differences between pilot trainees and typing students, but the similarity in acquisition times makes clear that learning to type at the professional level is not an easy task.

Like other motor skills, typewriting, once acquired, is remarkably resilient. In a classic series of motor learning studies, Hill (1934, 1957; Hill, Rejall, & Thorndike, 1913) recorded data from three month-long efforts to learn typewriting that were separated by lapses of 25 years. Hill found significant savings of skill at the beginning of the second and third learning efforts, despite the intervening 25 years between efforts. Salthouse (1984) studied the performance of professional typists ranging in age from 19 to 68 years. He measured performance of the typists on a battery of tasks, including a forced-choice reaction time task on the typewriter keyboard and a normal transcription typing task. Salthouse found that performance in the transcription typing task was not correlated with age, even though performance on supposedly similar motor tasks, such as tapping speed and forced-choice reaction time, showed the usual decline with age.

COMPARISONS OF EXPERT AND NOVICE TYPISTS

How do expert typists differ from novices? I've examined this question by comparing the performance of student typists and professional typists. For most of the studies reported here, the typists were asked to transcribe normal prose texts for about an hour. They typed on an electronic keyboard with a layout and "feel" similar to the IBM Selectric keyboard (Figure 1.1). Keystrokes and the corresponding times were recorded by a microcomputer with a resolution of 1 msec. The typists' finger movements were recorded on videotape.

The student typists were volunteers from the first semester typing class at a local high school. They came to the laboratory once a week between the fourth and eighth weeks of class. The expert typists were professional typists recruited from the university and from local businesses. Most of the experts were typical office secretaries, but a special effort was made to recruit a few very fast typists.

Faster Interstroke Intervals

The first measure of keystroke timing examined was the distribution of interstroke intervals. Figure 1.2 illustrates the range of distributions found among typists, showing the distribution of interstroke intervals for a student (Typist 21) at two points in time, a typical office typist (Typist 2), and an unusually fast typist (Typist 8). This figure also demonstrates the most obvious difference between novice and expert typists: experts type faster than novices. The typing speed of the students participating in this study ranged from 13 wpm for one student in the fourth week of class to 41 wpm for another student in the eighth week of class. (The typing speeds reported in this chapter are gross words per minute, with no correction for errors. A word was taken to be five characters, including spaces.) The typing

