Are We Ready for a Cognitive Engineering?

Proceedings of the Annual Meeting of the Cognitive Science Society, 3(0)

Card, S. K.

Newell, A.

1981

Are We Ready for a Cognitive Engineering?

S. K. CARD, *Xerox Palo Alto Research Center* AND
A. NEWELL, *Carnegie-Mellon University*

It is an interesting irony that psychology is often criticized for its impracticality and its concern with minutiae while at the same time the crucial problems impeding many engineering developments are problems of psychological performance. Nowhere is this more true than in the area of human-computer interaction, whether it is a matter of automation in the cockpit of advanced aircraft or in the tasks of office workers. In both cases it is the lack of understanding of the determinants of psychological performance in the user that currently sets limits on the use of technology.

Yet, anyone who has had the task of trying to obtain from the literature psychological guidance for the design of an interactive computer system is aware of the great frustrations engendered by the jumble of empirical results and micro-theories, tightly bound to experimental paradigms, which he finds. It is not that psychology has no information to offer. Indeed, while the literature admittedly contains quantities of ill-founded trivia, there has come to be established in cognitive psychology a solid set of verified facts about the working of the human mind. But these facts are difficult to retrieve and to apply in new circumstances. This is true partially because in psychology, unlike some other fields, there has been insufficient effort devoted to codifying and condensing the knowledge learned into simple forms, usable by workers with other specialties. The codification and simplification of established facts is important to the progress of science if the results from one specialty are to become the tools of another (Latour and Woolgar, 1979), if the results of cognitive psychology are to coalesce into a science base for cognitive engineering.

But are we ready for a cognitive engineering? Is it now within our reach to create a systematic methodology for designing machines optimized for human cognitive performance? There is certainly the engineering need and there are also the promising developments in cognitive psychology, but it is a great mistake to think that, by merely listing a miscellaneous collection of results, cognitive psychology is thereby rendered usable to support a discipline of cognitive engineering. What is required is a more radical departure from what is usual in psychological research. Whereas the point of experimental manipulations is often to discriminate between competing theories no matter how small the discrimination, by contrast, in a psychology useful for engineering design, small differences are lost in the many approximations always necessary. What is important is the ability to do task analysis (determining the specific, rational means of accomplishing various goals), calculation (zero-parameter predictions of behavior capable of parametric variation), and approximation (simplification of the task and of psychological theory).

In the remainder of this paper, we wish to suggest a way in which numerous results from cognitive psychology might be included in a single model, usable by computer system designers and others. Though limited, this model (which we shall dub the Model Human Processor) does make it possible to calculate predictions of user performance, albeit of an approximate kind. The purpose of the model is not to provide a precise description of what is in the head, but to provide an economical and sufficient basis for applied analysis.

THE MODEL HUMAN PROCESSOR

The Model Human Processor can be described by (1) a set of pr[illegible] (Figure 1) together with (2) a set of principles of operation (Figure 2). The principal properties of the processors and memories are summarized by a small set of parameters. A similar technique has proved successful for simplifying the analysis of electronic information-processing systems (see Siewiorek, Bell, and Newell, 1981). The memory parameters used in the model are are

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

The only processor parameter used is

τ, the cycle time.

The complete model is elaborated and argued in Card, Moran, and Newell (in preparation). Here, we wish only to give an illustration of how a single parameter τ from the model can be used to support system engineering analysis.

According to the Model Human Processor, the mind is comprised of three partially coupled processors, the Perceptual Processor, the Cognitive Processor, and the Motor Processor, each with a similar cycle time, derived from the literature.

