Wright, P. (1987*a*) Reading and writing for electronic journals. In: B. Britton (ed.) *Executive control processes in reading*, Hillsdale, N.J.: Lawrence Erlbaum Associates, 23-55.

Wright, P. (1987*b*) Issues of content and presentation in document design. In: M. Helander (ed.) *Handbook of human-computer interaction*, Amsterdam: North Holland Publishing Company.

Wright, P. (1987*c*) Writing technical information, *Review of Research in Education, 14*, (in press).

Wright, P., Hull, A.J. and Lickorish, A. (1984) Psychological factors in reading tables. *Proceedings of the 23rd International Congress of Psychology*, Mexico: Trillas, 194.

Wright, P. and Lickorish, A. (1984) Ease of annotation in proof-reading tasks, *Behaviour and Information Technology, 3*, 185-194.

Wright, P. and Lickorish, A. (1987) Colour cues as location aids in lengthy texts on screen and paper, *Behaviour and Information Technology*, (in press).

Young, R.M. and Hull, A.J. (1982) Cognitive aspects of the selection of viewdata options by casual users. In: M.B. Williams (ed.) *Proceedings of the 6th International Conference on Computer Communication*, London, 571-576.

Young, R.M. and Hull, A.J. (1983) Categorisation structures in hierarchical menus. *Proceedings of the 10th International Symposium on Human Factors in Telecommunications*, Helsinki, 111-118.

Reading as Goal-Driven Behaviour

Don G. Bouwhuis*

1 Reading between psychology and ergonomics

Ergonomics deals with the unhappiness that people experience in performing tasks, handling tools and objects, controlling equipment and understanding complex systems. Psychology, in general, concentrates on the perceptual, cognitive and motor processes that form the basis of these activities. Yet, despite this intricate mutual involvement of task and process, there is frequently a lack of complementarity and a mismatch of levels of description between psychology and ergonomics that impede proper understanding of psychological processes in human activity.

A prime example that may illustrate the dissociation between psychology and ergonomics is the reading task. Psychology and ergonomics study different aspects of the same reading activity either for fundamental or applied purposes. The studies discussed by Wright (*this volume*) are of particular interest here, as the main theme seems to be application-oriented, while the results have profound implications for fundamental research. For the purpose of this discussion it will be useful to define reading as *a means to an end*. First, *means* reflects the physical substrate, script, print, paper, and the reading environment that serves the purpose of reading. Second, reading is a way of information processing using sensory, perceptual and cognitive means. Both physical and perceptual means will be discussed and it will be argued that they are largely dependent on each other. Both involve fundamental and applied aspects that appear to be very difficult to disentangle. In the second part of this discussion the *end* of the means is analysed more closely. Although the means enable the end, it will be shown that the end, or goal, largely determines the means.

* Institute for Perception Research – IPO, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

WORKING MODELS OF HUMAN PERCEPTION
ISBN 0-12-238050-9

2 Origins

Research on reading may at first sight seem typically application-oriented. The underlying processes of reading, however, were judged so fundamental that it became one of the first comprehensive mental activities that psychology started to study (Javal, 1878; Grashey, 1885; Cattell, 1885; Goldscheider and Müller, 1893). As a matter of fact, quite a number of fundamental processes discovered in these early phases (Huey, 1968) set the scene for decades to come. Only when theoretical interest in reading started to wane after the turn of the century, did applied research that in some sense could be called ergonomic find its way into the field of reading. This was a major departure from both the structuralist and the more recent functionalist points of view, because the major motivation for these studies was to optimize the reading process. Experimental results were obtained in various areas of interest. It soon turned out that readers could not learn how to direct their visual selective attention in a more efficient way (Whipple, 1910). Neither did it prove possible to train readers' eye movements to improve reading. Studies on fatigue showed that readers could not be tired by a continuous reading task; in prolonged reading sessions (Carmichael and Dearborn, 1947), it was only sleep deprivation that caused the subjects to give up. In the field of individual differences it appeared impossible to determine those identifiable faculties that could account for reading performance. Studies on typography demonstrated that even wide variations in typeface affected reading performance only slightly, while letter size variation, too, produced only minor changes in reading speed (Tinker, 1964).

So the paradoxical situation was arrived at that a highly complex task, involving a host of resources of a visual, cognitive, memorial and linguistic nature, and requiring extensive training, proved refractory to most sorts of external manipulation. At least implicitly, it could be concluded that reading is a basically sound ergonomic activity. With some irony one might remark that this status justified reading as an object of valid psychological endeavour. Indeed, if perception, memory, thinking and language appeared to be the fundamental building blocks of cognitive functioning, reading was so closely tied to each of these that it could equally well be considered to be fundamental.

Evolution

Notwithstanding the dominant position of reading in psychological research, reading – and conversely writing – is a relatively recent invention, or technical development, with great applicability. Requirements for reading and writing have not been shaping human perception and cognition for a long time. In evolutionary terms vision, hearing and locomotion had already developed to a considerable extent when verbal language came into being, and it, in turn, was very likely to have been preceded by gestural and sign languages. The latter are already symbolic and definitely entail processes of abstraction. Verbal language came into being, in all probability between 5 million and 2 million years ago (Leakey and Lewin, 1977). Considering the highly effective expressive power of verbal language, this is a period that seems long enough to have shaped cognitive functioning of human beings substantially. This is much less the case for the making of artefacts and symbolic engravings, dating from only 100,000 years ago (Valladas, Reyss, Joron, Valladas, Bar-Yosef and Vandermeersch, 1988). Sculpture has its early roots about 50,000 years ago. It seems unlikely that they shaped human functioning, and specifically survival, to the same extent as language did, in as much as it seems rather more a product or consequence than a cause. A systematic form of what we may call writing arose only some 6000 years ago in Mesopotamia. The direct precursors of writing systems were three-dimensional tokens representing economic and metrological units of commodities; grain, land and animals. When these tokens were impressed in clay tablets a sign script evolved that could be used much more easily in mercantile transactions (Schmandt-Besserat, 1981). The actual basis of a still more abstract system where symbols denote language units rather than specific pictorial references, Sumerian writing, developed slowly from these first beginnings. To the extent that writing or reading is an even more abstract activity than imaging it can hardly be expected to have affected human cognitive functions greatly by genetic selection. Even nowadays illiterates may be charming partners and experience no problem at all in raising equally charming families. Conversely, it seems much more likely that the potentialities and restrictions of human functions have continually shaped and refined the reading and writing systems that were used. In this respect the relatively short existence of reading becomes a factor of importance, as even 4000 years provide ample opportunity for change, adaptation and optimization. It becomes a very long time in comparison with the existence of the telephone, the radio, or the database program on a computer. We

