University of Leicester

COPYRIGHT & COURSE PACKS SERVICE
DOCUMENT SUPPLY - MAIN LIBRARY

Copyright Notice

Staff and students of this University are reminded that copyright subsists in this extract and the work from which it was taken. This Digital Copy has been made under the terms of a CLA licence which allows you to:

- access and download a copy;
- print out a copy;

This Digital Copy and any digital or printed copy supplied to or made by you under the terms of this Licence are for use in connection with this Course of Study. You may retain such copies after the end of the course, but strictly for your own personal use.

All copies (including electronic copies) shall include this Copyright Notice and shall be destroyed and/or deleted if and when required by the University.

Except as provided for by copyright law, no further copying, storage or distribution (including by e-mail) is permitted without the consent of the copyright holder.

The author (which term includes artists and other visual creators) has moral rights in the work and neither staff not students may cause, or permit, the distortion, mutilation or other modification of the work, or any other derogatory treatment of it, which would be prejudicial to the honour or reputation of the author.

Course of Study: MSPSSOCDS_D - Occupational Psychology & Psychology of Work (Oct09)

Name of Designated Person
authorising scanning: Robert Melocha

Title: Chapter 1 'Acquiring and performing cognitive skills' in Acquisition and performance of cognitive skills, 1989, pp. 1-16.

Name of Author: A.M. Colley and J.R. Beech (Authors & Editors)

Name of Publisher: John Wiley and Sons Ltd

Name of Visual Creator (as appropriate): Tania Rowlett

Additional notes (if applicable): Possible underlining/marking visible which could not be removed from the original document.

Document scanning reference: MSPSSOCDS_DOCT09 – Doc 14

Date Scan Produced: 29 July 2010

Chapter 1

Acquiring and performing cognitive skills

Ann M. Colley and John R. Beech

INTRODUCTION

> . . . skill lies in the use of capacities efficiently and effectively as the result of practice and experience. (Welford, 1976, p. 14)

Researchers and laypeople alike would agree that activities are said to be skilled when the performance of them has reached a level where it appears to be effortless, where it is almost always accurate and where additional practice makes little apparent improvement. Fitts and Posner (1967) have proposed additional characteristics: skilled performance is organized spatially (if it has a motor component) and temporally; is goal-directed, and uses feedback for error correction unless there is insufficient time for detection and correction to take place.

Until relatively recently most studies of skilled performance focused on perceptual motor skills. Other kinds of activity which met the criteria outlined above were mostly neglected by investigators looking at learning mechanisms. One particular area of neglect has been that of cognitive or intellectual tasks which were studied both in applied educational contexts and in expert-novice comparisons in the laboratory but the mechanisms underlying their acquisition were not, until recently, studied from the perspective of trying to integrate theories and models with the skills literature in general. Welford (1976) identifies three types of skill which correspond to stages in information processing models of serial processing. Perceptual skills code and interpret incoming sensory information. Motor skills execute skilled movement

Acquisition and Performance of Cognitive Skills. Edited by A. M. Colley and J. R. Beech

efficiently but are reliant on appropriate links between sensory input and action routines. Intellectual skills link perception and action and are concerned with translating perceptual input into a skilled response by using appropriate descisions. Welford concludes that the majority of the most important skills fall into the last category. Why, then, have these intellectual skills received relatively little attention until the last decade? One reason for this may well be that much of their performance is invisible. The skilled problem-solver or mathematician produces a solution. In contrast, the skilled pianist produces a prolonged sequence of hand, finger and foot movements which are clearly visible and are outside the expertise of most of the spectators, in addition to a highly complex pattern of sound. Of course this is not the only reason why the area of cognitive skill developed only relatively recently. The interest in perceptual motor skills which arose in the 1950s and 1960s resulted from sponsorship by industrial and military sources interested in developing and improving weapon and radar systems which had been invented or improved during and after the Second World War. Essential to the design of these, was an understanding of the physical and mental capacities of the individuals operating them, and of the best way to train operators. Indeed, much of the contemporary literature on training is based firmly on research conducted at this time (see Annett's chapter).