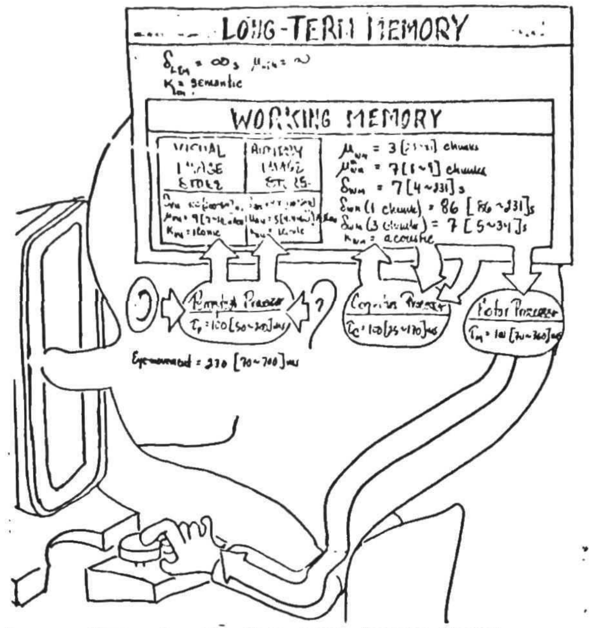

Figure 1. The Model Human Processor—Memories, processes, and Basic Principle of Operation.
Sensory information flows into Working Memory through the Perceptual Processor. Motor programs are set in motion through activation of chunks in Working Memory. Working Memory consists of activated chunks in Long Term Memory. Visual Image Store and Auditory Image Store can be thought of as special activations of the experiential and analogic properties of visual and auditory images. The basic Principle of Operation of the Model Human Processor is the Recognize-Act cycle of the Cognitive Processor.

On each cycle the contents of Working Memory activate actions associatively linked to them in Long Term Memory which, 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 for greater task demands and increased 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* $T_{pos} = I_M \log_2(2D/S)$, *where* $I_M = 100$ms/bit[70~120ms/bit].
- ***P6. Power Law of Practice.*** *The time* T_n *to perform a task on the nth trial follows a power law:* $T_n = T_1 n^{-r}$, *where* $r = .4[2 \sim 6]$.
- ***P7. The 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* $I_C = 150$ ms/bit [0~157 ms/bit]. *For n equally probably alternatives (Hick's Law),* $H = \log_2(n+1)$. *For n alternatives with different probabilities of occurring* p_i, $H = p_i \log_2(1/p_i + 1)$.
- ***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. The 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 operator movement, and (4) control knowledge for deciding which operator to apply next.*

Figure 2. The Model Human Processor—Additional Principles of Operation.

$$\begin{aligned}\tau_P &= 100 \text{ ms } [50{\sim}200]\\ \tau_C &= 100 \text{ ms } [25{\sim}170 \text{ ms}]\\ \tau_M &= 100 \text{ ms } [70{\sim}360 \text{ ms}].\end{aligned}$$

For some tasks (pressing a key in response to a light) the system must behave as a serial processor. For other tasks (typing, reading, and 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 lag showing that all three processors are working simultaneously.

Suppose a stimulus impinges upon the retina of the eye at time $t = 0$. At the end of one Perceptual Processor cycle, $t = \tau_P$, the image is assumed to be available in the Visual Image Store and the human able to see it. Shortly thereafter, a recognized, symbolic, acoustically- (or visually-) coded representation of at least part of the Visual Image Store contents is assumed to be present in Working Memory. In truth, this description is an approximation since different information in the image becomes available at different times, much as a photograph develops, and a person can react before the image is fully developed or he can wait for a better image, depending on whether speed or accuracy is the more important. According to the model, perceptual events occurring within a single cycle are combined into a single percept if they are sufficiently similar.

Once in Working Memory, information is processed by the Cognitive Processor's *recognize-act cycle*, analogous to the fetch-execute cycle of standard computers. 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 organized behavior can exist, but these are built up out of an organized set of recognize-act cycles.

Consequent to a decision to act by the Cognitive Processor, the action itself is controlled by the Motor Processor. Contrary to casual appearances, movement is not continuous, but consists of a series of discrete micromovements, each requiring one Motor Processor cycle of about $\tau_M = 100$ ms. The feedback loop from action to perception is sufficiently long (300~500 ms) that rapid behavioral acts such as [illegible] preprogrammed Motor Processor [illegible].