may note here, too, that the actual design of a reading/writing system is an open-ended issue and could lead to a sheer infinity of solutions, only limited by human imagination. This reminds one strongly of the infinite functionality claimed of computer systems. Although these have almost unlimited potential only a disappointing number of functions are realized in actual practice. It, therefore, comes as no surprise that writing systems are far more limited and restricted in their possibilities than one might have hoped for. Even a cursory inspection of some widely distributed writing systems reveals a striking uniformity. Alphabetic writing systems consist invariably of approximately 26 ± 2 characters; this holds for the Western Roman alphabet, the Greek, the Cyrillic[1] and the Arabic writing system. In all of these systems, except for a few characters in Cyrillic, the letters denote a single sound of speech, or, more specifically, a particular phoneme, but in all languages using these alphabets, the actual number of phonemes is much higher. What the alphabets have in common is that the symbol strings are in many cases just approximations of the speech sounds, and in many languages exceptions from basic pronunciation rules are the rule rather than the exception.

While a syllabary like the Japanese katakana might seem to have many more characters than the curious number 26, its construction actually reveals a simpler underlying organization. Syllables consist of CV combinations showing an orthogonal structure made up from a consonant and a following vowel. The number of identifiable or distinguishable symbols, therefore, comes close to the number of letters in an alphabet.

Writing and perception

It should not be underestimated to what extent reading habits are determined by requirements imposed by the writing tools, originally a wooden stylus and clay or wax (Powell, 1981). Repetitive writing mutated symbols to forms easier to produce, forcing increasing abstraction. Rotation through 90° to the left made top to bottom writing into left to right, which still is the way practised in most alphabets. The rotation was induced by the easier handling of the embossing stylus, the tip of which was shaped like a wedge. (Powell, 1981). Even nowadays it is not always clear whether the Sumerian lines of writing are arranged vertically or horizontally (Green, 1981). Be this as it may, linear arrangement of characters induces an inherent frame

[1] Although Cyrillic has formally 33 characters, a number of them are structurally similar. The actual number of truly different characters is 28.

of reference, a specific orientation, on all characters. This is a uniform property in almost all forms of script. It can be regarded as a major step of abstraction which departs radically from perceptual experience. No object in our visual environment changes identity on (non-negligible) rotation, or reflection; only letters do. A common example for the Roman alphabet is formed by the letters **b**, **d**, **p** and **q**. Probably part of this convention can be explained by the trade-off between complexity of characters and writing time, which works in the direction of simple but distinguishable characters. Acknowledging that this was a venture, it succeeded against all odds.

Script and speech

Another phenomenon that has not yet been resolved satisfactorily is why a generalized representation of the phoneme has become the elementary building block of the dominating alphabetic systems. Undoubtedly, there is a relation with the relatively limited number of articulatory movements, of which there are still many more than letters. If the number of distinctive visual symbols is to be minimized, the phoneme is a better choice than larger segments, of which there are many more, showing, moreover, a great variation because of coarticulation. Ease and speed of writing could tend to limit individual characters to a few strokes only. With six position-specific strokes 63 characters can be constructed, but many would be very confusable and there would be no redundancy in the symbol set at all. A significant step towards redundancy is to delete at least 50%, which brings us close to 26 ± 2. This is probably more than one would prefer from the visual or graphic point of view, but less than the number of phonemes. So a number slightly less than 30 would seem to be an uncomfortable but unavoidable compromise.

Yet even the choice of our current letter as the identifier of a phoneme was a venture. The phoneme is not a naturally occurring component of linguistic awareness. It is well documented that pre-readers have difficulty in phonetic segmentation and that phonetic segmentation ability is closely linked with reading acquisition (Liberman, 1973; Content, Kolinsky, Morais and Bertelson, 1986). Phonetic segmentation ability clearly also distinguishes normal readers from dyslexics (Morais, Cluytens and Alegria, 1984). In a recent series of studies among illiterates in Portugal Morais and co-workers found that illiterates just cannot segment separate phonemes, not even word-initial, out of spoken words. The maximal attainment seems the ability of some to rhyme at the end of lines. It should be noted that rhyming spans

generally at least two phonemes. Seen in this light the choice of phonemic units for the symbol set was a precarious one, that proved to be fortunate.

Perceptual frames of reference

Concatenation of letters into words has the same effect as the fixed frame of reference for letters had: the word, too, assumes also a fixed frame of reference aligned from left to right (or the reverse). The only exception was in the short-lived boustrophedon writing, where lines are written alternately from left to right and right to left. No current writing system employs this alternation. There is an empirical basis for this phenomenon. It is well known from studies on text rotation and transformations (Kolers and Perkins, 1975) that rotations and reflections seriously degrade reading performance. What Kolers (1976) also showed was that learning to read transformed text was highly specific with respect to the type of transformation, implying that a new, but still invariant frame of reference was embodied in the reading process.

The choice of clay as a writing substrate, easy to emboss and to correct and self-fixing through hardening[2] also brought its inconveniences. To cite Powell (1981) you might get a wholly different idea of the cuneiform Sumerian writing system if you had to carry a copy of Herodotus from Babylon to Damascus on your back.