The more recent emphasis on cognitive skills has arisen from the increasing use and power of computers. It is unlikely that many useful theoretical advances could be made in this area without the use of computer simulation. Measuring a weapons operator's ability to track a target, or a radar operator's ability to track an aeroplane on an oscilloscope is relatively easy to do using laboratory analogues. Studying how a chess master makes a decision to make a particular move, or how a consultant physician makes a diagnosis of a patient's ailments is more problematic. One methodology, which has gained in popularity since the publication in Ericsson and Simon's book in 1984, uses protocol analysis in which the investigator uses protocols as the basis for inferences about underlying cognitive processes. Ericsson and Oliver in Chapter 8 describe this methodology and how it can be used in the study of memory skills, particularly in exceptional cases. The growth of artificial intelligence has opened up new possibilities for the testing of theoretical ideas. Its influence can be seen in research on problem-solving and on computing. The chapters by Gilhooly and Green and by Elsom-Cook summarize the state of research in their respective fields, and Elsom-Cook argues strongly for the advantages of using simulations to study performance and to make predictions about behaviour.

Cognitive skills involve the effective and efficient translation of information into a response. To accomplish this, it is necessary to interpret the information in terms of current knowledge and to have procedures available to enact the steps necessary to make the translation. The same is true of any problem-solving task, and Anderson (1982) expresses the view that skill acquisition is synonymous with problem-solving. Acquisition of a cognitive skill involves the acquisition of a set of domain-specific rules which allow the solution of a particular problem. The end result of cognitive processing is a decision or a solution. In either case a problem has been solved. Even a fairly universal skill such as reading can be viewed in this way (see Beech's chapter). The framework offered by Van Dijk and Kintsch (1983) for understanding discourse comprehension assumes that the comprehension process consists of a hierarchy of strategies which perform various levels of analysis on written material.

Dimensions of variation among cognitive skills

A large number of the tasks which we perform can be described as cognitive skills. Four of these are discussed in some detail in later chapters: reading, problem-solving, computing, and also motor skills, many of which, Colley argues, have a substantial cognitive component. A mutually exclusive typology of cognitive skills is not possible, since cognitive tasks vary on a number of different dimensions all of which are relevant to the way in which they are learned. We outline some of these below.

(1) Simple-Complex: a simple task, such as a choice reaction time, requires a decision defined by procedural steps. A complex task, such as air traffic control, requires the integration of large amounts of information and a complex set of underlying rules to guide performance. Schneider (1985) points out that practice on simple tasks, such as remembering a phone number, makes perfect, but this is not necessarily true of more complex tasks. It is often the case that learners practise for many hours and do not improve their performance because they have failed to structure the task appropriately. Although the mechanisms underlying simple and complex tasks may be found to be similar, prescriptions for training must take into account the nature of the skill being learned.

(2) Divergent-Convergent: this is the traditional distinction between cognitive tasks which apply well-defined rules to find a single acceptable solution, such as applying a statistical test, and those which result in a novel product within a given domain. This product may be also have to fulfil criteria of being acceptable aesthetically, such as in writing a novel. Aesthetic evaluations are rule-based but use less well-defined and more individualized criteria than more formalized judgements of correctness. In order to produce a solution to meet these evaluations a considerable amount of domain-related knowledge is necessary, and learning this is a significant part of the acquisition process. Many complex tasks have both divergent and convergent elements, for

example, an architect must apply principles of basic engineering and building technology in producing a novel design for a building, which also has aesthetic appeal.

(3) Algorithmic-Heuristic: in algorithmic skills, the performer uses a set sequence of steps to arrive at a solution. An example would be where a car mechanic follows instructions in a manual to dismantle the air filter from an engine. In heuristic skills the performer works from knowledge of underlying principles to produce a solution, as in playing chess. Different kinds of information need to be presented during acquisition to reflect these two strategies.

(4) Inductive-Deductive: in deductive skills, such as solving a crossword clue, the performer works forward from evidence to solution. In inductive skills, inferences are made from particular instances to similar situations so that the solution is highly likely but not necessarily true. An example would be of a consultant physician attempting to make a diagnosis, based upon experience, of a patient presenting with a set of atypical symptoms.

(5) Open-Closed: this dimension was first proposed by Poulton (1957) to distinguish between performance in a predictable (closed) environment and performance in an unpredictable (open) environment. Writing takes place within a closed environment, whereas the decisions made while driving take place in a more open environment. Gilhooly and Green discuss adversarial and non-adversarial problem-solving in their chapter and this distinction relates closely to this dimension. In adversarial tasks, the presence of an opponent (such as in playing GO or chess) increases the unpredictability of the performing environment. The learner must acquire rules about performing the task itself and, in addition, must take into account the way in which the environment is likely to vary in establishing additional rules about the appropriateness of certain actions.

