00 01 02 03 04 05 06 07 08 09 0A

Attention Design: Seven issues to consider

Sharon Wood, Richard Cox & Peter Cheng
Representation & Cognition Group
Department of Informatics
University of Sussex, Falmer, BN1 9QH
[S.Wood|richc|P.C.H.Cheng]@sussex.ac.uk

Abstract. In HCI research there is a body of work concerned with the development of systems capable of reasoning about users' attention and how this might be most effectively guided for specific applications. The design of systems capable of assessing user attention, evaluating the effectiveness of focus of attention and capturing, shifting or maintaining attention, would best be informed by the literatures on attention and performance, psychophysical factors affecting attention, task demands, and an understanding of how errors typically arise and are dealt with in meeting those demands. Current work seeks to establish the relationship between attention, the user's task, its situational context, and the user's understanding of that context and the expectations this generates. In doing this, we have found it helpful to distinguish between benign situations and errorful situations. Our aim is, through better understanding of attention and its impact on the user's ability to extract pertinent information in a timely manner, to extract design principles for supporting user activity.

Introduction

Designing systems that can monitor the degree and focus of a user's attention and then adapt the interface, dynamically modifying it in some way to manage the user's attention is a major challenge for research in HCI and cognitive engineering. This endeavour is attempting to create a new class of interfaces that are unlike existing systems, so the process of design must be more than an incremental extension of existing technologies. From studies of other design oriented disciplines and investigations of the nature of design (Simon, 1981; Goel & Pirroli, 1992) it has been established that successful design in such innovative situations needs strong scientific bases to underpin the enterprise: the exploration of a vast space of imaginable design possibilities must be constrained. Fortunately for the new area of attention design in HCI there are bases (in psychology and cognitive science) upon which to build. The aim of this paper is a preliminary foray in to some of that work. The theories and findings that may be relevant are presented as seven issues, in order to stimulate debate rather than attempt to (prematurely) provide a coherent framework.

The structure of this paper is as follows - under seven separate headings we address issues concerned with the nature and measurement of visual attention, the role of external representations, human error, human memory annd task variables. We conclude by considering the implications of research findings in these areas for the design of adaptive interfaces for managing the user's attention.

Issue 1: What is attention? Attention is a process of selection and selective processing, required because the brain has a limited information processing capacity (eg. Allport, 1993). In this paper we discuss visual attention rather than auditory attention. This distinction is necessary because the nature of attentional processes in these two sensory modalities appears to differ in some important respects. For example, broadly speaking humans demonstrate a greater capacity for subconsciously monitoring (and processing) unattended events in the auditory modality compared to the visual modality (eg. Cocktail party phenomenon, Cherry, 1957). The visual attentional system seems, in contrast, to be prone to *inattentional blindness* (see *Issue 2* below and Pylyshyn, 2003 for a review).

In some respects visual attention operates in a manner akin to a camera's zoom lens - ie. a 'lens' of variable focal-length that can alter the scope of visual attention between 'wide-angle' and 'telephoto'. However a fixed amount of processing resources are applied to any particular scene within the range. Thus processing can either be intensive and local - dedicated to a small part of a scene with reduced attention to the rest of the scene, or more evenly distributed across a wider area. Moreover, recent research suggests that visual attentional profiles can be shaped within a scene. Farah (1989) used a cueing paradigm based on cells within 5x5 grids describing 'H' or 'T' letter shapes. Results suggested that subjects could spread their attention voluntarily over particular shapes (scene sub-regions). Other research suggests that such 'shaped attentional regions' correspond to visual objects rather than randomly shaped areas, for example (for a review, see Pylyshyn, 2003).