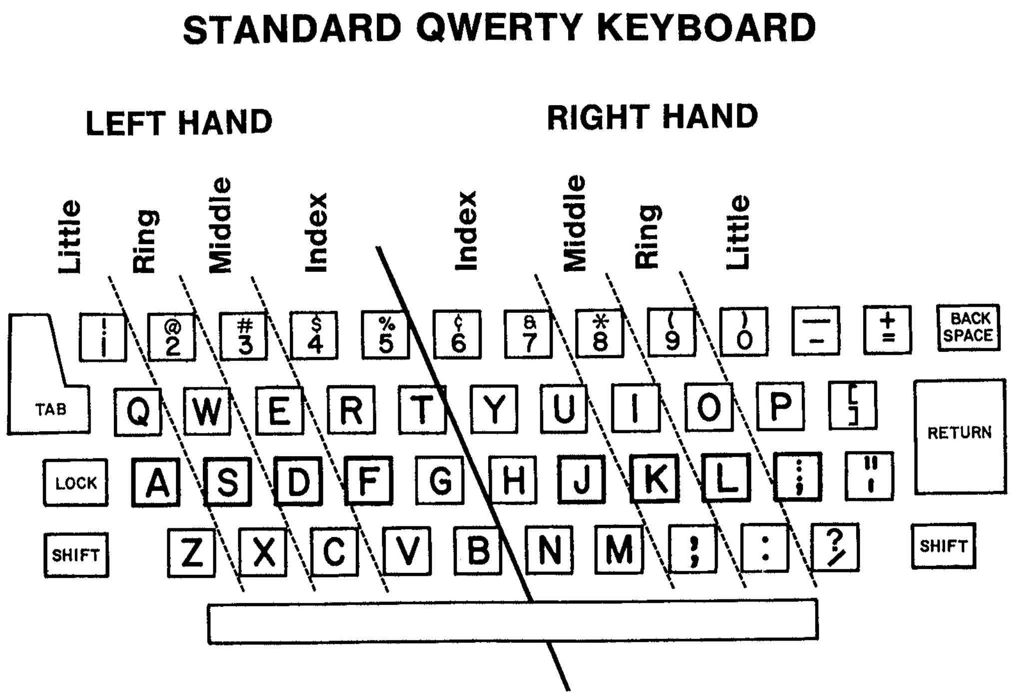

FIGURE 1.1 The layout of the keyboard used in these studies. This is the standard American ''qwerty'' keyboard and is identical to the layout of the IBM Selectric typewriter.

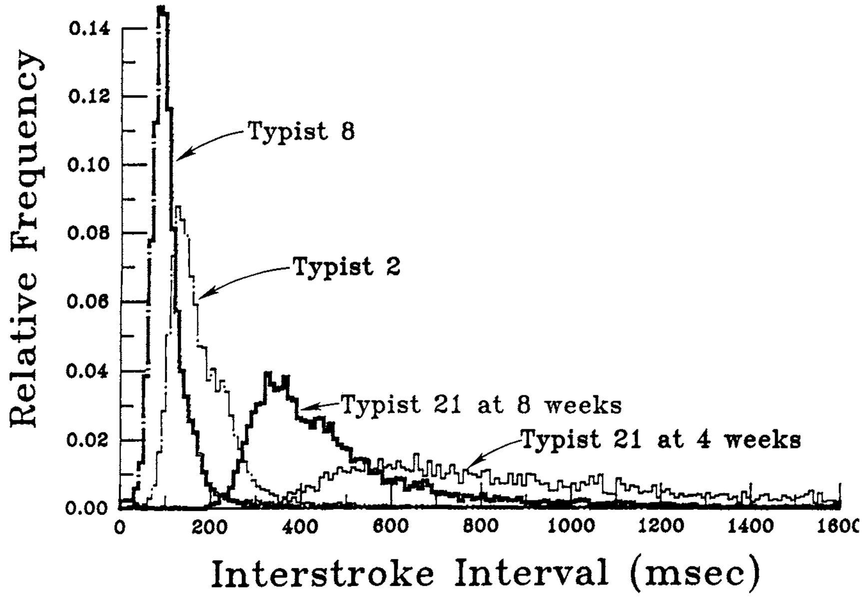

FIGURE 1.2 The distribution of all interstroke intervals for Typist 21 after 4 weeks (13 wpm) and 8 weeks (25wpm) of typing class, Typist 2 (66 wpm), and Typist 8 (112 wpm).

speed of the expert typists ranged from 61 to 112 wpm. In addition to being faster, the expert typists generally had a much lower error rate than the student typists.

How does the performance of the expert typists compare in detail with the performance of the student typists? Is expert performance simply a speeded-up version of student performance, or do qualitative changes in performance occur during acquisition of typing skill? As one approach to these questions, consider the simple movement required to type two letters in sequence with the same finger, such as *de*. The *de* interstroke intervals of experts were more than twice as fast as the *de* interstroke intervals of students. There are only three basic ways that an expert could type the *e* more quickly: (a) the finger movement to type the *e* could start earlier; (b) the finger could travel a shorter path; (c) the finger could move faster.

