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

Library of Congress Cataloging in Publication Data

Card, Stuart K.
The psychology of human-computer interaction.

Bibliography: p.
Includes index.
1. Interactive computer systems—Psychological aspects.
I. Moran, Thomas P. II. Newell, Allen. III. Title.
QA76.9I58C37 1983 001.64'01'9 82-21045
ISBN 0-89859-243-7

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

The Psychology of Human-Computer Interaction

Stuart K. Card
Thomas P. Moran
Xerox Palo Alto Research Center

Allen Newell
Carnegie-Mellon University

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1983 Hillsdale, New Jersey London

2. The Human Information-Processor

2.1. THE MODEL HUMAN PROCESSOR
- The Perceptual System
- The Motor System
- The Cognitive System

2.2. HUMAN PERFORMANCE
- Perception
- Motor Skill
- Simple Decisions
- Learning and Retrieval
- Complex Information-Processing

2.3. CAVEATS AND COMPLEXITIES

Our purpose in this chapter is to convey a version of the existing psychological science base in a form suitable for analyzing human-computer interaction. To be practical to use and easy to grasp, the description must necessarily be an oversimplification of the complex and untidy state of present knowledge. Many current results are robust, but second-order phenomena are almost always known that reveal an underlying complexity; and alternative explanations usually exist for specific effects. An uncontroversial presentation in these circumstances would consist largely of purely experimental results. Such an approach would not only abandon the possibility of calculating parameters of human performance from the analysis of a task, but would also fail in the primary purpose of giving the reader knowledge in a form relatively easy to assimilate.

Our tack, therefore, is to organize the discussion around a specific, simple model. Though limited, this model allows us to give, insofar as possible, an integrated description of psychological knowledge about human performance as it is relevant to human-computer interaction.

2.1. THE MODEL HUMAN PROCESSOR

A computer engineer describing an information-processing system at the systems level (as opposed, for instance, to the component level) would talk in terms of memories and processors, their parameters and interconnections.[1] By suppressing detail, such a description would help him to envision the system as a whole and to make approximate predictions of gross system behavior.

The human mind is also an information-processing system, and a description in the same spirit can be given for it. The description is approximate when applied to the human, intended to help us remember facts and predict user-computer interaction rather than intended as a statement of what is really in the head. But such a description is useful for making approximate predictions of gross human behavior. We therefore organize our description of the psychological science base around a model of this sort. To distinguish the simplified account of the present model from the fuller psychological theory we would present in other contexts, we call this model the *Model Human Processor.*

The Model Human Processor (see Figures 2.1 and 2.2) can be described by (1) a set of memories and processors together with (2) a set of principles, hereafter called the "principles of operation." Of the two parts, it is easiest to describe the memories and processors first, leaving the description of the principles of operation to arise in context.

The Model Human Processor can be divided into three interacting subsystems: (1) the *perceptual system*, (2) the *motor system*, and (3) the *cognitive system*, each with its own memories and processors. The perceptual system consists of sensors and associated buffer memories, the most important buffer memories being a Visual Image Store and an Auditory Image Store to hold the output of the sensory system while it is being symbolically coded. The cognitive system receives symbolically coded information from the sensory image stores in its Working Memory and uses previously stored information in Long-Term Memory to make decisions about how to respond. The motor system carries out the response. As an approximation, the information processing of the human will be described as if there were a separate processor for each subsystem: a Perceptual Processor, a Cognitive Processor, and a Motor Processor. For some tasks (pressing a key in response to a light) the human must behave as a serial processor. For other tasks (typing, reading, simultaneous translation) integrated, parallel operation of the three subsystems is possible, in the manner of three pipelined processors: information flows continuously from input to output with a characteristically short time lag showing that all three processors are working simultaneously.

[1] For a survey of computing systems in these terms see Siewiorek, Bell, and Newell (1981).

The memories and processors are described by a few parameters. The most important parameters of a memory are

μ, the storage capacity in items,
δ, the decay time of an item, and
κ, the main code type (physical, acoustic, visual, semantic).

The most important parameter of a processor is

τ, the cycle time.

Whereas computer memories are usually also characterized by their access time, there is no separate parameter for access time in this model since it is included in the processor cycle time.

We now consider each of the subsystems in more detail.

The Perceptual System

The perceptual system carries sensations of the physical world detected by the body's sensory systems into internal representations of the mind by means of integrated sensory systems. An excellent example of the integration of a sensory system is provided by the visual system: The retina is sensitive to light and records its intensity, wave length, and spatial distribution. Although the eye takes in the visual scene over a wide angle, not quite a full half-hemisphere, detail is obtained only over a narrow region (about 2 degrees across), called the *fovea*. The remainder of the retina provides peripheral vision for orientation. The eye is in continual movement in a sequence of saccades, each taking about 30 msec to jump to the new point of regard[2] and dwelling there 60~700 msec for a total duration of

[2] Russo (1978).