The zoom lens metaphor is compelling but is only one of several that have been used as the basis for theorizing about the nature of visual attention. Other metaphors include: attention as a spotlight with a moving fixed size diameter focus (Treisman, 1986); attention as a glue that binds together features of things that are being processed (Cowan, 2001); attention as a 'coherence field': a viewer-centred nexus maintained in short term memory, stabilising links between a selected set of low level visual structures (proto-objects) (Rensink, 2000); attention as a bottleneck through which a limited amount of information can be filtered (Broadbent, 1958); attention as a limited capacity for information processing (Broadbent, 1971); attention as a multimodal process in which different processing modes (top-down, bottom-up) are implicated at different stages (Pashler, Johnstone & Ruthruff, 2001). The marked differences between some metaphors, and similarity of others, is indicative of the lack of thoroughgoing theoretical consensus in this area.

A further distinction made by visual attention researchers is between parallel systems for establishing "what" (the categorical identity of a stimulus) and "where" (spatial location). The two subsystems seem to interact in a highly complementary manner, provided that both are concerned with the same object. In that case they can be used concurrently without extra processing cost (e.g. Duncan, 1993).

Most visual attention laboratory research has tended to employ external 2D visual stimuli and has tended to ignore the individual's motivational state and goals. Visual attention can certainly be captured by strong stimuli in the environment (a non-inhibitable orienting response) or it can be directed under voluntary control (we are capable of 'paying attention'). A well-documented example of involuntary attentional capture is the 'weapons focus' effect reported in the forensic psychology literature. The presence of a weapon at a crime scene draws an eyewitness's attention away from other features of the scene such as the culprit's face (see Wells & Olson, 2003 for a recent review). Clearly "what" and "where" subsystems interact to rapidly change attentional focus.

There seems to be reasonable consensus for a conception of visual attention as a system that is capable of switching between such wide-angle and telephoto modes as a function of task demands. The question of what drives such switching processes is a central concern of this paper. In one sense, our goal in this paper is to explore the potential for inducing 'weapons-focus' type attentional capture effects under, for example, emergency conditions in contexts such as process control displays.

Overviews of this extensive area can be found in Underwood (1993) and (more recently) in Pashler, Johnstone & Ruthruff (2001). The latter reviewers conclude that the last 10 years of research suggests that attention is controlled by top-down, cognitively driven processes to a much greater extent than was believed even 10 years ago. There are complex interactions between voluntarily adopted mental 'sets' and the attention-capturing attributes of stimuli. Research is also tending to show that the effects of practice upon attention and performance are less pronounced than was hitherto believed. Designers need to comprehensively target both top-down (cognitively-driven) and bottom-up (stimulus-driven) processes when they design support systems or information displays systems.

Issue 2: How can attention be measured? The user's failure to apprehend pertinent information in a timely manner may in some circumstances have serious implications for the user's task whilst in others may simply slow activity down (Rensink, 2002). Designing systems to compensate and support users operating under the constraints imposed by limitations in attention involves understanding the basis of these limitations and being able to measure attention effectively.

Research into the psychophysical factors affecting attention (e.g. Pylyshyn and Storm, 1988; Kahneman, Triesman and Gibbs, 1992; Rensink, 2000; Hayhoe, 2003; Triesch et al, 2003) reveals that visual activity predominantly arises as a consequence of actively scanning the field of view in a task directed manner. The phenomenon of inattentional blindness (Hayhoe, 2003; Mack and Rock, 1998; Rensink, 2000) or inattentional amnesia (Rensink, 2000; Wolfe, 1999) demonstrates the selective nature of vision. Even though entities are clearly within view, if they are not central to the task in hand, they frequently remain unseen (Rensink, 2000). By visually pursuing the selection of information about those entities central to the current task, other entities in the visual scene are actively ignored, no matter how conspicuous they may seem to the non-task-oriented viewer (Simons and Chabris, 1999).

A related phenomenon of change blindness further demonstrates aspects of natural vision which result in failure to notice changes to entities in the visual scene when these take place during a saccadic eye movement (Rensink, 2002; Grimes, 1996). It appears that change can only be detected when the changing object is fixated (Rensink, 2002; Rensink, O'Reagan and Clark, 2000).