To investigate this issue, I have examined the videotape records of the expert and student typists when typing the digraph *de*. The study included eleven expert typists ranging in speed from 61 to 112 wpm, and eight student typists in the seventh week of their typing class, ranging in speed from 17 to 40 wpm. For each typist, the 10 instances of *de* (5 instances in the case of student typists) with interstroke intervals closest to that typist's median *de* interstroke interval were selected for study. For each instance, the position of the left-middle fingertip was digitized on the videotape recording and the trajectory was calculated in three-dimensional space. The time of the first visible movement toward the *e* key was determined from a plot of the finger trajectory. Three measures were calculated for each trajectory: (a) the lag time—the time from the initial depression of the *d* key until the first visible movement toward the *e* on the top row; (b) the path length—the distance moved by the fingertip from the beginning of the movement until the *e* keypress; (c) the average speed of movement—the path length divided by the movement time. The results are shown in Table 1.1. Surprisingly, the mean path length of the students was slightly shorter than that of the experts, so the experts were not typing the *e* more quickly because of a shorter path. Instead, the experts started their finger movements with a shorter lag time after pressing the *d* (accounting for about 60 msec of the difference in interstroke intervals), and moved their fingers about twice as fast (accounting for the remaining 150 msec).

This picture develops an interesting twist when the data are examined for each group separately. An analysis of the correlations between the interstroke interval and the three measures (See Table 1.2) showed that the speed of finger movement was the primary determinant of the interstroke interval for the students ($r = -.92$). For the expert typists, however, speed was not correlated ($r = .06$) with the interstroke interval. Although there was considerable variation in speed among the experts (mean speeds ranged from 231 to 524 mm/sec), the typists with higher speeds also had longer path lengths, and the two factors cancelled out. Instead, the primary determinant of the interstroke intervals among experts was the lag times. The fastest experts had very short lag times between pressing the *d* key and starting the movement toward the *e* key.

Differential Speedup of Digraph Classes

The results described for the digraph *de* are consistent with the view that expert performance is simply a speeded-up version of student performance. Recall that the experts and students had similar path lengths, but the experts had shorter lag times and moved their fingers faster than the students. When other types of digraphs are examined, however, this simple picture

TABLE 1.1
Mean Characteristics of "de" Finger Movements

	Interstroke Interval (msec)	*Lag Time* (msec)	*Path Length* (mm)	*Average Speed* (mm/sec)
Students	384	104	38	152
Experts	178	46	45	353

TABLE 1.2
Correlations with Interstroke Inverval (Within-Group)

	Lag Time	*Path Length*	*Speed*
Students	−.18	+.51	−.92
Experts	+.74	+.41	+.06

of expert performance is no longer adequate. Although experts typed all sequences faster than students, the increase in speed was not equal for all interstroke intervals.

For this analysis, it is useful to divide the digraphs into classes according to the fingers used to type them (the keyboard is shown in Figure 1.1). Repeated letters, such as *dd*, are called doubles. The remaining (nondouble) digraphs typed by one finger, such as *de*, are called one-finger digraphs. Digraphs typed by two fingers on the same hand, such as *dr*, are called two-finger digraphs. And finally, digraphs typed by fingers on opposite hands such as *do*, are called two-hand digraphs.

Figure 1.3 shows the median interstroke intervals for the various digraph classes as a function of the typist's overall speed. Doubles were the fastest digraph class typed by students, but were among the slowest digraphs typed by experts. The other digraph classes were all typed at about the same speed by the slowest students, but were typed at significantly different speeds by experts. One-finger digraphs were typed the slowest by expert typists, and two-hand digraphs were typed the fastest. As overall typing speed increases, the median interstroke intervals get shorter for all digraph classes, but the amount of reduction in the interstroke interval varies, depending on the digraph class. Across this group of typists, the interstroke intervals for doubles decreased by a factor of 3 from the slowest to the fastest typists. By contrast, the interstroke intervals of two-hand digraphs decreased by a factor of 12. The interstroke intervals of one-finger and two-finger digraphs showed intermediate improvement, decreasing by factors of 6 and 10, respectively.