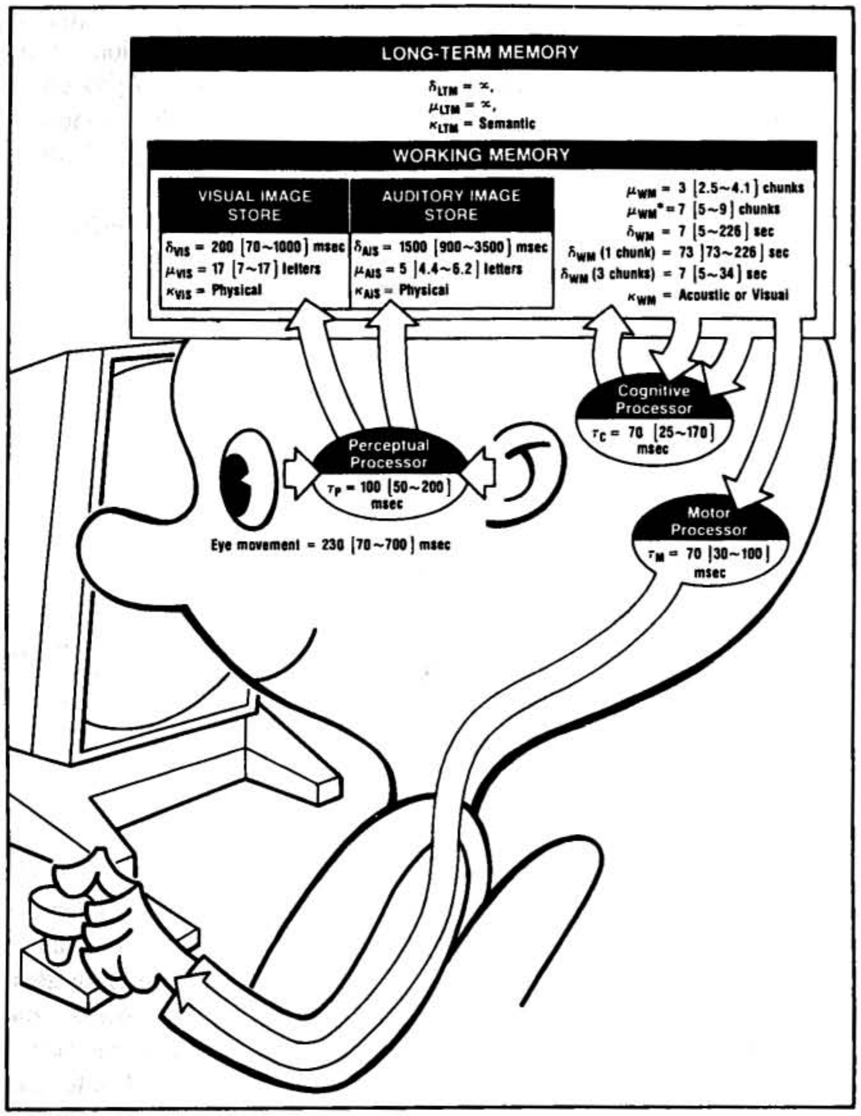

Figure 2.1. The Model Human Processor—memories and processors.

Sensory information flows into Working Memory through the Perceptual Processor. Working Memory consists of activated chunks in Long-Term Memory. The basic principle of operation of the Model Human Processor is the *Recognize-Act Cycle of the Cognitive Processor* (P0 in Figure 2.2). The Motor Processor is set in motion through activation of chunks in Working Memory.

P0. ***Recognize-Act Cycle of the Cognitive Processor.*** On each cycle of the Cognitive Processor, the contents of Working Memory initiate actions associatively linked to them in Long-Term Memory; these actions in turn modify the contents of Working Memory.

P1. ***Variable Perceptual Processor Rate Principle.*** The Perceptual Processor cycle time τ_P varies inversely with stimulus intensity.

P2. ***Encoding Specificity Principle.*** Specific encoding operations performed on what is perceived determine what is stored, and what is stored determines what retrieval cues are effective in providing access to what is stored.

P3. ***Discrimination Principle.*** The difficulty of memory retrieval is determined by the candidates that exist in the memory, relative to the retrieval clues.

P4. ***Variable Cognitive Processor Rate Principle.*** The Cognitive Processor cycle time τ_c is shorter when greater effort is induced by increased task demands or information loads; it also diminishes with practice.

P5. ***Fitts's Law.*** The time T_{pos} to move the hand to a target of size S which lies a distance D away is given by:

$$T_{pos} = I_M \log_2 (D/S + .5)\,, \tag{2.3}$$

where I_M = 100 [70~120] msec/bit.

P6. ***Power Law of Practice.*** The time T_n to perform a task on the nth trial follows a power law:

$$T_n = T_1 n^{-\alpha}\,, \tag{2.4}$$

where α = .4 [.2~.6].

P7. ***Uncertainty Principle.*** Decision time T increases with uncertainty about the judgement or decision to be made:

$$T = I_C H\,,$$

where H is the information-theoretic entropy of the decision and I_C = 150 [0~157] msec/bit. For n equally probable alternatives (called Hick's Law),

$$H = \log_2 (n + 1)\,. \tag{2.8}$$

For n alternatives with different probabilities, p_i, of occurence,

$$H = \Sigma_i\, p_i \log_2 (1/p_i + 1)\,. \tag{2.9}$$

P8. ***Rationality Principle.*** A person acts so as to attain his goals through rational action, given the structure of the task and his inputs of information and bounded by limitations on his knowledge and processing ability:

Goals + Task + Operators + Inputs
+ Knowledge + Process-limits → Behavior

P9. ***Problem Space Principle.*** The rational activity in which people engage to solve a problem can be described in terms of (1) a set of states of knowledge, (2) operators for changing one state into another, (3) constraints on applying operators, and (4) control knowledge for deciding which operator to apply next.

Figure 2.2. The Model Human Processor—principles of operation.

$$\text{Eye-movement} = 230\ [70 \sim 700]\ \text{msec}\ .$$ [3]

(In this expression, the number 230 msec represents a typical value and the numbers in brackets indicate that values may range from 70 msec to 700 msec depending on conditions of measurement, task variables, or subject variables.) Whenever the target is more than about 30 degrees away from the fovea, head movements occur to reduce the angular distance. These four parts—central vision, peripheral vision, eye movements, and head movements—operate as an integrated system, largely automatically, to provide a continual representation of the visual scene of interest to the perceiver.