This phenomenon has also been demonstrated in a virtual reality setting during activities in which the changed feature is central to the task in hand (Hayhoe, 2003). Participants asked to pick up blocks (pink or blue), and to place these in a particular location according to colour, failed to notice when the selected (virtual) object changed colour between initial selection and final placement. Most often the object was placed in the location appropriate for its colour during initial selection, rather than for the colour to which it had changed. Hayhoe (2003) argues that this demonstrates the 'micro-structure' of vision: that fixation of an object is not sufficient for apprehension of all the visual information associated with it. It would appear that during initial selection of the object, participants pay attention to colour, whilst during subsequent fixations they appear to be concerned with location in guiding the object to its resting place (Ballard, Hayhoe and Pelz, 1995).

In a separate study, Triesch, Ballard, Hayhoe and Sullivan (2003) varied the nature of the task in relation to the object feature (this time, size) undergoing change. The number of changes noticed varied significantly between tasks, depending upon whether size of object was relevant to the task and, in particular, whether this attribute was relevant to carrying out the task both before *and* after the change took place. Contrastingly, a detailed analysis of direction of gaze during all activities revealed that patterns of gaze did *not* vary between tasks, despite the evidence that participants varied significantly in terms of the information they visually apprehended. Consequently, it appears the user's awareness of a given situation is highly selective and task-oriented and, in particular, it seems that direction of gaze alone is not a sufficient indicator of the precise nature of the information penetrating the user's awareness.

These limitations on inferring focus of attention from direction of gaze firstly have implications for the development of systems which seek to enhance user-modelling through identifying focus of attention. Approaches to supporting the user in their task through enhanced user-modelling have attempted to contextualise eye-movements/gaze through various additional sources of information. These include the users interactions with software and devices, their prior interests and patterns of activity, and even using information from the users on-line calendar. Current studies seek to investigate the potential these combined sources of information offer for identifying focus of attention more accurately than through gaze alone (Horvitz, Kadie, Paek and Havel, 2003).

Limitations on inferring focus of attention from direction of gaze also have implications for studies where measures of attention form part of the experimental design. If direction of gaze is not necessarily synonymous with focus of attention, studies will need to substantiate focus of attention through further evidence, for example the transfer of information within a given task context, whereby evidence of the deployment of knowledge indicates the actual focus of attention.

Issue 3. Salience and free rides? The contribution of Gestalt psychology to our understanding of how visual displays are perceived, what structures we appear to naturally see in images, is well known. The various Gestalt principles concerning visual forms that are particularly salient can be found in any introductory cognitive psychology text and one might assume that this is a good place to begin considering interface design for attention. However, there have been substantial advances in our understanding, beyond that classical work, on the nature of representations and the design of representations to support inference and problem solving in cognitive science (e.g., Glasgow, Narayanan & Chandrasekaran, 1995; Anderson, Cheng, & Haarslev, 2000; Blackwell, Marriot, & Shimojima, 2004). For the attention designer there are valuable insights concerning how graphical representations may be used to support cognition. For instance the properties of diagrams that enable quick perceptual inferences to be used rather than laborious logical reasoning, so called free rides, are now reasonably understood and can be used to guide interface design. There is also a growing body of knowledge of how particular representations are well suited to different classes of information and types of inference. Such knowledge could be exploited in the design of interfaces that deliberately plays with the salience of the information for different goals and tasks.

Issue 4. Cognitive load ?, complexity ?, expertise ? Reason (1990) proposed a generic error modelling system (GEMS) framework for conceptualising the range of error types that humans exhibit when engaged in complex tasks such as controlling real-time, dynamic, multivariate systems. Three basic error types are proposed - skill-based slips, rule-based mistakes and knowledge-based mistakes.