The obvious choice for writing is two-dimensional only, and found its way into papyrus or specially treated animal skin. The dearth of this material necessitated putting as much as possible on a sheet. Essentially this fixes the length of the line, since the eye must be able to find the beginning of a line easily from the end of the preceding line. If line spacing is fixed, the length of the line must permit an approach angle that is sufficiently large to hit the start of the next one accurately. As sheets are rarely exactly square – and animals actively resist assuming this shape – the natural tendency will be to generate a block of text that is higher than it is wide: essentially what we have come to recognize as the standard page. Also it would be wasteful not to write on the other side of a sheet; visually this makes no difference at all. The habit of writing on scrolls does not contradict these observations, as scrolls have the standard 'pages' aligned sideways along their length. The advantages are the immutable order of pages and the place keeping facility. Another solution for storing larger quantities of writing is to stack

[2] As a rule clay writing tablets were not hardened by firing them, although this sometimes happened in incendiary accidents. Frequently tablets were reused.

approach angle

the sheets as in a book. The advantages of a stack are easier production and fast random access, assisted by the height of the stack as a place keeper. If sheets are to be collated permanently, a form of binding will be necessary. Disregarding mechanical wisdom, pages should be bound on one side only to render the stack accessible. Not all page sides are equally suitable for this (see figure 1). If the binding was on the top side, either the writing would be upside down after a page was turned over or one would have to write upside down. This would introduce ambiguity as to what was top and bottom of the sheet. A sheet itself, therefore, also has a built-in frame of reference consisting of top and bottom; the left and right apply only to a single face of the sheet. It may be pointed out as well that turning over pages along the top side produces a vertically extended reading space that is difficult to read because of the variability of distances between eyes and text, and also difficult to handle mechanically. Binding the right side

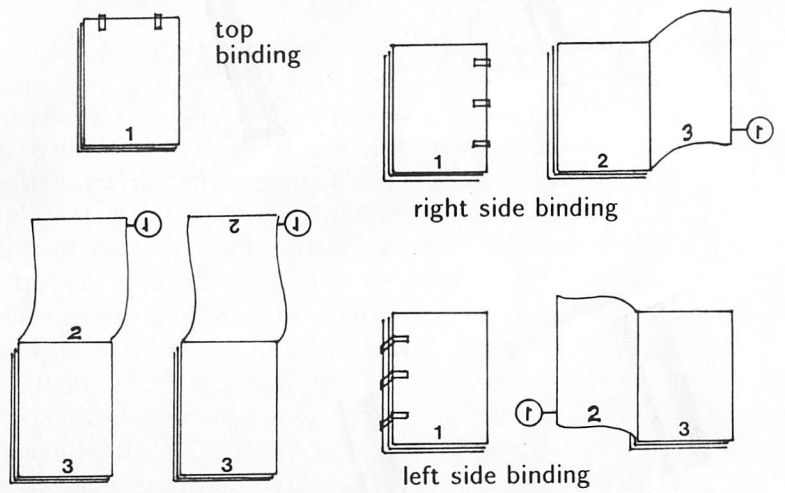


Figure 1: Text page organization as a result of binding methods. Numbers in circles refer to real orientation of page numbers overleaf.

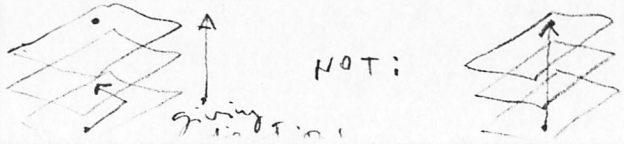
implies that on turning over the first page, the reader continues on the left page being uncovered, not on the right one which is the back of the first page. This essentially destroys the sequential integrity of text and pages. The disadvantages resulting from top binding and right side binding are combined in binding the bottom side, making this the worst option of all. So the only choice is really to have the binding on the left side for alphabets with a left-to-right writing system, and on the right side for right-to-left

writing systems. Again this implies the priority of the top-bottom property over that of the left-right property.

There is basically one exception to the single text block/page stack, which is the newspaper. This is a relatively recent development, not much older than 250 years, and it used to be much smaller in its early days. The single block of text has given way to narrow columns much less wide than they are high, which facilitate reading. The enlargement of the newspaper page has made random access possible on a single page, considerably facilitating scanning. Fewer turns of pages, of these unwieldy surfaces are needed to consult every section. Both the book, mostly devoted to a single topic, and the newspaper covering a variety of topics seem to have been adapted remarkably well to the human reader and his intellectual interests.

The 2++ dimensional man

There is one other factor in the success of reading a page that has received disappointingly scant attention. It is the spatial ability that man shares to some degree with his evolutionary ancestors. Basically, the ability of humans to grasp the spatial organization of physical objects is restricted to two dimensions. This seems likely as it is functional for surface-dwellers to deal successfully with the organism-centred spatial layout in the plane of locomotion. It wastes no neural complexities on things that are too high or too low. Things are in front, behind, left or right, and people can learn this fairly efficiently. Surprisingly, reading and understanding two-dimensional maps is one of the very few cognitive activities that does not show a difference between novices and experts (Thorndyke and Stasz, 1980). This suggests already that humans can cope quite adequately with maps, that are two-dimensional abstractions of space. Gilhooly, Wood, Kinnear and Green (1988) confirmed this and found that only in topographic vs planimetric maps experts were superior to novices, in that they could use detailed topographic and contour information better. In multi-storey buildings people give directions by mentally transferring the floor they are on to the intended one, and then proceed according to a two-dimensional lay out, rather than being able to point in a straight line. People are rather poor at imagining, or even recognizing, the shape of a three-dimensional rotated cube (Hinton, 1979; Bouwhuis, 1985) which by itself looks a most simple object. Even though a sphere looks much simpler, probably because it is rotationally invariant, human orientation on the three-dimensional globe is not impressive. Air travellers, knowing that New York lies on a latitude



some 1200 km south of Amsterdam, frequently cannot understand why the initial trajectory heading from Amsterdam is NNW rather than WSW. The lack of a southern component seems to them counterintuitive, yet on a globe the itinerary is the shortest possible. Note, however, that the air travellers' belief would hold on the surface of a cylinder, corresponding to the well-known Mercator projection. In fact, the subjective or imagined cylindrical projection takes the third dimension out of the trajectory, since the surface of a truncated cylinder has many similarities with a bounded plane, certainly with respect to its frame of reference. A striking illustration of the difficulty people have in combining the vertical with the horizontal plane is provided by Hinton (1987).