CALCULATING HUMAN PERFORMANCE

We now illustrate how the portion of the Model Human Processor we have described can be used to calculate answers to problems of human performance. To address uncertainties in the parameters of the model, we define three model versions: 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 typical performance (*Middleman*). Sensitivity analyses to see if the results of interest are affected by the true values of the parameters can be performed by making calculations for both Slowman and Fastman.

Example 1. Morse Code listening rate.

> What is the maximum rate, in words/min, at which Morse Code may be perceived? (Assume the old sort of American telegraph where dots and dashes are made by the clicks of the armature of an electromagnet, dots being distinguished from dashes by a shorter interval between armature clicks.)

Solution. If a dash takes less than one Perceptual Processor cycle $\tau_P = 100$ ms, there would be no way to tell it from a dot since, according to the model, durations smaller than τ_P cannot be perceived. Similarly, if two dots occur closer than about 100 ms they would be perceived as the same dot. So a reasonable approximation of the fastest rate is 100 ms/dot, 200 ms/dash, and 100 ms between letters.

If the probabilities for the various letters in English are multiplied by the calculated time per letter according to the above rule, we calculate a mean time of 459 ms/letter. Assuming 4.8 char/word (the value for Bryan and Harter's, 1898, telegraphic speed test),

$$\begin{aligned}\text{Max reception rate} &= (.459 \text{ s/letter} \times 4.8 \text{ letter/word})\\ &\quad + .200 \text{ s/word-space}\\ &= 2.4 \text{ s/word} = 25 \text{ words/min.} \;\blacksquare\end{aligned}$$

This number is in the range quoted by Bryan and Harter (1898) for very good, experienced, railroad telegraphers, 20~25 words/min.

An upper bound on the maximum rate is given by a Fastman calculation:

$$\begin{aligned}\text{Max rate} &= (100 \text{ ms}/50 \text{ ms}) \times 25 \text{ words/min}\\ &= 50 \text{ words/min.}\end{aligned}$$

Of course this calculation does not guarantee that 50 words/min is actually possible, only that it would be surprising if anyone were to be faster. In fact, the fastest performance known to Bryan and Harter was close to this rate, 49 words/min.

Example 2. Reaching to a button

> Suppose a user needs to move his hand D cm to reach a button S cm wide on a calculator (He cannot reach it by touch). How long will the movement require?

The movement of the hand, as we have said, is not continuous, but consists of a series of micro-corrections each with a certain accuracy. To make a correction takes at minimum one cycle of the Perceptual Processor to observe the hand, one cycle of the Cognitive Processor to decide on the correction, and one cycle of the Motor Processor to perform the correction, or $\tau_P+\tau_C+\tau_M$. The time to move the hand to the target is then the time to perform n of these corrections or $n(\tau_P+\tau_C+\tau_M)$. Since $\tau_P+\tau_C+\tau_M \simeq 300$ ms, n is the number of roughly 300 ms intervals it takes to point to the target.

Let X_i be the distance remaining to the target after the ith corrective move and X_0 (= D) be the starting point. Assume that the relative accuracy of movement is constant, that is, that $X_i/X_{i-1} = \alpha$, where α (< 1) is the constant error. On the first cycle the hand moves to

$$X_1 = \alpha X_0 = \alpha D.$$

On the second cycle, the hand moves to

$$X_2 = \alpha X_1 = \alpha(\alpha D) = \alpha^2 D.$$

On the nth cycle it moves to

$$X_n = \alpha^n D.$$

The hand stops moving when it is within the target area, that is when

$$\alpha^n D \leq ½S.$$

Solving for n gives

$$n = -\log_2(2D/S) \;/\; \log_2 \alpha.$$

Hence the total movement time T_{pos} is given by

$$\begin{aligned}T_{pos} &= n(\tau_P+\tau_C+\tau_M)\\ T_{pos} &= I_M \log_2(2D/S) \qquad (1)\\ &\text{where } I_M = -(\tau_P+\tau_C+\tau_M) \;/\; \log_2 \alpha.\end{aligned}$$

Equation 1 is called Fitts's Law (this derivation based on Keele, 1968). It says that the time to move the hand to a target depends only on the relative precision required, that is, the ratio between the target's distance and its size.