The difference between the three levels of processing has been nicely encapsulated in an example provided by Felciano (1995). Opening a door is usually a skill-based activity performed unconsciously without the need to 'think about' it. If, however, the door refuses to open then we switch to rule-based processing - perhaps the door is locked, perhaps it pushes instead of pulls, etc. Following a limited set of rule-based but unsuccessful attempts, we then move to the knowledge-based level of processing. To diagnose why the door won't open we begin to consider whether the door is being held shut by someone on the other side, has had its lock changed, has become jammed, *etc*.

Rules and knowledge-based processing are both conscious processes. Knowledge-based processing requires a mental model of the problem, and analysis of more abstract relations between structure and function. Hence it requires more cognitive effort than rules-based processing. Shifting from rule to knowledge-based levels can result, *inter alia*, from the detection of countersigns. These are present when inputs indicate that a more general rule is inapplicable (Reason, 1990). In the door example, the fact that unlocking the door and pushing/pulling fail is a countersign and a signal to shift to knowledge-based levels of processing.

The attention and change blindness literatures have major implications for several aspects of Reason's model, particularly those concerned with the need to switch from skill-based to rule-based levels when error states in a system are detected. For example, change blindness may result in failure to detect 'countersign' information. Reason (1990) argues that the difficulty of detecting countersigns is further compounded by 'information overload' of the cognitive system by a high volume of information. Change blindness phenomena and top-down/bottom-up interactions in attention provide a useful basis for operationalising Reason's rather vague concept of 'information overload' in his GEMS model.

Issue 5. Working memory The nature of working memory is a major and active area of research in psychology and cognitive science and should be central in the thinking about designing for attention. An excellent overview of what is known, and what is far from certain, about working memory is Miyake and Shahs (1999) volume. They posed eight questions about the form, mechanisms, functions and implementation of working memory to leading researchers in the area. The answers to the one concerning the relation of working memory to attention and consciousness are especially pertinent here. Some consider working memory, attention and consciousness as largely synonymous constructs. Others makes distinctions between them or considering them as overlapping or being in subset relations. What is clear is that understanding and accounting for working memory is important. It is widely accepted that humans exhibit a limited working memory capacity. Millers (1956) classic paper put the number of chunks that can be stored in working memory at seven plus or minus two chunks. A chunk consists of items of information that are strongly associated with each other and weakly associated with items of information not in the chunk. A chunk may be considered, in general terms, as a concept. Since Millers paper, the actual capacity of working memory has been challenged, particularly in circumstances that are not ideal. Cowan's (2001) review of findings puts the realistic capacity at four chunks. The implications for the attention designer are various. For instance, if an estimate of the number of chunks required to do a task can be made, this may indicate the cognitive load placed on the user of an interface and hence may indicate whether attention is substantially deployed. One might deliberately attempt to structure the task or interface to manage the number of chunks so that there are not more chunks than can reasonably be attended to. There are many explanations for the limited capacity of working memory, including limited supply of activation to spread over concepts, finite processing speed, decay of chunks, interference among concepts and others. Rather than providing details of the different theories what appears to be useful for those interested in designing for attention is a model that operationally reflects what is known about the relation of working memory and attention. Working memory can be considered to be that part of long term memory representations that is currently active and readily available for processing. The working memory part of long term memory is made active by the context of the current task which provides cues, pieces of information, to retrieve related chunks. In turn only part of the content of working memory will be directly under the, so-called, focus of attention and engaged with immediate task activity. The activation of chunks in working memory will fade and no longer be available to attention, unless they are reactivated by rehearsal or the activation of related chunks brings them back. The implications for designing for attention are various. How will relevant chunks brought into working memory so they are at least available? If there is insufficient situational or contextual information to cue the concepts they will not available for processing. What is to be done to keep the concepts active during long tasks where they are likely to decay and not be available or require deliberated effort to be recalled?