However, people can do better than two-dimensional organization, though the claim here is that this would not go as far as the $2\frac{1}{2}$D sketch of Marr and Nishihara (1978). In their description the normals to any three-dimensional plane are part of the $2\frac{1}{2}$D-sketch. Here we suppose that the organization that can be immediately perceived is the proximal order of objects, basically in all directions. So a human observer perceives and knows which object occludes which other object, or is closer. This reflects only an ordinal representation of proximity with a rather low-resolution representation of physical distance. People are well aware that they are standing *on* a floor, walking *beneath* the ceiling, what is lying *on* top of or *under* what. Remarkably, object occlusion does not at all impede object perception; on the contrary, occluded objects automatically undergo cognitive completion (Gerbino and Salmaso, 1987) and appear subjectively to be an integral whole. This phenomenon implies among other things that objects have a frontal plane as well as an inferred distal rear plane, and this holds equally well for two-dimensional objects such as a page. The only difference between a page and its rear is that left and right are reversed. On a single page readers generally know fairly well where they have read particular sections. It is much harder to locate these in a book, as the stack is too large to address it accurately. Yet the thickness of the remaining stack indicates approximately how far the reader is removed from either the beginning or the end. Again, this is much simpler with a newspaper where a specific page can usually be found quickly. It is obvious that in virtual two-dimensional structures such as stacked pages, sequential addressing is much more efficient if page units are used rather than units of physical distance. So perhaps we should refer to human subjective space dimensionality only as 2++, while realizing that subjective structuring of this space is a powerful faculty of the perceptual system. The importance of this is that there is an

almost perfect correlation between the temporal organization of the reading process and the spatial organization of text. Conceptual components of the text thus acquire a system of spatial correlates. It is tempting here to speak of neighbourhood, a spatial denomination. A neighbourhood can be defined on many levels of text; graphemic, lexical, syntactic, semantic and thematic. All of them are reflected in the spatial neighbourhood of written and printed text. Spatial integrity maintains integrity of text.

3 Screen versus sheet

Considering how reading texts are displayed, or have to be even on well-engineered CRT screens as opposed to conventional print, we can observe a number of dramatic differences.

First, the displayed text is always 'front', as there is no legible backside. This prevents the reader from organizing text sections into a left/right structure as in a book, increasing the actual text planes to be remembered by a factor of two. Worse still, there usually is no well-defined top and bottom, preventing a subjective vertical structure. The possibility of two-way scrolling, hailed as a technical breakthrough, is in fact disastrous for an effective cognitive organization of long passages of text. When used it obliterates the previously established vertical frame of reference, replacing it continuously by a new one. It is, therefore, quite understandable that WrightWright, P. (*this volume*) found that screen readers wanted to have information on the actual length of the text to be read. Usually any indication on length is lacking as the screen presentation either stacks the pages invisibly, as in a window system, or gives the impression of a continuous vertical scroll with invisible ends. No reading medium shows so little of the whole text as the current CRT screen. Wright and her colleagues (Tombaugh, Lickorish and Wright, 1987) provide empirical evidence that stresses this point. They had readers relocate information in 'stacked', slightly offset windows of sequential text. This meant that individual windows, or pages, could be selected for presentation. This proved to be a significant help to readers looking for specific information in lengthy texts. These experiments show that the reading task is much more dependent on a number of purely physical conditions than the complexity of the reading process itself would suggest. No one would propose that, for example, thinking, depends on the clothes you are wearing. It is essentially those physical parameters which seem only accidental to the reading process that have never been studied

seriously in a comprehensive account of the reading task as a whole.

A factor that has not yet been mentioned in the comparison between print and CRT display is the physical quality of the imaged characters. In print, characters have uniformly solid strokes, with edge blurring lower than the eye's resolution and rather high subjective contrast. None of these properties are attainable, even singly, with CRT screens. The effects of low contrast have been investigated in a number of studies (Van Nes and Jacobs, 1981; Leermakers and Boschman, 1984) showing both performance degradation and considerable loss of 'visual comfort'. Although originally typeface was not found to affect legibility seriously (Tinker, 1965), new computer and television typefaces challenge this position. Character design is frequently subject to numerous technical restrictions and does not follow typographic tradition. The use of a sparse dot matrix is a case in point. Reduction of letter feature prominence was studied by Schiepers (1976), who found a marked decrease of letter identifiability. The effects of degrading even single letter identifiability may be expected to be rather serious. Straightforward application of existing quantitative word recognition models demonstrates that even slight letter degradation may reduce dramatically the efficiency of word recognition as a whole (Bouwhuis and Bouma, 1979). Reading turns out to be a fight for percentages.

How ergonomic is reading?

