Attention aware systems: Theories, applications, and research agenda

Claudia Roda [a,*], Julie Thomas [b]

[a] *Computer Science Department, The American University of Paris, 31, Avenue Bosquet, 75007 Paris, France*
[b] *International Communication Department, The American University of Paris, 31, Avenue Bosquet, 75007 Paris, France*

Abstract

Human perceptual and cognitive abilities are limited resources. Attention is the mechanism used to allocate such resources in the most effective way. Current technologies, in addition to allowing fast access to information and people, should be designed to support human attentional processes on which they impose further strain. This paper analyses the issues related to the design of systems capable of such support: attention aware systems. We introduce the research aimed at understanding and modelling human attentional processes, including perceptual and cognitive processes as studied in cognitive psychology, as well as rhetorical, aesthetic, and social aspects related to attentional mechanisms. We analyse current approaches to the design of attention aware systems along three major features: detection of user's current attentional state, detection and evaluation of possible alternative attentional states, strategies for focus switch or maintenance. Finally, we discuss the most promising research direction for the development of systems capable of supporting human attentional mechanisms.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: Attention aware systems; Adaptive systems; Interruptions; Cognitive systems

1. Introduction

Consider the following scenario.

Ann has come home from basketball practice. Her computer welcomes her and notifies her of a message from her basketball trainer and one from her mum, it also informs her

---

* Corresponding author. Tel.: +33 1 40 62 06 50.
E-mail address: croda@aup.fr (C. Roda).

that the homework for the "operating systems" class is due tomorrow. On the screen there is a note about email messages from two friends and about a special offer lasting two days at her favourite bookshop. When she gets to the Web site of the operating systems course, the due assignment comes up on screen, the only other items currently visible are links immediately related to the assignment. As Ann starts working at her assignment several emails arrive but she only receives notification for an email arrived much earlier from the course teacher about a change of venue for next meeting, with a "that's fine" she confirms that the information can be stored in her diary. After concentrating for some time on a few lines of the assignment, Ann appears frustrated. The computer displays the summary of a preparation exercise that Ann has not yet completed, and a link to a further explanation of the subject of the assignment. In the meantime Ann's brother walks into the room and asks her whether she could pick him up from his football game tomorrow afternoon. After a few exchanges with him, Ann turns to the computer and asks for her diary which is immediately displayed, along with notification of messages from two friends. As Ann turns back to the computer, after arrangements have been made with her brother, the course Web site is visible and the diary has faded. Ann completes the preparation exercise and reads the text but she still cannot advance on the assignment. The computer tells her about two lectures related to the assignment that she may want to review, and also informs her that her teacher and two students from her class are online. Two emails from the bookshop about the special offer are not immediately notified to Ann but a message coming from a business partner of her employer interrupts her work.

The system Ann is using is capable of reasoning about her current goals and attentional focus. It presents her with the information needed to achieve her goals, it hides information currently irrelevant, it follows her in her choices of focus shift, and it interrupts her when necessary. This system must be aware, amongst other things, of Ann's current, planned, and possible activities, of the load and importance of these activities, of her learning preferences, of her social networks, and of her most likely reactions to a set of perceptual inputs.

The system Ann is using seems particularly desirable because current technology enables access, communication, creation, and use of more information than our information-processing capacity allows us to handle. This problem is normally referred to as *information overload*. Heylighen (2004, p.1) argues that "ephemeralization, [i.e.] an accelerating increase in the efficiency of all material, energetic and informational processes [forces individuals] to consider more information and opportunities than they can effectively process. [and] It leads to anxiety, stress, alienation, and potentially dangerous errors of judgement". The concept of information overload is quite interesting especially if we reflect on the fact that we have lived, and learned to cope with information overload well before the advent of technology. What we see, hear, and generally perceive around us (in the physical world) exceeds, probably by several orders of magnitude, what we are actually capable of processing. We therefore have developed cognitive mechanisms that enable us to select the information we process. Considering visual stimuli for example, Rensink (2000) offers a fascinating and well-documented discussion about how little we "see" compared to the stimuli on the retina. Chun and Wolfe explain that "What you see is determined by what you attend to. At any given time, the environment presents far more perceptual information than can be effectively processed. [...] Complexity and information overload characterize almost every visual environment [...] To cope with this potential overload, the brain is equipped with a variety of attentional mechanisms". (Chun & Wolfe, 2001, p. 273)

These mechanisms collectively form what we call *attention* and allow us to cope with the massive amount of stimuli we receive in the physical world. Why then are we now so overwhelmed by the information made available by technology? Certainly, this does not significantly increase the input to our perceptual system as compared to the real world stimuli. In certain situations, in fact, the introduction of technology reduces rather than increasing the perceptual input. For example the perceptual input in most types of device-mediated communication is lower than in face-to-face communication.