Issue 6. Are core attentional processes open to change (immutable)? Evidence is provided by dual-task research in the attention and performance literature. An example of a dual-task paradigm might be requiring someone to verbally indicate whether a tone is low or high in pitch while concurrently indicating (via keyboard key selections) which letter appears on a computer screen. The effects of dual-task interference diminish with practice but not to the point where performance levels are identical to those observed when the tasks are performed singly (see Pashler et al., 2001). Moreover, training effects differ in their nature from those predicted by traditional ideas of 'automaticity'. The challenges for designers are a.) to minimise the extent to which tasks interfere with each other and b.) to try and minimise the errors due to between-task interference.

Issue 7. Are attentional processes different for different tasks (e.g. vigilance vs diagnosis)? Monitoring the state of variables in a complex system, for example, entails stimulus-driven (bottom-up) processing to a greater extent than a task like diagnosis does. In tasks such as *monitoring*, subjects are typically required to keep track of displays and gauges. If they are well-designed, such displays make error states salient to the point that bottom-up, stimulus-driven capture of attention is inevitable. Recent evidence (e.g. Pashler et al., 2001) suggests that novel objects capture attention to a greater degree than other types of manipulation such as abrupt changes in luminance. Such findings are commensurate with the research cited earlier which shows that recognisable objects can prime attentional 'regions' within a scene and allow the categorical identity and spatial location attentional subsystems to be used concurrently and without extra processing cost.

In contrast, in a problem-solving task such as *diagnosing* the cause of a fault, attention interacts with mode-of-execution. The process is primarily cognitively-driven. The person doing the diagnosis must first notice an anomalous state and would typically attempt a rule-based solution in the first instance. This could be followed by a sequence of cycling to and from a knowledge-based solution level if necessary. Reason (1990) states that a key feature of the GEMS model is that "human beings are strongly biased to search for and find a prepackaged solution at the rule-based level before resorting to the far more effortful knowledge-based level, even where the latter is demanded at the outset." (p. 65). To return to our 'door' example, this might correspond to a perseveration with the 'door is locked' rule based processing level perhaps taking the form of an assumption that the (correct) key tried was the wrong one.

Discussion

We have posed seven questions highlighting issues which might be usefully inform a theoretical framework for designing for attention.

Interest in the design of systems which take account of the user's focus of attention might naturally take the perspective that such systems should be capable of reasoning about users' attention and how this might be most effectively guided for specific applications, through being capable of assessing user attention, evaluating the effectiveness of focus of attention and gaining, shifting or maintaining attention. This presents a particular view on how design might best address issues regarding the psychophysical limitations of attention.

The change blindness (enactive perception) and visual attention findings are not yet integrated sufficiently to allow prescriptive design recommendations to be made. However there are some indications that information display designers should consider using novel and recognisable visual objects for attention capture rather than abstract shapes or information channels such as colour. Designers might also consider using displays for slowly changing data that decouple the display from the information it depicts, by using periodic step changes in the display in place of real-time tracking, as these are likely to be more readily noticed. It is also clear from this literature that the very nature of attention makes it a difficult phenomenon to gauge and study and that experimental design should supplement eye gaze data with other forms of evidence for substantiating focus of attention. Finally, designs decisions made with the aim of managing the user's attention should build on existing, well-established principles for the design of effective representations (i.e. exploiting graphical free rides, judicious selection of representational systems in terms of expressiveness, etc).

References

A. Allport, (1993) Visual attention. In M. I. Posner (ed.) *Foundations of cognitive science*. Cambridge, MA: MIT Press.

M. Anderson, P. C.-H. Cheng, & V. Haarslev, *Theory and Application of Diagrams: First International Conference, Diagrams 2000*. Berlin: Springer, 2000.

A. Blackwell, K. Marriot, & A. Shimojima (Eds.), *Proceedings of the 3rd International Conference on Diagrams, 2004*. Berlin: Springer-Verlag, 2004.

D. Ballard, M. Hayhoe & J. Pelz, Memory representations in natural tasks, *Journal of Cognitive Neuroscience*, 7, 1995, pp. 66-80.