Taking all arguments together, reading appears not to be the ergonomically sound activity that it is mostly implicitly assumed to be. Centuries of trials, redesign, adaptation and refinement, continuously watched by countless generations of readers, have resulted in a tradition for print that is optimized as far as it can be with the available means. Few will argue that printing is not an art or specialism, to say nothing of writing. This situation has given rise to the belief that reading as a process is a natural and optimal way of information uptake in general. Even if it is, it is only attained at great cost. Most effort in basic education is spent on learning to read, both in the technical and in the conceptual sense. Nevertheless, it appears that alphabetic writing systems carry a percentage of dyslexia, mostly developmental in nature, with an estimated minimum of 5%. None of this applies to activities such as walking, eating, seeing or hearing, which are not less important. In view of the many counterintuitive perceptual properties of text there is even reason to be surprised about the relative rarity of reading problems. It seems that the reading process is a compromise, and we do

not know exactly of what. One general conclusion to be drawn is that one would be ill-advised to try and change the reading process beyond the conditions we know to be optimal, or at least adequate. This applies to spelling reforms, to typographic design and to reading education. It applies, too, to reading from a computer screen.

None of this touches the issue of the need to read, which there obviously is. This is not a single variable and it determines the course and nature of the reading process to a very great extent. It is to these determining variables that we turn now.

4 Reading for needs

Without aiming at completeness we distinguish three types of reading: reading for *entertainment*, for *information* and for *consultation*. People read novels, and newspaper columns for entertainment; even thrillers. They are supposed to start at the beginning and work their way towards the end, only to be interrupted temporarily. During reading there is supposed to be continuous, regular progress. Interestingly, some books will not be read to the end; readers may stop sooner or later when they get bored and lose interest or find the text too difficult. This fate is claimed for many bestsellers; it is said to hold for 'The Name of the Rose' (Eco, 1983) and 'Gödel, Escher, Bach' (Hofstadter, 1979) respectively. While the arguments for stopping may seem valid to the stranded readers, an important question is why they started to read in the first place. Ignoring the uninteresting issue of social desirability, the most probable factor is the need for entertainment, which can be satisfied most easily by reading. Nobody reads just for the fun of the saccades.

The second type of reading, viz. for information, can be found in reading the news, or studying a textbook. There is in principle no need to proceed from the beginning to the end. The reading process usually is selective, and there need not be regular progress. Instead, sections may be skipped, or repeated, the reader may go back to earlier sections, or discontinue altogether and start somewhere else. Most probably the reader is either looking for specific or relevant information, or, on finding the informational value too low, moves to a potentially more informative target.

The last type, reading for consultation, can be found in looking up a dictionary, finding a telephone number or searching in the Radio or TV Times. The reader looks for one specific item of information, but has to

make his or her way through the book or magazine in order to recover it. It is entirely natural to stop reading when the desired information has been found. It is even considered outright silly to continue reading in a dictionary or a telephone directory without an intended target.

These three examples may suffice to show that reading

1. need not to be continuous and regular,
2. may not always be clearly motivated, and
3. may stop to give way to higher priority activities.

Task-action protocols

Without adopting a teleological perspective from the outset it will be convenient for the following reasoning to assume that cognitive behaviour is purposeful or goal driven. The importance of this assumption is demonstrated by Rosenbloom, Laird and Newell's description of the SOAR system (*this volume*) and Bösser's discussion of the same (*this volume*). This assumption implies that there must be an intention, leading to an activity that results in goal attainment. Sequences like these are called task-action protocols and instances have been described by Card, Moran and Newell (1980), Norman (1984) and Epstein (1986), shown in table I.

All three protocols show great similarity even if Epstein constructed his in connection with motor control. The first two were designed explicitly to deal with human-computer interaction. What all three protocols are saying is that actions are started by an intention, or goal. The intention can in principle be satisfied in several ways. In the GOMS model OPERATORS represent single actions, the concatenation of which represents a METHOD to attain the GOAL. Users select from alternative METHODS in order to reach it. As the model, unlike a production system (see Rosenbloom, Laird and Newell, *this volume*), does not deal specifically with errors there is no evaluation stage or feedback. The other two protocols are closed in the sense that feedback is present on some or all levels. When a selection has been made the user/actor may, in performing the action, assess, whether the goal has been attained. This assessment occurs on-line in motor control, which is why Epstein's protocol does not show an evaluation phase.

Table I: The action protocols by Card *et al.*, Norman and Epstein.

Card, Moran & Newell GOMS	Norman (ATS)	Epstein
GOALS	forming the *intention*	intention
OPERATORS		representation of actor-environment setting
METHODS		perceptual assessment of compatibility
SELECTION	*selecting* an action *executing* the action *evaluating* the outcome	selection of action execution

Every task can be considered to form part of a hierarchy in which tasks higher up in the hierarchy represent more global activities, and those lower down in the hierarchy more detailed ones. Tasks contain subtasks to an almost infinite extent. Higher-order tasks are increasingly under more cognitive control than lower-order ones, cf. the desire for food or entertainment vs eye movements in reading, or pupillary reactions. Neither Norman's nor Epstein's action protocol is clearly located within the task hierarchy. One example given by Norman is that of the user of a text editor wanting one page to look nicer. Of the many options a two-column layout may be chosen. Once the layout change has been carried out the user evaluates whether the result matches the original intention. A lower-order task is to modify a mistyped letter. One could move back the cursor and overstrike; zap the character and insert the new one, change the character by a change command or whatever means is available. The motivating intention is also to make the page look nicer, but in a specific and much more detailed way. The intention to make the page look better can be safely assumed to be dependent on a higher-order goal: e.g., to increase readability, to satisfy a customer, to demonstrate differences in layout, etc. In fact, these task-action protocols can be extended practically indefinitely both upwards and downwards. If this property is attributed to Norman's action protocol – it is slightly more difficult for Epstein's – it shows certain similarities with the layered protocols as defined by Taylor (1987*ab*). The basic property of the layered protocol is that intentions can only be compared at the same level. In discussing his Action-Trigger Schema (ATS) system (see table I) which formed the origin of his task-action protocol, Norman (1981) points out that the intentions are only specified on the highest level of memory representations. The intention can be realized only by an action, residing on a lower, less abstract level. In that sense the action is independent of the intention.

Turning now to reading, we have mentioned three kinds of reading: entertainment, information acquisition and consultation. All three goals satisfy the requirements for being intentions to read.