(6) Universal-Specialized: some cognitive skills, such as reading, are acquired by almost everyone, and to a high level of competence. Others, such as computer programming, are learned by relatively few, and even fewer still acquire a high level of competence. This may be partly a function of the universality of teaching and large amounts of practice, but it may also be the case that certain tasks are easier to acquire because the basic abilities required are possessed by the majority of individuals.

Acquisition and performance

The dimensions outlined above illustrate the importance of considering the nature of a task in making prescriptions about the way in which acquisition should proceed. This does not mean that a generalized theory of performance can not be established, simply that the way in which a learner structures a task, the amount of domain-related information available and the way in which the procedural rules are presented to the learner are of significance for training individual skills.

A generalized theory of acquisition must take into account the way in which performance changes during acquisition. One important change is in terms of the apparent use of attentional resources. Early in acquisition, only a small amount of the available information can be attended to, while later on, the performer can accomplish the task easily and apparently has capacity to spare. With sufficient practice, under certain circumstances, two complex tasks can be performed simultaneously. Barber's chapter discusses the implications of this dual task performance for attentional theory. Some investigators (e.g. Schneider and Fisk, 1983; Schneider and Shiffrin, 1977; Shiffrin and Dumais, 1981) have proposed a two-process theory of attention to account for these changes. Early in acquisition, processing is *controlled*, that is, it uses general processing capacity, it is also slow, effortful, generally serial, under intentional control and involves awareness. Later in practice, apparently effortless performance results from *automatic* processing, which is fast, parallel, obligatory, does not involve awareness and has low demands on processing capacity. These two-process theories have been criticized for the lack of internal consistency of the definitions they give (e.g. Cheng, 1985; Phillips and Hughes, 1988). The notion of automatic processing has been described as circular (e.g. Allport, 1980); its presence is inferred from the type of performance that it is invoked to explain. It is probable that several mechanisms underlie the changes in attentional deployment that the controlled-automatic dichotomy describes (Colley and Beech, 1988) , but no current theoretical framework deals with these in a completely satisfactory way.

Speed of performance across a wide range of tasks (motor and cognitive) changes in a characteristic way with practice: large increases in speed occur initially, then performance stabilizes and increases only slightly in speed over a long period. This relationship can be described by a power law: the logarithm of the time to complete a response is a linear function of the logarithm of the number of trials (Newell and Rosenbloom, 1981, discuss this law in some detail). Neves and Anderson (1981) have demonstrated that the law describes the learning of solutions to geometry proofs and have interpreted it within the framework offered by ACT* (Anderson, 1982; see below).

So far we have considered acquisition rather than performance but, of course, the two are intimately related. Performance of a skill varies as a function of a number of factors. Perhaps the most important of these is the amount of practice that a learner has had, but other factors interact with this.

The most notable of these are developmental stage (see the chapter by Sincoff and Sternberg) and individual/environmental factors such as circadian rhythms (see Smith's chapter) and the presence or absence of various stressors (see the chapter by Hartley, Morrison and Arnold). Understanding the way in which these factors interact provides clues to the nature of underlying mechanisms as well as being of considerable practical importance.

RECENT APPROACHES TO SKILL ACQUISITION

In this section we shall briefly outline two areas of theoretical development which seem to us to have important implications for understanding cognitive skill acquisition and performance. Two productions systems theories will be presented. Production systems consist of rules for executing procedures which have an 'if . . . then . . .' form, i.e. if a set of conditions is satisfied, then an operation or operations will be executed. Anderson's ACT* theory, discussed first, has an advantage over many other similar theories, of a clear focus on the mechanisms of learning. Its influence within the literature is evident in several of the chapters in this volume. Hunt and Lansman's Production Activation Model focuses on the distinction between controlled and automatic processing, so has the potential to explain the differences in the way that attentional resources are deployed at different stages of acquisition.

The second area which will be discussed is connectionism, which, although it does not make explicit prescriptions concerning skill acquisition, nevertheless provides a clear theoretical basis for understanding skill.

Production systems approaches

Anderson's ACT theory*

Anderson (1982) has provided a framework for understanding observations made previously by Fitts (1964) on the development of skill. Fitts outlined three main stages: the cognitive stage, in which the learner makes an initial approximation to the skill, based upon background knowledge, observation or instruction; the associative stage, in which performance is refined through the elimination of errors; and the autonomous stage, in which skilled performance is well-established but still continues to improve, albeit very gradually.