The constant α has been found to be about .07 (Vince, 1948), so I_M can be evaluated:

$$\begin{aligned}I_M &= -300 \text{ ms}/\log_2(.07) \text{ bits}\\ &= 78 \text{ ms/bit.}\end{aligned}$$

Results from various experiments give values in the in the I_M = 70~120 ms/bit range.

Example 3. Reaction time.

> The user is presented with two symbols, one at a time. If the second symbol is identical to the first, he is to push the key labeled Yes, otherwise he is to push No. What is the time between signal and response for the Yes case?

Solution The first symbol is presented on the screen where it is observed by the user and processed by his Perceptual Processor giving rise to associated representations in the user's Visual Image Store and Working Memory. The second symbol is now flashed on the screen and is similarly processed. Since we are interested in how long it takes to respond to the second symbol, we now start the clock at 0. The Perceptual Processor processes the second symbol to get an iconic representation in Visual Image Store and then a visual representation in Working Memory, requiring one cycle, τ_P. If not too much time has passed since the first symbol was presented, its visual code is still in Working Memory and the Cognitive Processor can match the visual codes of the first and second symbols against each other to see if they are the same. This match requires one Cognitive Processor cycle, τ_C. If they match, the Cognitive Processor decides to push the Yes button, requiring another cycle τ_C for the decision. Finally, the Motor Processor processes the request to push the Yes button, requiring one Motor Processor cycle τ_M. The total elapsed reaction time, according to the Model Human Processor, is

$$\begin{aligned}\text{Reaction time} &= \tau_P + 2\tau_C + \tau_M \\ &= 100\ [50\sim200] + 2\times(100\ [25\sim170]) \\ &\quad + 100\ [70\sim360]\ \text{ms} \\ &= 400\ [170\sim900]\ \text{ms}.\end{aligned}$$

This analysis could be repeated for the case ("name match") where the user is to press Yes if the symbols were both the same letter, although one might be in upper case, the other in lower case. Here, an extra Cognitive Processor cycle is required to get the abstract code for the symbol (Computed reaction time = 500 [195~1070] ms). Likewise, if the user were to press Yes when the symbols were only of the same class ("class match"), say both letters, yet another Cognitive Processor cycle would be required (Computed reaction time = 600 [220~1240] ms).

Experiments have been performed to collect empirical data on the questions presented in these examples. The finding is that name matches are about 70 ms slower than physical matches and that class matches are about 70 ms slower yet, a number in line with our 100 ms [25~170 ms] value for τ_C.

The forgoing examples start from task analysis of a problem and proceed through approximation and calculation to make predictions of human performance of the sort that might be used in an engineering analysis of cognitive behavior. Although it is hoped that the Model Human Processor itself will be useful for engineering, the real point is in the spirit of the enterprise: that knowledge in cognitive psychology and collateral sciences is sufficiently advanced to allow the analysis and improvement of common mental tasks. In short, that a cognitive engineering is now feasible, provided there is a disciplined understanding of how the knowledge must be structured to be useful. Of course, suggestions for improvement to the present model will occur all around. But then that is part of the idea and the challenge: to use the Model Human Processor as a framework into which new research results and insights can be fit in a way amenable to use in the cognitive engineering of practical mental tasks.

REFERENCES

Bryan, W. L. and Harter, N. (1898) Studies in the physiology and psychology of the telegraphic language. *Psychological Review 4*, 27-53

Card, S. K.; Moran, T. P.; and Newell, A. (in preparation) *The Psychology of Human-Computer Interaction.* Hillsdale, N. J.: Erlbaum Associates.

Keele, S. (1968). Movement control in skilled motor performance. *Psychological Bulletin*, 387-403

Latour, B. and Woolgar, S. (1979) *Laboratory Life the Social Construction of Scientific Facts.* London: Sage Publications

Siewiorek, D.; Bell, G.; and Newell, A. (1981) *Computer Structures.* New York: McGraw Hill

Vince, M. A. (1948) Corrective movements in a pursuit task. *Quarterly Journal of Experimental Psychology, 1*, 85-103