Scheduling in reading for entertainment

Any reading session must ultimately come to a stop. Even the most determined readers in the prolonged reading experiments of Carmichael and Dearborn (1947) had to stop because of sleep deprivation, the need for food, drink and bodily movement. These factors also represent the result of an intention that is mediated by elementary physiological requirements. If stopping reading is a conscious decision, then it represents a choice, in which preference is given to the satisfaction of needs other than that for entertainment, or, in this case, the curiosity of the experimenter. In less extreme cases reading for entertainment is meshed in with normal daily activities. Which action is chosen in preference to reading depends on the time of day and external needs. Description of this continuously changing action pattern is not primarily part of reading theory, but might well fit in with the layered protocol description. From this point of view a theory of not reading is a theory of the other actions that take its place.

The factors just discussed are external to reading; but there are also factors within the reading process that determine its continuity or interruption. On a global level reading consists of a perceptual task (fixation + recognition), an eye-movement control task (saccades) and a cognitive-integration task (understanding). While the perceptual tasks may proceed smoothly, it may well be that understanding is impeded by the occurrence of unfamiliar words, complex style or incompleteness of presentation (frequent in detective stories). As a case in point an example is presented here from an experiment by Buswell (1920) who measured eye movements in texts with closely occurring homonyms. The passage to be read was:

... coming
down the road. She had tears in her dress
and also tears in her eyes...

The word *tears* occurs twice in close succession, with different meanings, and this caused many of Buswell's subjects to show erratic saccades in the vicinity of both *tears*. Figure 2 shows the saccades and fixation times of Buswell's subject H13. What is approximately visible in the functional visual field in one fixation has been displayed horizontally; the vertical connecting lines indicate the fixation time. After a regular progress from: *down the road...*, it is around the word *tears* that regressions occur which persist throughout this line, and continue up to the next. If reading is measured by progressive saccades, then the reader has stopped three times in the first line and twice in the next. The simplest interpretation is that this subject aimed at understanding by ambiguity resolution, and had to discontinue reading at various points in time in order to attain consistency of interpretation. The intention to stop, regress and proceed is mediated by lack of local understanding. Subjects describe this by commenting that they have to take a good look again; in fact this good look takes the form of a number of fixations in the area of ambiguity. The word *tears* is actually within the functional visual field six times in the first line and five times in the second.

Also, reading for entertainment implies reading for understanding. Both textual difficulties and individual performance limits may influence heavily the progress of reading. Even when text has been completely understood, the reader may pause momentarily to elaborate the consequences of what has just been read. It might be the case that these moments are what the reader is trying to attain by reading.

To summarize, three reasons have been discussed for stopping to read for entertainment.

1. No human action can last indefinitely, but the reason for stopping reading can only be described by a theory of human action - not reading *per se*.

2. Reading is stopped in its forward motion when comprehension lags behind perceptual processing. Discontinuation can only be described by a theory of reading.

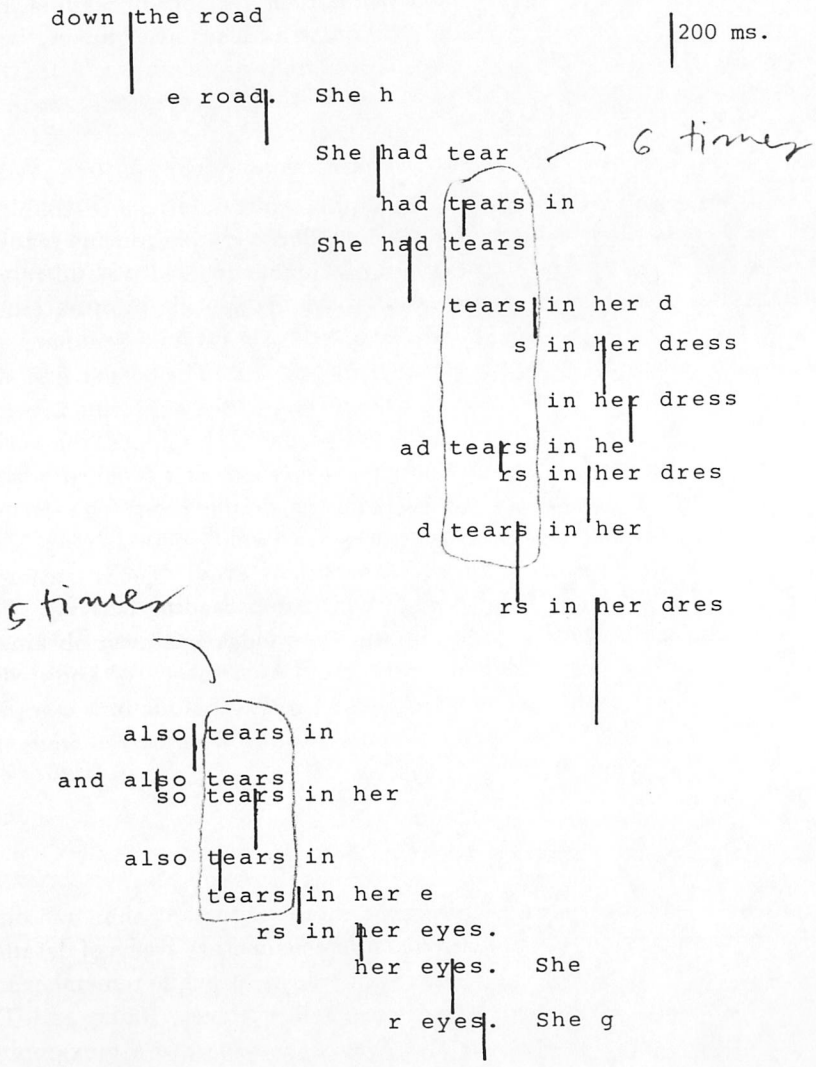


Figure 2: Sequential eye fixation locations of Buswell's subject H13 (1920). Vertical lines indicate fixation durations; saccade times are omitted. The functional visual field is assumed to span six letters to the left and eight to the right of fixation.