PERCEPTUAL MEMORIES

Very shortly after the onset of a visual stimulus, a representation of the stimulus appears in the *Visual Image Store* of the Model Human Processor. For an auditory stimulus, there is a corresponding *Auditory Image Store*. These sensory memories hold information coded *physically*, that is, as an unidentified, non-symbolic analogue to the external stimulus. This code is affected by physical properties of the stimulus, such as intensity. For our purposes we need not enter into the details of the physical codes for the two stores but can instead just write:

$$\kappa_{VIS} = \text{physical}\ ,$$
$$\kappa_{AIS} = \text{physical}\ .$$

For example, the Visual Image Store representation of the number 2 contains features of curvature and length (or equivalent spatial frequency patterns) as opposed to the recognized digit.

The perceptual memories are intimately related to the cognitive Working Memory as Figure 2.1 depicts schematically. Shortly after a physical representation of a stimulus appears in one of the perceptual memories, a recognized, symbolic, acoustically-coded (or visually-coded)

[3] Actual saccadic eye-movement times (travel + fixation time) can vary quite considerably depending on the task and the skill of the observer. Russo (1978, Table 2, p. 94) lists 70 msec as the minimum time and 230 msec as a typical time. The largest time given by Busswell (1922, p. 31) for eye-movements in reading is 660 msec (for first-grade children), which we round to 700 msec.

representation of at least part of the perceptual memory contents occurs in Working Memory. If the contents of perceptual memory are complex or numerous (for example, an array of letters) and if the stimulus is presented only fleetingly, the perceptual memory trace fades, and Working Memory is filled to capacity before all the items in the perceptual memory can be transferred to representations in Working Memory (for letters the coding goes at about 10 msec/letter). However, the Cognitive Processor can specify which portion of the perceptual memory is to be so encoded. This specification can only be by physical dimensions, since this is the only information encoded: after being shown a colored list of numbers and letters, a person can select (without first identifying what number or letter it is) the top half of the Visual Image Store or the green items, but not the even digits or the digits rather than the letters.

Figure 2.3 shows the decay of the Visual Image Store and the Auditory Image Store over time. As an index of decay time, we use the half-life, defined as the time after which the probability of retrieval is less than 50%. While exponential decay is not necessarily implied by the use of the half-life, Figure 2.3 shows that it is often a good approximation to the observed curves. The Visual Image Store has a half-life of about

$$\delta_{VIS} = 200\ [90 \sim 1000]\ \text{msec}\ ,$$ [4]

but the Auditory Image Store decays more slowly,

[4] A least-squares fit to data estimated from figures appearing in Sperling (1960) and Averbach and Coriell (1961) yields the following facts. The half-life of the letters in excess of the memory span that subjects could report in the partial report condition of Sperling's (1960) experiment was 621 msec (9-letter stimulus) and 215 msec (12-letter stimulus). Averbach and Coriell's (1961) experiment gives a half-life of 92 msec (16-letter stimulus). The typical value for δ_{VIS} has been set at 200 msec, representing the middle of these. The lower and upper bounds for δ_{VIS} are set at rounded-off values reflecting the fastest subject in the condition with the shortest half-life and the slowest subject in the condition with the longest half-life. The shortest half-life in these experiments was 93 msec for Averbach and Coriell's Subject GM (16-letter condition); the longest half-life was 940 msec for Sperling's Subject ROR (9-letter condition). It is possible to have the average half-life be 92 msec, shorter than the half-life of any subject, because this average is computed by first taking the mean of each point across subjects, then computing the slope of the best least-square fitting line in semilog coordinates.

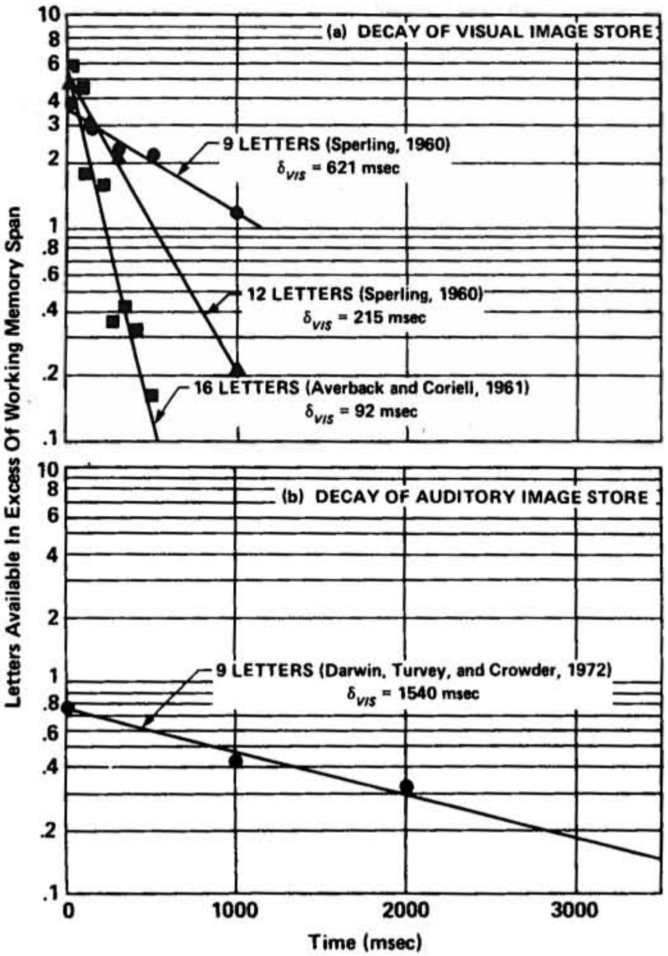

Figure 2.3. Time decay of Visual and Auditory Image Stores.
(a) Decay of the Visual Image Store. In each experiment, a matrix of letters was made observable tachistoscopically for 50 msec. In the case of the Sperling experiments, a tone sounded after the offset of the letters to indicate which row should be recalled. In the case of the Averbach and Coriell experiment, a bar appeared after the offset of the letters next to the letter to be identified. The percentage of indicated letters that could be recalled eventually asymptotes to $\mu_{WM}{}^*$. The graph plots the percentage of letters reported correctly in excess of $\mu_{WM}{}^*$ as a function of time before the indicator.
(b) Decay of the Auditory Image Store. Nine letters were played to the observers over stereo earphones arranged so that three sequences of letters appear to come from each of three directions. A light lit after the offset of the letters to indicate which sequence should be recalled. The graph plots the percentage of the relevant 3-letter sequence in excess of $\mu_{WM}{}^*$ reported correctly as a function of time before the light was lit.