Anderson bases his framework on his ACT production system. Three types of memory are distinguished in ACT. Declarative memory contains factual information in a propositional network. Procedural memory contains the procedural steps required to accomplish tasks in *productions*, which are production rules. Working memory is a blackboard for the transfer of information between declarative and procedural memory, and for the intake and rehearsal of information from the environment. Anderson's theory has two main stages. The *declarative stage* is similar to Fitts' cognitive stage, while the *procedural stage* is equivalent to Fitts' autonomous stage. Anderson regards Fitts' associative stage as a transition between assembling facts about skill and enacting and refining procedures. He calls this process of converting facts into procedures *knowledge compilation*.

In the declarative stage, knowledge about how to perform a skill is assembled from declarative memory, and from instruction or guidance, into working memory. General problem-solving procedures then turn this declarative knowledge into productions. As Anderson (1982) points out, instruction rarely specifies a procedure for the performance of a skill. The learner must establish a procedure using existing strategies, or *weak problem-solving methods*. These are very general strategies such as, for example, the use of analogy with a similar problem or working backward from a solution. Analogy is a particularly widely used strategy. Anderson illustrates the way in which students learning geometry use worked examples as a basis for solving unfamiliar problems. Rumelhart and Norman (1981) discuss the way in which procedures incorporating task knowledge (schemata in their theoretical framework) are created or extended by analogy. They also discuss the features of a good analogy from a pedagogic perspective: it should be from an area familiar to the learner and within which he or she can reason well, its domain should be similar to that of the new task, and the same operations should be appropriate or inappropriate in both domains.

Knowledge compilation has two subprocesses of *composition* and *proceduralization*. Composition collapses successive productions into a single production which has the same effect. Proceduralization removes clauses in the condition of a production that require matching from long-term memory via working memory. Compilation is a gradual process which allows for errors in procedural information to be corrected over practice. In the procedural stage of acquisition productions are *tuned*, that is, made more appropriate and efficient for the task in hand. Subprocesses of tuning are first, *generalization*, in which common aspects of specific productions are used to create a more widely applicable production which can then be used in novel situations; second, *discrimination*, which restricts the use of productions to instances where they are successful; third, *strengthening*, where productions are strengthened with repeated application so that the time taken to apply them diminishes. Strengthening produces speedup in the performance of simple tasks such as choice reaction times. Speedup in complex tasks results from strengthening and algorithmic improvement, which is the reduction in the number of productions required through composition, generalization and discrimination.

Anderson's theory exploits the advantages of having both declarative and procedural representations (Neves and Anderson, 1981). Having a declarative representation allows for the changing of procedures to suit prevailing circumstances. Procedural representation provides methods of accomplishing tasks with slightly different requirements by allowing different variables to be entered into productions, but knowledge represented in this way cannot be accessed for inspection.

Hunt and Lansman's Production Activation Model

Hunt and Lansman (1986) produced a production systems model which does not separate declarative and procedural information but which attempts to integrate findings on changes in the use of attentional resources over skill acquisition with those on problem-solving. Hunt and Lansman propose that productions can be triggered either by spreading activation between them or by matching their condition with information in working memory, which acts as a blackboard for the transfer of information between productions or its acquisition from the environment. This allows a distinction to be drawn between automatic and controlled processing in a similar way to that of the two-process theories of attention mentioned earlier. Controlled processing involves the use of working memory. An initial match is made between information in the environment and the conditions of a production rule in long-term memory. When a match is found, the condition is transferred to working memory, then the production is enacted. Automatic processing takes place via the spread of activation between productions. Hunt and Lansman have produced successful computer simulations of various tasks including choice reaction times, divided attention tasks and the Stroop test.

Hunt and Lansman point out that the empirical consequences of the difference between their model and ACT* in terms of the lack of separation of declarative and procedural memory in their Production Activation Model are unclear. The only means of testing such models and distinguishing between them is in terms of the success of computer simulations in replicating a wide range of robust findings in the literature, and more of such simulations are clearly required. One advantage that Hunt and Lansman's model does have over ACT* is in the sharp distinction it draws between controlled and automatic processing. ACT* is based very firmly in the problem-solving domain and has not given full consideration to attentional phenomena, although in ACT* the features of automatization are seen as being solved by proceduralization (Anderson, 1983). Given the earlier discussion of two-process theories of attention, it seems reasonable not to put too much emphasis on explaining a distinction between automatic and controlled processing which is in reality, not absolute. One advantage the ACT* has over the Production Activation Model, however, lies in its focus on skill acquisition. It is not clear how the Production Activation Model deals with skill acquisition, and particularly how the transition from controlled to automatic processing occurs for well-learned tasks.