3. Reading may give way to mental states that are implicitly aimed at. Like the first reason, it does not form part of a theory of reading, but again a theory of action, reading becomes a means after having been a goal.

Reading for information

If we think of reading for information, exemplified by studying a textbook or cookbook, there is the implicit understanding that a simple routine reading pass will not suffice. Studying a text is a continuous interactive and repetitive issue, undertaken for the purpose of obtaining specific information or knowledge. Tinker (1965) provides the example of a chemist scanning the chemical formula CaO_2H_2 within a paragraph of text. The reader first fixates on the C, jumps to the H, next to the O, then to the subscript 2 in O_2, and finally to the 2 in H_2. Just before, the reader had skipped the entire word *hydroxide*; later on, the compound *amonium sulphate* (spelled with a single m!) gets only one fixation. This indicates that reading for information can be a highly selective process, where very small sections may draw considerable and repeated scrutiny and others none at all. The intention is to obtain the desired knowledge through continuous reading activity. The reason for stopping is simply that sufficient knowledge has been obtained, which was the prime goal. This stopping requires a theory of knowledge acquisition, not of reading. A theory of reading might include how complex information, containing unknown elements, can be or is extracted from the text. Stopping forms no part of this theory.

Reading for consultation

This is a situation where search becomes more important than reading. Reading for consultation takes place only when elementary items of detailed information are needed, which are not considered worthwhile remembering. Telephone numbers, train departure times, Yellow Pages, Radio and TV Times, card files, stock quotations and most database data are examples. All of these items refer to activities logically connected with them, such as calling somebody, leaving home, warning the plumber, watching a TV show, looking up an address, selling stock or checking stores. The intention in reading and scanning this kind of material is essentially to stop reading. If it does not stop, there is something fundamentally wrong either with the person who searches, or with the search tables.

5 Conclusion

It has been argued here that reading and writing constitute an even more abstract form of coding than spoken language. To the extent that reading requires considerable cognitive resources, it may be regarded as reflecting fundamentals of cognitive functioning. The extensive training time it demands, and its many counterintuitive perceptual properties shed doubt on this generality. The fundamental nature of the reading process has been strengthened by its relatively long history, which has provided the proper conditions for optimal functioning. The fact that reading performance is adversely affected by CRT-screen display should not come as a surprise once the specific conditions under which reading and writing evolved have been analysed. Wright's statement (*this volume*) that a theory of not-reading is needed is addressed with an analysis of reading situations. In many cases reading is not the highest priority activity; in fact, it rarely is. Reading is generally a means to attain a goal and is performed as long as it is deemed necessary or desirable. Consequently, in most cases 'not reading' cannot be properly described in terms of a theory of reading, but rather in a theory of actions. For these theories task-action protocols have been proposed (Norman, 1984; Epstein, 1986) and formally defined (Taylor, 1987*ab*) which can successfully address action transitions. Only one instance of discontinuity in reading can, and should be addressed by a theory of reading; it is the temporal asynchrony between perception and understanding. Yet, this case would be understood by most students of reading behaviour as 'reading', not as 'not reading'.

References

Bösser, T. (1988) A discussion of 'The chunking of skill and knowledge' by Paul S. Rosenbloom, John E. Laird and Allen Newell. In: B.A.G. Elsendoorn and H. Bouma (eds) *Working models of human perception*, London: Academic Press.

Bouwhuis, D.G. (1985) Psychologie van de informatieverwerking, *Informatie en Informatiebeleid, 11*, 48-58.

Bouwhuis, D.G. and Bouma, H. (1979) Visual recognition of three-letter words as derived from the recognition of the constituent letters, *Perception and Psychophysics, 25*, 12-22.

Buswell, G.T. (1920) An experimental study of the eye-voice span in reading, *Supplementary Educational Monographs, No. 17.*

Card, S.K., Moran, T.P. and Newell, A. (1980) Computer text editing: An information processing analysis of a routine technical skill, *Cognitive Psychology, 12*, 32-74.

Carmichael, L. and Dearborn, W.F. (1947) *Reading and visual fatigue.* Boston: Houghton Mifflin Company.

Cattell, J. McK. (1885) Über die Zeit der Erkennung und Benennung von Schriftzeichen, Bildern und Farben, *Philosophische Studien, 2*, 635-650.

Content, A., Kolinsky, R., Morais, J. and Bertelson, P. (1986) Phonetic segmentation in pre-readers: Effects of corrective information, *Journal of Experimental Child Psychology, 42*, 49-72.

Eco, U. (1983) *The name of the rose.* London: Martin Secker & Warburg Ltd. (*Il nome della rosa.* Sonzogno: Gruppo Editoriale Fabbri Bonpiani, 1980).

Epstein, W. (1986) Contrasting conceptions of perception and action, *Acta Psychologica, 63*, 103-115.

Gerbino, W. and Salmaso, D. (1987) The effect of amodal completion on visual matching, *Acta Psychologica, 65*, 25-46.

Gilhooly, K.J., Wood, M., Kinnear, P.R. and Green, C. (1988) Skill in map reading and memory for maps, *Quarterly Journal of Experimental Psychology, 40A*, 87-107.

Goldscheider, A. and Müller, R.F. (1893) Zur Physiologie und Pathologie des Lesens, *Zeitschrift für Klinische Medizin, 23*, 131-167.

Grashey, H. (1885) Über Aphasie und ihre Beziehungen zur Wahrnehmung, *Archiv für Psychiatrie und Nervenkrankheiten, 16*, 654-688.

Green, M.W. (1981) The construction and implementation of the cuneiform writing system, *Visible Language, XV*, 345-372.

Hinton, G. (1979) Some demonstrations of the effects of structural descriptions in mental imagery, *Cognitive Science, 3*, 231-250.