$$\delta_{AIS} = 1500\ [900\sim3500]\ \text{msec}\,,^{5}$$

consistent with the fact that auditory information must be interpreted over time. The capacity of the Visual Image Store is hard to fix precisely but for rough working purposes may be taken to be about

$$\mu_{VIS} = 17\ [7\sim17]\ \text{letters}\,.^{6}$$

The capacity of the Auditory Image Store is even more difficult to fix, but would seem to be around

$$\mu_{AIS} = 5\ [4.4\sim6.2]\ \text{letters}\,.^{7}$$

PERCEPTUAL PROCESSOR

The cycle time τ_P of the Perceptual Processor is identifiable with the so-called *unit impulse response* (the time response of the visual system to

[5] The half-life of the letters in excess of the memory span that subjects could report in the partial report condition of Darwin, Turvey, and Crowder's (1972) experiment was 1540 msec, which we have rounded to δ_{AIS} = 1500 msec. The difference in decay half-life as a function of letter order in their experiment (963 msec for the third letter, 3466 msec for the first letter) has been rounded to give lower and upper bounds of 900 and 3500. Other techniques have been used to obtain values for the "decay time" of the Auditory Image Store. For example, use of a masking technique gives estimates of around 250 msec full decay (Massaro, 1970), but these experiments have been criticized by Klatzky (1980, p. 42) because they may only measure the time necessary to transmit categorical information to Working Memory. On the other end, experiments that measure the delay at which there is still some facilitation of the identification of a noisy signal (Crossman, 1958; Guttman and Julesz, 1963) give very wide full-decay estimates: from 1000 msec to 15 minutes!

[6] Sperling (1963, p. 22) estimates the capacity of the Visual Image store in terms of the number of letters available at least 17 letters and possibly more. The fewest number of letters available for any subject immediately after stimulus presentation in the 9-letter condition (Sperling, 1960) was 7.4 letters for Subject NJ.

[7] Range is from the number of letters or numbers that could be reported by Darwin, Turvey, and Crowder's (1972) subjects in an experiment in which they had to give the trio of letters coming from one of three directions (indicated by a visual cue shortly after the end of the sounds). Lowest value, 4.4 letters, is for accuracy of recalling second letter of triple when subjects had to name all items coming from a certain direction (Figure 1, p. 259). Highest number, 6.2 letters, is for recall by category when no location was required (Figure 2(B), p. 262).

a very brief pulse of light)[8] and its duration is on the order of

$$\tau_P = 100\ [50\sim200]\ \text{msec}\ .^9$$

If a stimulus impinges upon the retina at time $t = 0$, at the end of time $t = \tau_P$ the image is available in the Visual Image Store and the human claims to see it. In truth, this is an approximation, since different information in the image becomes available at different times, much as a photograph develops.[10] For example, movement information and low spatial frequency information are available sooner than other information. A person can react before the image is fully developed or can wait for a better image, according to whether speed or accuracy is the more important.

Perceptual events occurring within a single cycle are combined into a single percept if they are sufficiently similar. For example, two lights occurring at different nearby locations within 60~100 msec combine to give the impression of a single light in motion. A brief pulse of light, lasting t msec with intensity I, has the same appearance as a longer pulse of less-intense light, provided both pulses last less than 100 msec, giving rise to Bloch's Law (1885):

$$I \cdot t = k, \quad t < \tau_P\ .$$

Two brief pulses of light within a cycle combine their intensities in a more complicated way, but still give a single percept.[11] Thus there is a basic quantum of experience; and the present is not an instantaneous dividing line between past and future, but has itself duration.

Figure 2.4 shows the results of an experiment in which subjects were presented with a rapid set of clicks, from 10 to 30 clicks per second, and were asked to report how many they heard. The results show that they heard the correct number when the clicks were presented at 10 clicks/sec, but missed progressively more clicks at 15 and 30 clicks/sec. A simple

[8] See Ganz (1975).

[9] The source of the range is the review by Harter (1967), who also discusses the suggestion that the cycle time can be identified with the 77~125 msec alpha period in the brain.

[10] See Ericksen and Shultz (1978), Ganz (1975).

[11] See Ganz (1975).

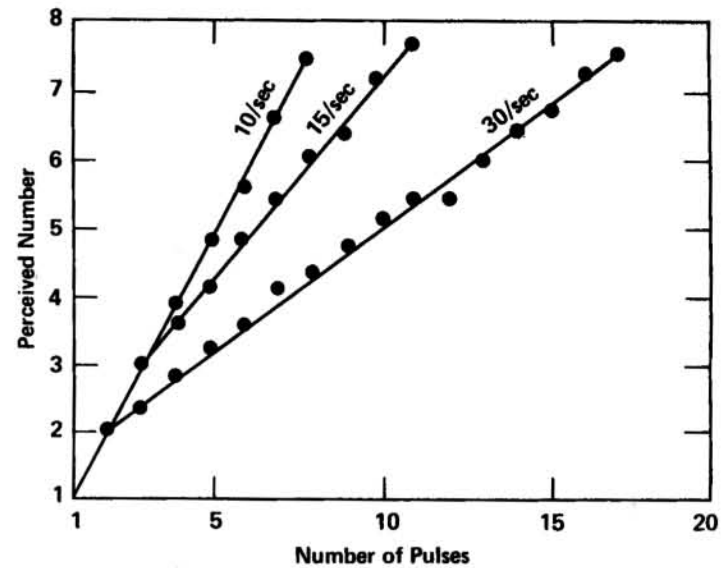

Figure 2.4. Fusion of clicks within 100 msec.
A burst of sound containing an unknown number of auditory clicks at the uniform rate of 10/sec, 15/sec, or 30/sec was presented to the subject. The graph plots the number of clicks/burst reported as a function of the number presented. After Cheatham and White (1954, Figure 1, p. 427).