Connectionist approaches

The study of cognition gathered momentum from the 1960s using the analogy between the programming operations in the computer and cognitive processing. The processing of information was considered to operate in discrete stages, usually in a serial sequence (e.g. Sternberg, 1966). More recently, a new framework has emerged which is more concerned with modelling cognitive architecture, rather than modelling the programming taking place within the hard wiring of the conventional computer. Previously, some cognitive psychologists had argued that their main focus was on the nature of the programs operating within cognition, while the nature of the 'wiring' was more the province of the neuropsychologist. In the new connectionism the modeller develops *networks* of *connections* between *units* which can give the appearance of a model which is actually simulating neuronal networks, and indeed some modellers have explicitly set out to do this (e.g. Gluck and Thompson, 1987). However, this is not the primary intention of most modellers, at least for the present (McClelland, 1988). In this section we shall describe connectionism and how it can be relevant to explaining how skills develop. We shall then briefly consider the connectionist position in relation to the type of symptoms and deterioration sustained by brain damage. Issues raised by these considerations are whether skills are organized within modules and whether all skill operations can be accounted for in terms of the connectionist framework.

Connectionism

McClelland, Rumelhart and Hinton (1986) and Phillips (1988) have briefly outlined the various strands of development leading to the present connectionist approach in cognitive psychology. In the early 1980s and slightly earlier the first cognitive connectionist models began to appear independently of each other. For example, one of us proposed a network model to account for the phenomenon of visual image scanning, proposing that the increase in reaction time as a function of distance scanned across an image could be accounted for by a model involving triggering signals, domino-fashion, across an array of elements representing an image (Beech, 1979a, b); Feldman (1981) proposed what he called a 'connectionist' model of visual memory; an edited book by Hinton and Anderson (1981) highlighted the importance of neural net models for cognitive psychologists; McClelland and Rumelhart

(1981) produced a connectionist model of word recognition; and by 1986 two influential edited volumes had appeared by McClelland and Rumelhart (McClelland and Rumelhart, 1986; Rumelhart and McClelland, 1986a).

A crucial property of the network within the connectionist model is the nature of the individual units. These are 'simple processing devices which take on activation values based on a weighted sum of the inputs from the environment and from other units' (McClelland, 1988). In other words, these units are capable of, for instance, computing the relative importance of a set of inputs. The role of these units varies considerably. A unit might be the representation of a single word in the lexicon, and there could also be units to represent letters and features of letters (e.g. McClelland and Rumelhart, 1981). On the other hand, a conceptualization within cognition might entail a particular pattern of activation over a large network of units (e.g. Beech, 1979a, b).

To illustrate how a weighting system might work, consider the model of Paap, Newsome, McDonald and Schvaneveldt (1982). In this model of word recognition a confusion matrix was used which had been obtained by giving subjects brief visual presentations of individual letters and noting their errors. Paap *et al.* thus derived a set of probabilities for visually presented letters when confused with all possible alternative letters. A lexicon was also stored with the visual confusion matrix in the program to represent the visual lexicon of the reader. In the simulation, single words were 'presented' to the program. All the words in the lexicon were activated on this presentation in accordance with their corresponding probability values in the confusion matrix. Then the geometrical mean was taken (by multiplying the probabilities together and dividing by the number of letters) for each word in the lexicon. This had the effect of attenuating any word which contained letters with a value close to zero. In other words, if a word in the lexicon contained a letter visually very dissimilar (e.g. *M* vs *O*), this would dramatically reduce the geometrical mean for that particular word entry. However, a word visually similar to the presented word (e.g. *PORE* vs *PORK*) would have a reasonably high geometrical mean. It can be seen that this model involves a detailed simultaneous activation of all the words in the lexicon with each unit representing a word being involved in computational activity. Such models are often referred to as modelling parallel distributed processes or simply PDP models. The actual simulation of a PDP model on a conventional computer involves computing activation levels for each word in a serial manner, which is a time-consuming process. Until adequate parallel computer architectures are developed (and accompanying programming languages), these simulations can not take place in real time. Nevertheless, they do provide new ways of looking at experimental phenomena which had not hitherto been contemplated.