Rather, the problem resides in the format of this information that – being highly formalised, conventional, and often lacking context – requires high level cognitive mechanisms in order to be processed. Whilst we could have developed through evolution particular skills to process certain complex scenes very fast, we certainly do not have this ability for instance with text. Some recent studies in neuropsychology have demonstrated that we can "rapidly categorize highly variable natural scenes outside the focus of attention" (Fei Fei, VanRullen, Koch, & Perona, 2002, p. 9601). This results in our ability to process – e.g. search, look at, and describe – a huge landscape faster than a page of text. There is a possibility that, as Heylighen suggests, humans will eventually evolve to be able to process formal/conventional information much faster. However, information systems may also be designed so that they better support our current attentional capabilities.

This paper analyses the issues related to the design of systems capable of such support: attention aware systems.

We start (Section 2) by reviewing research in cognitive psychology aimed at understanding and modelling human attentional processes. Whilst this research concentrates on perceptual and cognitive processes, other processes related to attentional mechanisms are reviewed in Section 3, these include rhetorical, aesthetic, and social aspects. In Section 4, we analyse current approaches to the design of attention aware systems along three major features: detection of current user's attentional state, detection and evaluation of possible alternative attentional states, strategies for presentation to the user of these alternative states. Finally, in Section 5, we discuss the most promising research directions for the design of attention aware systems.

2. Human attention: perceptual and cognitive processes

Cognitive psychologists and, more recently, neuropsychologists have extensively studied human attention. Within these disciplines attentional mechanisms have been considered as guided by *perceptual* and *cognitive* processes. The research results obtained so far and their connections and implications are far too many to be discussed in a single paper; however, in this section we attempt to give an overview of those themes that are particularly relevant to the design of attention aware systems.

2.1. Attention as selection

There is no agreed upon definition of attention. However, most researchers refer to attention as the set of processes enabling and guiding the selection of incoming perceptual information. For example, Posner (1982) stresses the fact that attention is selective, has limited capacity, is related to both reactive and deliberative processes, and it is associated with both inhibitory and facilitating effects. Driver (2001, p. 53) defines "Research on attention [as] concerned with selective processing of incoming sensory information". Lavie

and Tsal (1994, p. 183) see ''Selection of information [as] the primary concern of attention research''. Chun and Wolfe (2001, p. 273) propose that ''First, attention can be used to select behaviorally relevant information and/or to ignore the irrelevant or interfering information. [...] Second, attention can modulate or enhance this selected information according to the state and goals of the perceiver. With attention, the perceivers are more than passive receivers of information. They become active seekers and processors of information, able to interact intelligently with their environment''. Three aspects of attention seem to be commonly recognised as fundamental: selection, awareness, and control (Baddeley & Weiskrantz, 1993; Parasuraman & Davis, 1984).

Studies in experimental and cognitive psychology essentially confirm our naïve understanding of attention control. Roughly one can say that attention can either be controlled voluntarily by the subject, or it can be captured by some external event. The former type of control mechanism is referred to as endogenous, top-down, or goal driven attention (Arvidson, 2003; Posner, 1980; Yantis, 1998). An example of endogenous attentional mechanism is the attention you are paying to this page as you are reading. The latter type of mechanism is referred to as exogenous, bottom-up, or stimulus-driven. An example of this would be the attention shift from your reading due to a sudden noise. These two mechanisms have different characteristics, ''endogenous attention is voluntary, effortful, and has a slow (sustained) time course; [...] exogenous attention draws attention automatically and has a rapid, transient time course'' (Chun & Wolfe, 2001, p. 279).