analysis in terms of the Model Human Processor shows why. When the experimenter plays the clicks at 10 clicks/sec, there is one click for each $\tau_P \simeq 100$ msec interval and the subject hears each click. But when the experimenter plays the clicks at 30 clicks/sec, the three clicks in each 100 msec cycle time are fused into a single percept (perhaps sounding a little louder) and the subject hears only one click instead of three, or 10 clicks/sec. The data in Figure 2.4 show that the number of clicks/sec perceived by the subjects does in fact stay approximately constant in the 10 clicks/sec range (the measured values of the slopes are 9~11 clicks/sec) for the three rates of presentation.

As a second-order phenomenon, the processor time τ_P is not completely constant, but varies somewhat according to conditions. In particular, τ_P is shorter for more intense stimuli, a fact derivable from a more detailed examination of the human information-processor using linear systems theory, but which we simply adopt as one of the principles of operation (Figure 2.2):

> ***P1. Variable Perceptual Processor Rate Principle.*** *The Perceptual Processor cycle time* τ_P *varies inversely with stimulus intensity.*

The effect of this principle is such that τ_P can take on values within the 50~200 msec range we have given. Under very extreme conditions of intense, high-contrast stimuli or nearly invisible, low-contrast stimuli, τ_P can take on values even outside these ranges.

The Motor System

Let us now consider the motor system. Thought is finally translated into action by activating patterns of voluntary muscles. These are arranged in pairs of opposing "agonists" and "antagonists," fired one shortly after the other. For computer users, the two most important sets of effectors are the arm-hand-finger system and the head-eye system.

Movement is not continuous, but consists of a series of discrete micromovements, each requiring about

$$\tau_M = 70\,[30\sim100]\text{ msec},$$ [12]

which we identify as the cycle time of the Motor Processor. The feedback loop from action to perception is sufficiently long (200~500 msec) that rapid behavioral acts such as typing and speaking must be executed in bursts of preprogrammed motor instructions.

An instructive experiment is to have someone move a pen back and forth between two lines as quickly as possible for 5 sec (see Figure 2.5). Two paths through the processors in Figure 2.1 are clearly visible: (1) The Motor Processor can issue commands ("open loop") about once every τ_M = 70 msec; in Figure 2.5 this path leads to the 68 pen reversals made by the subject in the 5 sec interval, or τ_M = 74 msec/reversal. (2) The subject's perceptual system can perceive whether the strokes are staying within the lines (the perception process requires τ_P msec) and send this information to the cognitive system, which can then advise (the decision process requires τ_C msec) the motor system to issue a correction (the motor process requires τ_M msec). The total time, therefore, to make a correction using visual feedback ("closed loop") should be on the order of $\tau_P + \tau_C + \tau_M$ = 240 msec; in Figure 2.5, this path leads to the roughly 20 corrections about the ruled guidelines as indicated by the dotted line tracing the contours of the bottoms of the strokes, or (5 sec)/(20 movements) = 250 msec/movement.

[12] The limit of repetitive movement of the hand, foot, or tongue is about 10 movements/sec (Fitts and Posner, 1967, p. 18). Chapanis, Garner, and Morgan (1949, p. 284) cite tapping rates of 8~13 taps/sec (38~62 movements/sec, assuming 2 movements/tap). Fox and Stansfield (1964) cite figures of 130 msec/tap = 65 msec/movement. Repetition of the same key in Kinkead's data (Figure 2.15*b*) averages to 180 msec/keystroke = 90 msec/movement. The scribbling rate in Figure 2.5 was 74 msec/movement. We summarize these as 70 [30~100] msec/movement.

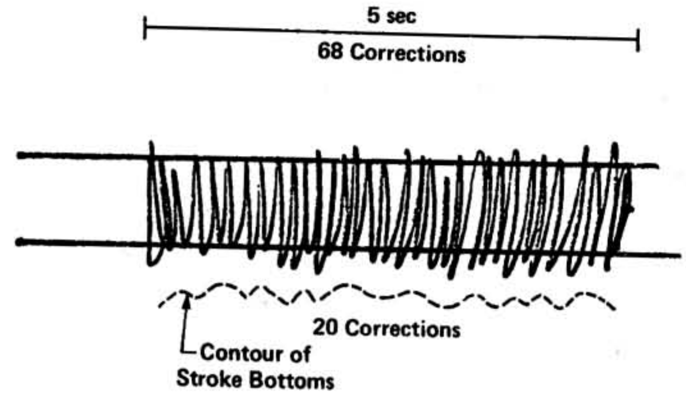

Figure 2.5. Maximum motor output rate.
Marks made by subject moving pen back and forth between two lines as fast as possible for 5 sec.

The Cognitive System

In the simplest tasks, the cognitive system merely serves to connect inputs from the perceptual system to the right outputs of the motor system. But most tasks performed by a person are complex and involve learning, retrieval of facts, or the solution of problems. As would be expected, the memories and the processor for the cognitive system are more complicated than those for the other systems.