Connectionist accounts of skill acquisition

Two examples of applications of connectionism to acquiring a skill will now be considered. The first example is by Gluck and Bower (1988) who have proposed a connectionist account of students learning to make a diagnosis based on descriptions of medical symptoms of patients, using the Rescorla–Wagner model of associative learning (Rescorla and Wagner, 1972; Wagner and Rescorla, 1972). The Gluck and Bower model involves a network with the input units or nodes representing the medical symptoms (e.g. bleeding gums).

Before going further it should be noted that connectionist accounts frequently use linear algebra to describe their models. In many cases this enables components of a vector to be represented in n-dimensional space, where n is the number of components. In the case of the basic Gluck–Bower model, a vector represents the patient, p, who has four binary components corresponding to the medical symptoms. Each patient has either of two diseases. The model operates by being given a series of symptoms for each patient and informing it which disease is operating. The model multiplies the activation of each symptom x_{pj}, with its corresponding weight w_j and then sums these values across all symptoms, in this case, four. This results in the output for the patient, o_p, which is the degree to which that disease is preferred over disease 2, as shown in equation 1:

$$o_p = \sum_{k=1}^{n} x_{pk} w_k \tag{1}$$

Then a change in the value of the weight is calculated which is equivalent to learning about the relationship of the symptoms to the disease. This uses the delta rule, or the Windrow–Hoff rule (Sutton and Barto, 1981), in which the extent of what is learned is proportional to the *difference* (this is the delta part) between the level of actual activation and the target level of activation. This is expressed in equation 2, in which λ_p is the desired disease output, o_p is the calculated output derived from equation 1 and β is the learning rate:

$$\Delta w_j = \beta(\lambda_p - o_p)x_{pj} \tag{2}$$

The result of this function is to produce a learning rate which is negatively accelerated reaching asymptote at a rate determined by the magnitude of β. The constant β must be small, otherwise there will be exaggerated oscillations in the weight changes.

One important aspect common to both the Rescorla–Wagner model and the simulation model of Paap *et al.* (in which the geometrical mean was used) is the use of products to provide a *gating* mechanism. When there is a multiplicative connection this means that one unit is capable of gating the other or several other units, because as mentioned before, if one has a value of zero, this means that the other members are effectively blocked as well. If some weights have a value of unity, they are effectively neutralized and if they are positive but below one, they reduce output according to their value. The state of being positive or negative will also be important; for instance, one negative value will have an inhibitory effect on overall output.

There are several other cases where PDP models have been used to give an account of skill development. One further example to be described here is that of the development of skilled typing. The connectionist model of Rumelhart and Norman (1982) proposed that on reading a word within a sequence, a unit corresponding to that word is activated which in turn activates individual units representing component letters of the word. The activation of the first letter unit to be typed within the word inhibits the rest of the units, the second unit inhibits the remainder, and so on. The end result is the activation of the units in descending order in relation to their serial position. Hand position for each letter-press is relatively unchanged if the letter is on the home row, but can move if the top or bottom rows need to be reached. But the extent of this movement is modified by the activation levels of other units which need to be typed, with the succeeding letter unit to the current letter unit exerting the greatest activation in the context of the overall activation of the remaining letters. One important aspect to note about this model, which has been quite good at predicting inter-keystroke time, is that it proposes that typing is not a serial process in which each letter unit is activated in succession. The role of a central executive function, in the sense of a serial conscious process initiating behaviour, is reduced in this model because important processing functions, involving the computation of weighting levels, are distributed within the operating system. The chapter on motor skills discusses this issue further.

Connectionism and brain damage

The impact of neurological damage on behaviour is of considerable interest for those interested in skill development. In many cases damage can result in the impairment and sometimes the complete disappearance of skills which had previously been acquired. Some connectionists make the strong claim that their PDP approach can provide an account of these effects in a way that other theories are unable to do. Given that brain damage appears to eliminate whole categories of subskills in certain cases, this suggests that whole specialized modules have been eliminated. However, modularity is not a feature of connectionism. This issue is explored briefly and will also be germaine to the chapter on processing resources by Paul Barber and to Beech's chapter which proposes that the development of component subskills is an important feature of learning to read.

Rumelhart and McClelland (1986b), advocating the connectionist position, concede that there are *regions* of specialized brain activity, particularly for lower levels of processing, but generally speaking patients usually experience *graceful degradation* in performance corresponding to deterioration in increasing numbers of neurological units. This is because in neurological structures there is a considerable amount of redundancy, in the same way as in connectionist models. Alzheimer's Disease is a particularly good example of this phenomenon. This position contrasts with the serial cognitive models of the past in which the elimination of one stage of processing would result in a hypothesized catastrophic deterioration in performance, analogous to the performance in the conventional computer in which one error in the program can make the remainder of the operations meaningless.