Hinton, G. (1987) The horizontal–vertical delusion, *Perception, 16*, 677-680.

Hofstadter, D.R. (1979) *Gödel, Escher, Bach: An eternal golden braid.* New York: Basic Books.

Huey, E.B. (1968) *The psychology and pedagogy of reading.* Cambridge, Mass.: M.I.T. Press. (First published in 1908).

Javal, E. (1878) Sur la physiologie de la lecture, *Annales d'Oculistique.*

Kolers, P.A. (1976) Reading a year later, *Journal of Experimental Psychology, Human Learning and Memory, 2*, 554-565.

Kolers, P.A. and Perkins, D.N. (1975) Spatial and ordinal components of form perception and literacy, *Cognitive Psychology, 7*, 228-267.

Leakey, R. and Lewin, R. (1977) *Origins.* London: Rainbird Publishing Group Ltd.

Leermakers, M.A.M. and Boschman, M.C. (1984) Eye movements, performance and visual comfort using VDTs, *IPO Annual Progress Report, 19*, 70-75.

Liberman, I.Y. (1973) Segmentation of the spoken word and reading acquisition, *Bulletin of the Orton Society, 23*, 65-77.

Marr, D. and Nishihara, H.K. (1978) Representation and recognition of the spatial organisation of three-dimensional shapes, *Proceedings of the Royal Society of London, B, 200*, 269-294.

Morais, J., Cluytens, M. and Alegria, J. (1984) Segmentation abilities of dyslexics and normal readers, *Perceptual and Motor Skills, 58*, 221-222.

Nes, F.L. van and Jacobs, J.C. (1981) The effect of contrast on letter and word recognition, *IPO Annual Progress Report, 16*, 72-80.

Norman, D.A. (1981) Categorization of action slips, *Psychological Review, 88*, 1-15.

Norman, D.A. (1984) Four stages of user activities. In: B. Shackel (ed.) *Interact 1984, First IFIP Conference on 'Human-Computer Interaction'*, London: Imperial College.

Powell, M.A. (1981) Three problems in the history of cuneiform writing: Origins, direction of script, literacy, *Visible Language, XV*, 419-440.

Rosenbloom, P.S., Laird, J.E. and Newell, A. (1988) The chunking of skill and knowledge. In: B.A.G. Elsendoorn and H. Bouma (eds) *Working models of human perception*, London: Academic Press.

Schiepers, C.W.J. (1976) *Global attributes in visual word recognition.* Doctoral thesis, University of Nijmegen.

Schmandt-Besserat, D. (1981) From tokens to tablets: A re-evaluation of the so-called 'Numerical tablets', *Visible Language, XV*, 321-344.

Taylor, M.M. (1987*a*) Layered protocols for computer-human dialogue: I. Principles, *International Journal of Man-Machine Studies*, in press.

Taylor, M.M. (1987*b*) Layered protocols for computer-human dialogue: II. Practical issues, *International Journal of Man-Machine Studies*, in press.

Thorndyke, P. and Stasz, C. (1980) Individual differences in procedures for knowledge acquisition from maps, *Cognitive Psychology, 12*, 137-175.

Tinker, M.A. (1964) *Legibility of print.* Ames: Iowa State University Press.

Tinker, M.A. (1965) *Bases for effective reading.* Minneapolis: University of Minnesota Press.

Tombaugh, J., Lickorish, A. and Wright, P. (1987) Multi-window displays for readers of lengthy texts, *International Journal of Man-Machine Studies, 26*, 597-615.

Valladas, H., Reyss, J.L., Joron, J.L., Valladas, G., Bar-Yosef, O. and Vandermeersch, B. (1988) Thermoluminescence dating of Mousterian 'Proto-Cro-Magnon' remains from Israel and the origin of modern man, *Nature, 331*, 614-616.

Whipple, G.M. (1910) The effect of practise upon the range of visual attention and of visual apprehension, *The Journal of Educational Psychology, 1*, 249-262.

Wright, P. (1988) The need for theories of NOT reading. In: B.A.G. Elsendoorn and H. Bouma (eds) *Working models of human perception*, London: Academic Press.

How Good a Bet is the Likelihood Principle?

Emanuel Leeuwenberg* and Frans Boselie*

1 Introduction

Visual perception is a remarkably powerful process. It merely begins from two-dimensional retinal patterns. Yet it enables us to see a three-dimensional object which turns around or moves to another position although the corresponding projections on the retina vary greatly. We interpret these projections or proximal stimuli as a composition of a single distal object and its motion. Distal objects may even have varying shapes and yet be classified into one category. Besides this amazing ability to pick up 'invariants', perception appears to predict the world veridically. Because of this property, we may run through woods or drive a car without tactual checks on our path. A driving licence depends largely on one's visual ability, not on other perceptual skills. One rarely asks the witness of an accident: 'Did you touch the cars?' Illusions do indeed occur, but seldom in everyday life. They are produced by scientists in poor laboratory conditions.

If we focus on the veridicality and predictive power of perception, we are almost forced to conclude that perception uses inferences and knowledge. On the other hand, if we focus on the ability to represent the world in terms of invariants, we are inclined to think that perception aims at simplicity. According to the first option, that specific perceptual organization is chosen which corresponds to the most likely distal object. According to the second option, the simplest pattern representation will be the preferred organization.

Before we consider the question of which option is the better one, an antecedent issue should be considered: These two options may simply be just two sides of the same coin. Such a view is defended by Mach (1906), and it indeed seems pertinent if one considers the following example, taken from Rock, (1983): If one suddenly hears a loud noise and, at the same time sees a bright flash of light, this pair of stimuli can be interpreted as being

* Psychologisch Laboratorium, Katholieke Universiteit Nijmegen, Postbus 9104, 6500 HE Nijmegen, The Netherlands.

WORKING MODELS OF HUMAN PERCEPTION
ISBN 0-12-238050-9