COGNITIVE MEMORIES

There are two important memories in the cognitive system: a *Working Memory* to hold the information under current consideration and a *Long-Term Memory* to store knowledge for future use.

Working Memory. Working Memory holds the intermediate products of thinking and the representations produced by the perceptual system. Functionally, Working Memory is where all mental operations obtain their operands and leave their outputs. It constitutes the general registers of the Cognitive Processor. Structurally, Working Memory consists of a subset of the elements in Long-Term Memory that have become *activated;* this intimate association between Working Memory and Long-Term Memory is represented in Figure 2.1 by the placement of Working Memory inside Long-Term Memory. Although Working Memory information can be coded in many ways, the use of symbolic *acoustic* codes is especially common, related, no doubt, to the great importance of verbal materials to the tasks people frequently perform. The user of a telephone, for example, is especially liable to dial numbers mistakenly that sound like the numbers he has just looked up. *Visual* codes, if required by the the task, are also possible (as are some other types of codes). For purposes of the Model Human Processor we consider the predominant code types to be

$$\kappa_{WM} = \text{acoustic or visual}\,.$$

It is important to distinguish the symbolic, nonphysical acoustic or visual codes of Working Memory, which are unaffected by physical parameters of the stimulus (such as intensity), from the nonsymbolic, physical codes of the sensory image stores, which *are* affected by physical parameters of the stimulus.

The activated elements of Long-Term Memory, which define Working Memory, consist of symbols, called *chunks*, which may themselves be organized into larger units. It is convenient to think of these as nested abstract expressions: **CHUNK1 = (CHUNK2 CHUNK3 CHUNK4)**, with, for instance, **CHUNK4 = (CHUNK5 CHUNK6)**.[13] What constitutes a chunk is as much a function of the user as of the task, for it depends on the contents of the user's Long-Term Memory. The sequence of nine letters below is beyond the ability of most people to repeat back:

B C S B M I C R A

[13] It is also possible to think of these as semantic networks, such as those in Anderson (1980) and other recent publications. At the level of our discussion, any of these notations will suffice about equally well. See also Simon (1974) for a technical definition of chunk.

However, consider the list below, which is only slightly different:

C B S I B M R C A

Especially if spoken aloud, this sequence will be chunked into CBS IBM RCA (by the average American college sophomore) and easily remembered, being only three chunks. If the user can perform the recoding rapidly enough, random lists of symbols can be mapped into prepared chunks. A demonstration of this is the mapping of binary digits into hexadecimal digits:

0 1 0 0 0 0 1 0 0 0 0 1 0 0 1 1 0 1 1 0 0 1 1 0 1 0 0 0
0100 0010 0001 0011 0110 0110 1000
4 2 1 3 6 6 8

This last can be easily remembered. The coding must be done in both directions, binary to hexadecimal and hexadecimal to binary, and takes substantial practice before it can be carried out as part of a regular memory-span test, but it can be done. Indeed, with extended effort, the digit span can be increased enormously. A Carnegie-Mellon University student holds the current record at 81 decimal digits, presented at a uniform rate of 1 digit per second.[14] This particular event occurred as part of a psychological study, where it could be verified that all the gain was due to elaborate recoding and immense practice in its use and development, rather than any physiological endowment.

Chunks can be related to other chunks. The chunk **ROBIN**, for example, sounds like the chunk **ROBERT**. It is a subset of the chunk **BIRD**, it has chunk **WINGS**, it can chunk **FLY**. When a chunk in Long-Term Memory is activated, the activation spreads to related chunks and to chunks related to those. As the activation spreads to new chunks, the previously activated chunks become less accessible, because there is a limited amount of activation resource. The new chunks are said to *interfere* with the old ones. The effect of this interference is that the chunk appears to fade from Working Memory with time (unless reactivated), as the decay curves in Figure 2.6 show. The curves are significantly affected by other variables, including the number of other chunks the user is trying to remember, retrieval interference with similar

[14] Ericsson, Chase, and Faloon (1980); Chase and Ericsson (1981).

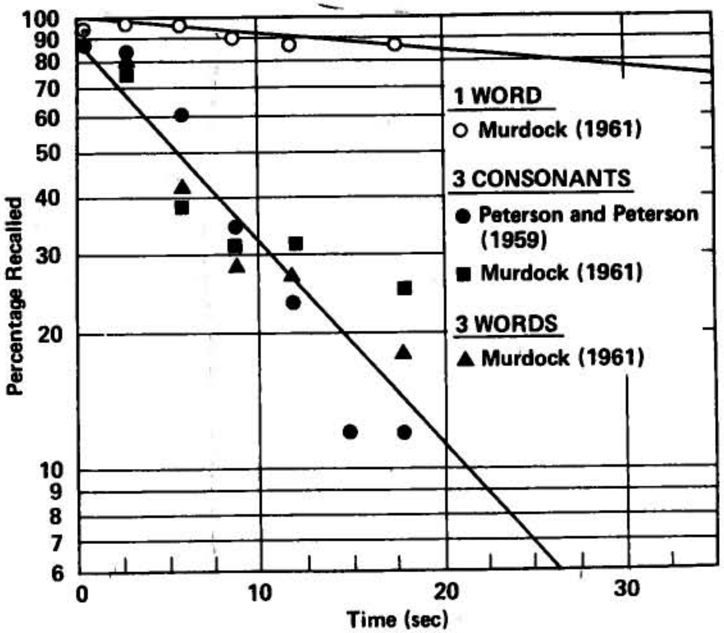

Figure 2.6. Working Memory decay rate.