Consideration of the finger movements required to type these digraphs suggests a mechanism for the differential improvement in interstroke intervals. With two-hand digraphs, which showed the greatest improvement, it would be possible to overlap finger movements, so that the finger movement for the second letter could be started before the first letter was typed. Alternatively, at least the movement to type the first letter with one hand should not interfere with the movement to type the second letter with the other hand. In contrast, doubles and one-finger digraphs, which showed the least speed improvement, are typed by a single finger and thus no overlapped movements are possible.

The possibility of overlapped movements for two-finger and two-hand

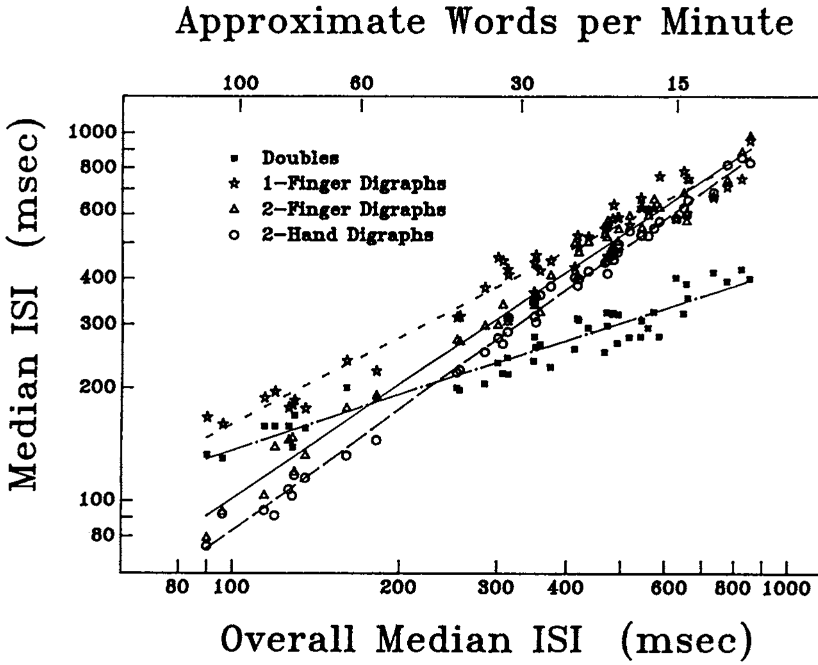

FIGURE 1.3 The median interstroke interval for double, one-finger, two-finger, and two-hand digraphs plotted as a function of the typists' overall median interval. The fastest typist (112 wpm) is on the left; the slowest typist (13 wpm) is on the right. The data on the left are from 10 skilled typists; the data at center and right are from 37 sessions with 8 student typists in the fourth through eighth week of a beginning typing class. The typists were copying normal prose. The data plotted are based on approximately 12,000 keystrokes per typist for the skilled typists, and from 3,000 to 6,000 keystrokes per typist for the student typists. Note that one-finger doubles were among the slowest for skilled typists but fastest for the students.

digraphs was confirmed by analysis of videotape and high-speed film records of typists' finger movements (Gentner, 1981; Gentner, Grudin, & Conway, 1980). Numerous instances were found in the videotapes of expert typists when two, or occasionally three, keystrokes were in progress at one time. The overlapping of finger movements in time is not the only way a typist can take advantage of the ability to move fingers independently. When successive letters are typed on different rows of the keyboard, moving the whole hand to type the first letter can carry the other finger out of position for the second letter. There were many cases on the videotape records where no overlapped movement was seen, but digraphs typed by different fingers on the same hand were faster because the second finger was not pulled out of position by the keystroke of the first finger (Gentner, 1983).