Hinton, McClelland and Rumelhart (1986) described a connectionist model of learning to read which produced semantic errors (e.g. responding 'apricot' when presented with *peach*) similar to the kinds of errors produced by deep dyslexic patients. This is an interesting result because the acquired dyslexias usually produce quite distinctive symptoms often suggestive of a loss of distinctive modules or specialized functions (e.g. Marshall, 1987). But the connectionist view of Rumelhart and McClelland does not account for different specialities and yet their model is operating in a way approximating that of a deep dyslexic (at least, in one aspect). On this point Phillips (1988), perhaps reading more into their position than is intended, imputes that 'No attempt is made to indicate possible roles for the modules. . . . It grossly undervalues the evidence obtained by other approaches, such as neuropsychology' (p. 396). A middle course is that subsystems could each operate in a connectionist manner. The model described by Hinton, McClelland and Rumelhart produced the symptoms of a deep rather than a surface dyslexic because one layer of the system was concerned with semantic representation. If there had been a grapheme-phoneme conversion layer of units no doubt errors would have been generated more analogous to one of the symptoms of the surface dyslexic. Phillips suggests that modules with different types of specialities will need different kinds of properties. This is a view reinforced by Fodor and Pylyshyn (1988) who are also critical of the connectionist acounts, but in a more radical way. It is not surprising that there are critics, especially as the connectionist view has been expressed in somewhat extreme terms. Nevertheless, as far as skill acquisition is concerned, the approach promises some interesting advances.

CONCLUSIONS

In this introductory chapter we have outlined some general issues concerning cognitive skill acquisition and performance, many of which will be elaborated upon in more detail in the chapters that follow. Two major conclusions can be drawn from the preceding discussion. First, that cognitive psychology is, at last, placing some emphasis on mechanisms of learning. Not only do advances in cognitive psychology suggest new ways of describing these mechanisms, but also, as Langley and Simon (1981) conclude 'Learning theory . . . [has] a central role to play in formulating parsimonious, nearly invariant laws of cognition' (p. 378), since stage of learning is one of the major determinants of performance. Second, that artificial intelligence methodology is central to advances in this area, and complements the more traditional laboratory-based experimentation as well as the more ecologically valid use of protocol analysis to understand cognitive skills in the environment in which they are performed.

REFERENCES

Allport, D. A. (1980). Attention and performance, in G. L. Claxton (ed.), *Cognitive Psychology: New Directions*. London: Routledge.

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

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University.

Beech, J. R. (1979a). Scanning visual images: implication for the representation of space, *Perception*, **8**, 621–628.

Beech, J. R. (1979b). A chronometric study of the scanning of visual representations. D. Phil. thesis submitted to the New University of Ulster.

Cheng, P. W. (1985). Restructuring versus automaticity: Alternative accounts of skill acquisition, *Psychological Review*, **92**, 414–423.

Colley, A. M. & Beech, J. R. (1988). Discussion: Changing representations of skill, in A. M. Colley & J. R. Beech (eds), *Cognition and Action in Skilled Behaviour*. Amsterdam: North-Holland.

Ericsson, K. A. & Simon, H. A. (1984). *Protocol Analysis*. Cambridge, MA: MIT/Bradford.

Feldman, J. A. (1981). A connectionist model of visual memory, in G. E. Hinton & J. A. Anderson (eds), *Parallel Models of Associative Memory*. Hillsdale, NJ: Erlbaum.

Fitts, P. M. (1964). Perceptual-motor skill learning, in A. W. Melton (ed.), *Categories of Human Learning*. New York: Academic Press.

Fitts, P. M. & Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks/Cole.

Fodor, J. A. & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis, *Cognition*, **28**, 3–71.

Gluck, M. A. & Bower, G. H. (1988). Evaluating an adaptive network model of human learning, *Journal of Memory and Learning*, **27**, 166–195.

Gluck, M. A. & Thompson, R. F. (1987). Modeling the neural substrates of associative learning and memory: A computational approach, *Psychological Review*, **94**, 176–191.

Hinton, G. E. & Anderson, J. A. (1981). *Parallel Models of Associative Memory*. Hillsdale, N.J.: Erlbaum.