Subject is given either one or three words or consonants to remember. He counts backwards (preventing rehearsal) for a time and then recalls stimulus. Graph plots proportion of items correctly recalled as a function of the time elapsed until recall began.

chunks in Working Memory, and input and retrieval memory strategies of the user. As a working value we take the half-life of 7 sec from the curve in Figure 2.6, which together with other data gives

$$\delta_{WM} = 7\ [5\sim226]\ \text{sec}.[15]$$

The decay parameter δ_{WM} has a wide range, because most of the apparent decay comes about from the details of interference, as we have noted above. But these details are difficult to analyze, so it is most convenient to accept the range and talk in terms of decay. Since the

[15] For three chunks, Peterson and Peterson's (1959) data (Figure 2.6) give a half-life of about 5 sec. Murdock's data (Murdock, 1961) in Figure 2.6 give a half-life of about 7 sec for 3 words and also 9 sec for 3 consonants. On the other hand, Melton's (1963) data give a much longer half-life of 34 sec. For one chunk, Murdock's data in Figure 2.6 and Melton's (1963) give half-lives of 73 sec and 226 sec, respectively.

decay rate is particularly sensitive to the number of chunks in the recalled item, it is useful to record the decay rate of representative item sizes:

$$\delta_{WM}(1\ \text{chunk}) = 73\ [73\sim226]\ \text{sec},[15]$$
$$\delta_{WM}(3\ \text{chunks}) = 7\ [5\sim34]\ \text{sec}.[15]$$

When people are asked to recall information a few seconds after hearing it, they use both Working Memory and Long-Term Memory to do so. Experimentally, these two systems have been teased apart showing that there is a *pure capacity of Working Memory* (example: number of immediately preceding digits recallable from a long series when the series unexpectedly stops),

$$\mu_{WM} = 3\ [2.5\sim4.1]\ \text{chunks}.[16]$$

When this pure capacity is augmented by the use of Long-Term Memory, the *effective capacity of Working Memory* μ_{WM}^* (example: longest number that can be repeated back) extends to the familiar 7±2 chunks,

$$\mu_{WM}^* = 7\ [5\sim9]\ \text{chunks}.[17]$$

Long-Term Memory. Long-Term Memory holds the user's mass of available knowledge. It consists of a network of related chunks, accessed associatively from the contents of the Working Memory. Its contents comprise not only facts, but procedures and history as well.

Apparently, there is no erasure from Long-Term Memory,

$$\delta_{LTM} = \infty.$$

However, successful retrieval of a chunk depends on whether associations to it can be found. There are two reasons the attempt to retrieve a chunk might fail: (1) effective retrieval associations cannot be found, or

[16] Crowder (1976) reviews several methods. Estimates are Waugh and Norman (1965) method, 2.5 items; Raymond (1969) method, 2.5 items; Murdock (1960*b*, 1967) method, 3.2~4.1 items; Tulving and Colatla (1970) method, 3.3~3.6 items. See also Glanzer and Razel (1974).

[17] Miller (1956).

(2) similar associations to several chunks interfere with the retrieval of the target chunk. The great importance of these links between particular chunks in Long-Term Memory, that is, the *semantic* coding of information, leads us to list it as the predominant code type,

$$\kappa_{LTM} = \text{semantic}\,.$$

To be stored in Long-Term Memory, information from the sensory memories must ultimately be encoded into symbolic form: a pattern of light and dark might be coded as the letter A, an extended pattern coded as a system error message. When the information from Working Memory becomes part of Long-Term Memory, the precise way in which it and the coincident Working Memory contents were encoded determines what cues will be effective in retrieving the item later. Suppose a user names a computer-imaging file LIGHT (as opposed to DARK). If he later scans a directory listing of file names to identify which ones were the ones he created and thinks of LIGHT (as opposed to HEAVY), he will not be able to recognize the file, because he will be using a different set of retrieval cues. As a principle of operation,

> ***P2. Encoding Specificity Principle.***[18] *Specific encoding operations performed on what is perceived determine what is stored, and what is stored determines what retrieval cues are effective in providing access to what is stored.*

Because of interference with other chunks in memory that are more strongly activated by the associations used as retrieval cues, information, despite being physically present, can become functionally lost. Stated as a principle,

> ***P3. Discrimination Principle.*** *The difficulty of memory retrieval is determined by the candidates that exist in the memory, relative to the retrieval cues.*

Items cannot be added to Long-Term Memory directly (accordingly, Figure 2.1 shows no arrow in this direction); rather, items in Working

18 Tulving and Thompson (1973).

Memory (possibly consisting of several chunks) have a certain probability of being retrievable later from Long-Term Memory. The more associations the item has, the greater its probability of being retrieved. If a user wants to remember something later, his best strategy is to attempt to associate it with items already in Long-Term Memory, especially in novel ways so there is unlikely to be interference with other items. Of course this activity, by definition, activates more items in Long-Term Memory, causing new items to appear in Working Memory, and use capacity. On a paced task, where a user is given items to remember at a constant rate, the percentage of the items recalled later increases as the time/item increases (the probability the item will be stored in Long-Term Memory and linked so it can be retrieved increases with residence time in Working Memory), until the time allowed per item is of the same magnitude as the decay time of Working Memory (after which, more time available for study does not increase the time the item is in Working Memory), around δ_{WM} sec/chunk = 7 sec/chunk.[19]

Storing new chunks in Long-Term Memory thus requires a fair amount of time and several Long-Term Memory retrievals. On the other hand, Long-Term Memory is accessed on every 70 msec cognitive-processing cycle. Thus the system operates as a fast-read, slow-write system. This asymmetry puts great importance on the limited capacity of Working Memory, since it is not possible in tasks of short duration to transfer very much knowledge to Long-Term Memory as a working convenience.