Increase in Overlapped Movements

The extent of overlapped finger movements varied considerably from one expert typist to another, and was moderately correlated with typing speed. Unfortunately, the direct determination of the extent of overlapping finger movements from the videotape records is very time consuming. I have, therefore, tried to estimate the extent of overlapped movements from the interstroke intervals. Although this is an indirect measure, it at least has the virtues of ease and objectivity. The basic assumption, in this measure of the extent of overlapped finger movement, is that the interstroke interval for a normal one-finger digraph represents the time for a keystroke with no overlap. Interstroke intervals for two-finger and two-hand digraphs are normally shorter than for one-finger digraphs, and this estimate assumes that these shorter intervals are the result of overlapped movements. Thus, for each typist, I calculated a "cross-hand overlap index" by taking the difference between the median interstroke intervals for one-finger digraphs and for two-hand digraphs, and dividing by the median one-finger interstroke interval. I also calculated a "within-hand overlap index" in an analogous fashion, as a measure of the amount of overlapping movement between two fingers on the same hand, by comparing the median two-finger and one-finger interstroke intervals. Figure 1.4 shows these cross-hand and within-

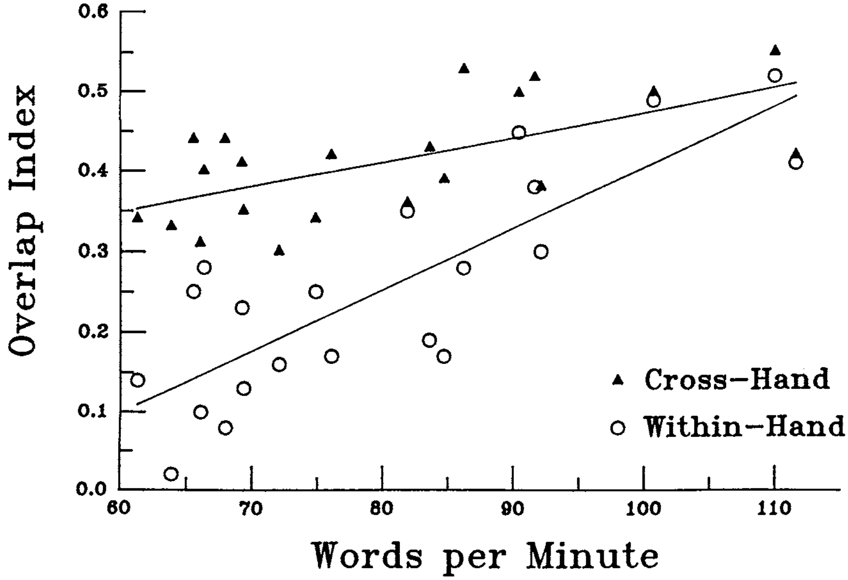

FIGURE 1.4 Cross-hand and within-hand overlap indices for expert typists, as a function of typing speed. The overlap index, plotted on the vertical axis, is a rough estimate of the fraction of a keystroke that overlaps the previous keystroke. Although the amount of cross-hand overlap increases with increasing speed, the increase in within-hand overlap is much greater.

hand overlap indices for a group of 21 expert typists with varying typing speeds. Although the absolute values of these overlap indices should not be taken too seriously, they appear to be a reliable measure of the relative extent of overlapped movements exhibited by different typists for cross-hand and within-hand movements. Figure 1.4 indicates that there was a modest increase in cross-hand overlapped movements as expert typists increased in speed from 60 to 112 wpm ($r = +.63$). The increase in within-hand overlapped movements with typing speed ($r = +.82$) was much greater, however, and appears to be a major contributor to the high speed of the fastest experts. Within-hand overlapped movements are negligible for the typists in the range of 60 wpm, but the fastest typists show as much overlapped movement within-hand as across-hand. This trend is also evident on close examination of Figure 1.3: The fastest typists have almost identical interstroke intervals for two-finger and two-hand digraphs.

The large differences in within-hand overlap among expert typists is related to another finding. The median interstroke intervals for two-finger digraphs were more variable among expert typists than any other digraph class. This variability was based on differences in the degree to which the fingers within a hand were moved independently. Analysis of the videotape records showed that typists with rapid interstroke intervals for two-finger digraphs moved their fingers independently or actually overlapped finger movements within a hand. Typists who had slow interstroke intervals for two-finger digraphs tended to move all the fingers on a hand together and thus their other finger was often out of position to easily type the second letter of a two-finger digraph (Gentner, 1981).