Hinton, G. E., McClelland, J. L. & Rumelhart, D. E. (1986). Distributed representations, in D. E. Rumelhart & J. L. McClelland (eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition (vol. 1 Foundations)*. Cambridge, MA: MIT Press.

Hunt, E. & Lansman, M. (1986). Unified model of attention and problem solving, *Psychological Review*, **93**, 446–461.

Langley, P.E. & Simon, H. A. (1981). The central role of learning in cognition, in J. R. Anderson (ed.), *Cognitive Skills and their Acquisition*. Hillsdale, NJ: Erlbaum.

Marshall, J. (1987). The cultural and biological context of written languages, in J. R. Beech & A. M. Colley (eds), *Cognitive Approaches to Reading*. Chichester: Wiley.

McClelland, J. L. (1988). Connectionist models and psychological evidence, *Journal of Memory and Language*, **27**, 107–123.

McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings, *Psychological Review*, **88**, 375–407.

McClelland, J. L. & Rumelhart, D. E. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition (vol. 2 Psychological and Biological Models)*. Cambridge, MA: MIT Press.

McClelland, J. L., Rumelhart, D. E. & Hinton, G. E. (1986). The appeal of parallel distributed processing, in D. E. Rumelhart & J. L. McClelland (eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (vol. I). Cambridge, MA: MIT Press.

Neves, D. M. & Anderson, J. R. (1981). Knowledge compilation: mechanisms for the automatization of cognitive skills, in J. R. Anderson (ed.), *Cognitive Skills and their Acquisition*. Hillsdale, NJ: Erlbaum.

Newell, A. & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice, in J. R. Anderson (ed.), *Cognitive Skills and their Acquisition*. Hillsdale, NJ: Erlbaum.

Paap, K.R., Newsome, S. L., McDonald, J. E. & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word-superiority effect, *Psychological Review*, **89**, 573–594.

Phillips, J. & Hughes, B. (1988). Internal consistency of the concept of automaticity, in A. M. Colley & J. R. Beech (eds), *Cognition and Action in Skilled Behaviour*. Amsterdam: North-Holland.

Phillips, W. A. (1988). Brainy minds, *Quarterly Journal of Experimental Psychology*, **40A**, 389–405.

Poulton, E. C. (1957). On prediction in skilled movements, *Psychological Bulletin*, **54**, 467–478.

Rescorla, R.A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement following prior inhibitory conditioning, in A. H. Black & W. F. Prokasy (eds), *Classical Conditioning: II. Current Research and Theory*. New York: Appleton–Century–Crofts.

Rumelhart, D.E. & McClelland, J. L. (1986a). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. (vol. 1 Foundations)*. Cambridge, MA: MIT Press.

Rumelhart, D. E. & McClelland, J. L. (1986b). PDP models and general issues in cognitive science, in D. E. Rumelhart & J. L. McClelland (eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition (vol. 1 Foundations)*. Cambridge, MA: MIT Press.

Rumelhart, D. E. & Norman, D. A. (1981). Analogical processes in learning, in J. R. Anderson (ed.), *Cognitive Skills and their Acquisition*. Hillsdale, NJ: Erlbaum.

Rumelhart, D. E. & Norman, D. A. (1982). Simulating a typist: A study of skilled cognitive-motor performance, *Cognitive Science*, **6**, 1–36.

Schneider, W. (1985). Training high-performance skills: fallacies and guidelines, *Human Factors*, **27**, 285–300.

Schneider, W. & Fisk, A. D. (1983). Attention theory and mechanisms for skilled performance, in R. A. Magill (ed.), *Memory and Control of Action*. Amsterdam: North-Holland.

Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic information processing: I. Detection, search, and attention, *Psychological Review*, **84**, 1–66.

Shiffrin, R. M. & Dumais, S. T. (1981). The development of automatism, in J. R. Anderson (ed.), *Cognitive Skills and their Acquisition*. Hillsdale, NJ: Erlbaum.

Sternberg, S. (1966). High-speed scanning in human memory *Science*, **153**, 652–654.

Sutton, R. S. & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectations and prediction, *Psychological Review*, **88**, 135–170.

Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.

Wagner, A. R. & Rescorla, R.A. (1972). Inhibition in Pavlovian conditioning: applications of a theory, in R. A. Boakes & S. Halliday (eds). *Inhibition and Learning*. New York: Academic Press.

Welford, A. T. (1976). *Skilled Performance: Perceptual and Motor Skills*. Glenview, Ill: Scott Foresman.