COGNITIVE PROCESSOR

The *recognize-act cycle*, analogous to the fetch-execute cycle of standard computers, is the basic quantum of cognitive processing. On each cycle, the contents of Working Memory initiate associatively-linked actions in Long-Term Memory ("recognize"), which in turn modify the contents of Working Memory ("act"), setting the stage for the next cycle. Plans, procedures, and other forms of extended organized behavior are built up out of an organized set of recognize-act cycles.

Like the other processors, the Cognitive Processor seems to have a cycle time of around a tenth of a second:

19 Newell and Simon (1972, p. 793) reviews experiments that gives times of 8~13 sec/chunk.

$$\tau_C = 70\ [25 \sim 170]\ \text{msec}\ .^{20}$$

The cycle times for several types of tasks are given in Figure 2.7. The times vary in the 25~170 msec/cycle range, depending on the specific experimental phenomenon and experimental circumstances with which one wishes to identify the cycle. We have chosen as a nominal value 70 msec, about at the median of those in Figure 2.7, but have included within the upper and lower limits all the estimates from the figure. As with the Perceptual Processor, the cycle time is not constant, but can be shortened by practice, task pacing, greater effort, or reduced accuracy.

> ***P4. Variable Cognitive Processor Rate Principle.*** *The Cognitive Processor cycle time τ_C is shorter when greater effort is induced by increased task demands or information loads; it also diminishes with practice.*

The cognitive system is fundamentally parallel in its recognizing phase and fundamentally serial in its action phase. Thus the cognitive system can be aware of many things, but cannot do more than one deliberate thing at a time. This seriality occurs on top of the parallel activities of the perceptual and motor systems. Driving a car, reading roadside advertisements, and talking can all be kept going by skilled intermittent allocation of control actions to each task, along the lines of familiar interrupt-driven time-sharing systems.

Summary. This completes our initial description of the Model Human Processor. To recapitulate, the Model Human Processor consists of (1) a set of interconnected memories and processors and (2) a set of

[20] On the fast end, memory scanning rates go down to 25 msec/item (Sternberg, 1975, p. 225, Figures 8 and 9, lower error bar for LETTERS). Michon (1978, p. 93) summarizes the search for the "time quantum" as converging on 20~30 msec. On the slow end, silent counting, which takes about 167 msec/item (Landauer, 1962), has sometimes been taken as a minimum cognitive task. It has sometimes been argued (Hick 1952) that the subject in a choice reaction time experiment makes one choice for each bit in the set of alternatives, in which case a typical value would be 153 msec/bit (Figure 2.22). Welford (1973, in Kornblum) has proposed a theory of choice reaction in which the subject makes a series of choices, each taking 92 msec. Blumenthal (1977) reviews an impressively large number of cognitive phenomena with time constraints in the tenth of a second range.

Rate at which an item can be matched against Working Memory:		
Digits	33 [27~39] msec/item	Cavanaugh (1972)
Colors	38 msec/item	Cavanaugh (1972)
Letters	40 [24~65] msec/item	Cavanaugh (1972)
Words	47 [36~52] msec/item	Cavanaugh (1972)
Geometrical shapes	50 msec/item	Cavanaugh (1972)
Random forms	68 [42~93] msec/item	Cavanaugh (1972)
Nonsense syllables	73 msec/item	Cavanaugh (1972)
Range =	27~93 msec/item	
Rate at which four or fewer objects can be counted:		
Dot patterns	46 msec/item	Chi & Klahr (1975)
3-D shapes	94 [40~172] msec/item	Akin and Chase (1978)
Range =	40~172 msec/item	
Perceptual judgement:		
	92 msec/inspection	Welford (1973)
Choice reaction time:		
	92 msec/inspection	Welford (1973)
	153 msec/bit	Hyman (1953)
Silent counting rate:		
	167 msec/digit	Landauer (1962)

Figure 2.7. Cognitive processing rates.
Selected cycle times (msec/cycle) that might be identified with the Cognitive Processor cycle time.

principles of operation. The memories and processors are grouped into three main subsystems: a perceptual system, a cognitive system, and a motor system. The most salient characteristics of the memories and processors can be summarized by the values of a few parameters: processor cycle time τ, memory capacity μ, memory decay rate δ, and memory code type κ. Each of the processors has a cycle time on the order of a tenth of a second.

A model so simple does not, of course, do justice to the richness and subtlety of the human mind. But it does help us to understand, predict, and even to calculate human performance relevant to human-computer interaction. To pursue this point, and to continue our development of the Model Human Processor, we now turn to an examination of sample phenomena of human performance.

2.2. HUMAN PERFORMANCE

We have said that in order to support cognitive engineering of the human-computer interface, an applied information-processing psychology should be based on task analysis, calculation, and approximation. These qualities are important for the Model Human Processor to possess if we are to address the practical prediction of human performance. Although it might be argued that the primitive state of development in psychological science effectively prevents its employment for practical engineering purposes, such an argument overlooks the often large amounts of uncertainty also encountered in fields of engineering based on the physical sciences. The parameters of soil composition under a hill, the wind forces during a storm, the effects of sea life and corrosion on underwater machinery, the accelerations during an earthquake—all are cases where the engineer must proceed in the face of considerable uncertainty in parameters relevant to the success of his design.

A common engineering technique for addressing such uncertainty is to settle on nominal values for the uncertain parameters representing low, high, and typical values, and to design to these. Thus a heating engineer might calculate heating load for a building at design temperatures of 10°F. for winter, 105°F. for summer, and a more common 70°F. day.

A similar technique helps us to address the uncertainties in the parameters of the Model Human Processor. We can define three versions of the model: one in which all the parameters listed are set to give the worst performance (*Slowman*), one in which they are set to give the best performance (*Fastman*), and one set for a nominal performance (*Middleman*).

The difference between the results of the Middleman (nominal) and the Fastman-Slowman (range) calculations must be kept clearly in mind. Secondary effects, outside the scope of the model, may mean that the