Simulation of Acquisition

This view, that the differential speedup of digraph types is based on the possibility of overlapped movements, is supported by results from the computer simulation of a typist developed by Rumelhart and Norman (1982). Their typing simulation model is based on a parallel, distributed view of cognitive processes, and does not have any central planning or timing control. Instead, their simulation model attempts to type several characters at once, and the interstroke intervals are a result of competition and collaboration among concurrent goals to move the fingers to the different keyboard locations.

Producing a sequence of events in the proper serial order has always been a problem for theories of action (Lashley, 1951). In the Rumelhart and Norman simulation model, the proper serial order is obtained by having each letter inhibit all following letters. Thus the first letter in a sequence, because it is not inhibited by any other letters, will normally be the most highly activated letter and will be typed first. The second letter will be inhibited by only one letter to the left and will normally have the next highest activation, and so fourth.

D. Rumelhart (personal communication, 1982) found that, if the amount by which a given letter inhibited the following letters was varied, the simulation model showed a pattern of changes similar to the pattern of changes found in going from student to expert typists (see Figure 1.5). Decreasing the amount of inhibition between successive letters in the model has the effect of increasing the degree to which several letters tend to be typed at once. When the simulation model had a high level of inhibition between successive letters, and thus tended to type one letter at a time, the pattern of interstroke intervals was similar to the pattern observed with student typists. Whereas when the level of inhibition between successive letters was low, causing the simulation model to attempt to type several letters simultaneously, the pattern of interstroke intervals was similar to the pattern observed with expert typists. Thus, this simulation result suggests that an important component of the acquisition of typewriting skill is the change toward a less sequential and more overlapped mode of performance.

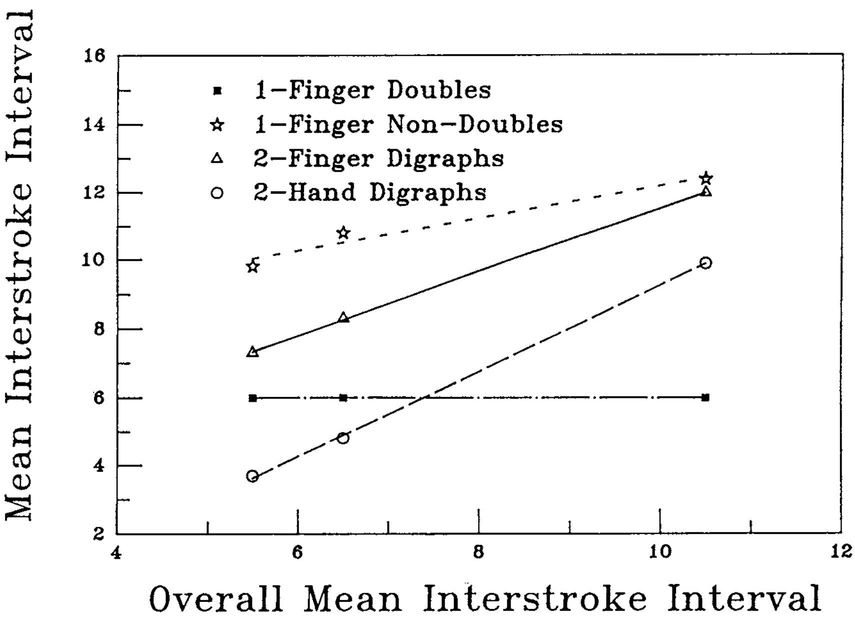

FIGURE 1.5 The effect of changing the amount of inhibition between successive letters in the Rumelhart and Norman (1982) simulation model of a typist. Points on the right have the most inhibition; points on the left have the least inhibition. Decreasing the amount of inhibition decreases the average interstroke intervals and also changes the pattern interstroke intervals for the different digraph classes. Compare this figure with Figure 1.3.