D. E. Broadbent, *Perception and communication.* London: Pergamon, 1958.

D.E. Broadbent, *Decision and stress.* London: Academic Press, 1971.

C. Cherry, (1957) *On human communication.* (2nd ed.) Cambridge, MA: MIT Press.

N. Cowan, The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behaviour and Brain Sciences.* 2001 24(1):87-114.

J. Duncan, (1993) Coordination of what and where in visual attention. *Perception*, **22(11)**, 1261–1270.

M.J. Farah, (1989) Mechanisms of imagery-perception interaction. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 203-211.

R.M. Felciano, (1995) Human Error: Designing for Error in Medical Information Systems. Web URL
http://www.smi.stanford.edu/people/felciano/research/humanerror/
Accessed 13 Dec 2004.

J. Glasgow, N. H. Narayanan, & B. Chandrasekaran, *Diagrammatic Reasoning: Cognitive and Computational Perspectives*. Menlo Park, CA: AAAI Press, 1995.

J. Grimes, On the failure to detect changes in scenes across saccades, in: K.A. Atkins, ed., *Perception, Vancouver Studies in Cognitive Science*, 5, OUP, Oxford, 1996.

V. Goel & P. Pirolli, The structure of design problem spaces, *Cognitive Science*, 16, 395-429, 1992.

M.M. Hayhoe, What guides attentional selection in natural environments? in: *Abstract Proceedings of Fifth Workshop on Active Vision*, University of Sussex, September 22, 2003.

E. Horvitz, C. Kadie, T. Paek and D. Havel, Models of Attention in Computing and Communication: From principles to Applications. *Communications of the ACM*, Vol 46(3), 52-9, 2003.

D. Kahneman, A. Triesman, & B. Gibbs, The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175-219, 1992.

A. Mack and I. Rock, *Inattentional blindness*. Cambridge, MA: MIT Press, 1998.

G. A. Miller, The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, 1956, 63, 81-97.

A. Miyake, & P. Shah, (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge University Press.

H. Pashler, J.C. Johnstone, & E. Ruthruff, Attention and performance. *Annual Review of Psychology*, 52, 629-51, 2001.

Z. Pylyshyn, (2003) *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press.

Z.W. Pylyshyn, & R.W. Storm, Tracking multiple independent tartets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3, 179-197, 1988.

J. Reason, *Human error*. Cambridge, UK: Cambridge University Press, 1990.

R.A. Rensink, Seeing, Sensing and Scrutinizing. *Vision Research*, 40, 1469-1487, 2000.

R.A. Rensink, The dynamic representation of scenes. *Visual Cognition*, 7, 17-42, 2000.

R.A. Rensink, Internal vs. external information in visual perception. *Proceedings of the 2nd ACM International Symposium on Smart Graphics*, 63-70, 2002.

R. A. Rensink, J.K. O'Regan, & J.J. Clark, On the failure to detect changes in scenes across brief interruptions, *Visual Cognition*, 7,127-145, 2000.

H. A. Simon, *Sciences of the Artificial*, 2nd ed. Cambridge, MA: MIT Press, 1981.

D.J. Simons & C.F. Chabris, Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, 28, 1059-1074, 1999.

A. Treisman, Features and objects in visual processing. *Scientific American*, 254, 114-124, 1986.

J. Triesch, D.H. Ballard, M.M. Hayhoe and B.T. Sullivan, What you see is what you need. it Journal of Vision, 3, 86-94, 2003.

G. Underwood (ed). *The Psychology of Attention (Vol 1)*. Aldershot: Elgar. 1993.

J.M. Woolfe, Inattentional amnesia. In V. Coltheart (Ed) *Fleeting Memories*, Cambridge, MA: MIT Press, 1999.

G.L. Wells, & E.A. Olson, (2003) Eyewitness testimony. *Annual Review of Psychology*, **54**, 277–295.