Research on visual attention has also reported that exogenous mechanisms may have different degrees of power so that certain stimuli become basically impossible to ignore (e.g. sudden luminance changes), whilst others are more controlled by volition. This latter type of stimuli has highlighted the interaction between exogenous and endogenous mechanisms. In fact the endogenous mechanisms in place (e.g. what one is looking for in a visual field, and how this search is performed at the voluntary level) seem to determine whether one will automatically be able to ignore certain stimuli. Or, as Chun and Wolfe put it, ''the guidance of attention is determined by interactions between the bottom-up input and top-down perceptual set'' (Chun & Wolfe, 2001, p. 280).

In order to support these two (top-down, and bottom-up) mechanisms it is necessary to understand how they work, and more importantly how they interact. This is what we will analyse in the next two sections. First we briefly present the classic theories of attention that, by observing reactions to perceptual stimuli, try to infer bottom-up processes and possibly top-down ones. Then we turn to more general cognitive theories that explain bottom-up processes in the frame of more general models of cognitive, top-down, processes.

2.2. From perception to cognition

Allocating attention to a set of stimuli is seen as selecting those stimuli for processing. But how, and when, are those stimuli selected? In other words, is it possible to break down the selection process into sub-processes, or stages? The first answer to this question came from Broadbent's *early perceptual selection* theory (Broadbent, 1958) which proposed that selection takes place in two stages: a pre-attentive one, and an attentive one. First, a parallel (several stimuli may be processed at the same time) preattentive stage encodes simple physical properties of incoming stimuli. The recognition of the simple physical properties allows one to filter out all the irrelevant (unattended) stimuli and to pass on only selected stimuli to the second stage. This second, serial, attentive stage is capable of only limited

processing and it encodes more abstract properties such as semantics of the attended stimuli. Later experiments demonstrated that unattended stimuli received higher level of processing than those predicted by Broadbent's theory (e.g. Lewis (1970) & MacKay (1973) reported experiments demonstrating semantic processing of unattended stimuli). As a response to these findings some authors (see for example Deutsch & Deutsch, 1963; Duncan, 1980; Norman, 1969) proposed a *late selection theory* where all stimuli are analysed but only pertinent stimuli are selected for awareness and memorisation. Other studies (Neisser, 1967; Pashler, 1998; Treisman, 1960, 1969) have suggested *modified early selection theories* where non-attended stimuli are not completely filtered out but their contribution to the attentive stage is somehow limited. Treisman (1960, 1969), for example, proposed that unattended stimuli might make it to the second serial, attentive stage, but in an attenuated mode. These attenuated stimuli would then be elaborated on only if they had a low "threshold" for identification. Such a low threshold may be due either to the stimuli's particular significance in the current environment or personal experience.

Modified early selection theories are particularly interesting because they relate attentional mechanisms to personal experience and to the current environment (which includes immediately preceding stimuli). Partial information and context, for example, seem to be at the root of *interference effects*. Interference effects are delays in the processing of stimuli due to unwanted stimuli called *distractors*. Interference effects help explain why we do not always manage to keep attention on a target stimulus nor can we consistently avoid distractors. How much interference a distractor will generate appears to be a function of personal experience and the environment. Rafal and Henik (1994), for instance, observe that distractors may delay the processing of stimuli especially when they are temporally near and conceptually related to the target stimulus. Classic examples of interference are *negative priming* (Tipper, 1985) and the *Stroop effect* (Stroop, 1935). Negative priming is an effect by which "it is more difficult to select a stimulus, belonging to a given category, for the control of action, if that same category of object was actively ignored on the preceding trial" (Allport, 1989, p.659; quoted in Arvidson, 2003, p. 114). The *Stroop effect* (Stroop, 1935) is a visual inhibition effect that occurs when a presented word indicates a different colour than the actual colour of the font used to spell it.

From the point of view of system design, these studies allow us to better understand how one could evaluate the likelihood that a stimulus be recognised by a user – and the effort that this recognition may require – in the context of the specific users and their activities. Also these theories hint that it is possible to provide users with information in a format such that it will be noticed only if relevant to the user's current purpose. For example, Hyrskykari, Majaranta, and Räihä (2003) report that, in their system supporting the reading of foreign language documents, messages providing translations often went unnoticed if not needed by the reader.

Although the debate on early and late selection has been central to research in attention and has provided many insights into the processes underlying attention, modern theories of attention try to integrate the early and late selection approaches. Lavie and Tsal (1994), for example, integrate the late selection assumption that perceptual processing takes place automatically and cannot be deliberately limited – as suggested by Deutsch and Deutsch (1963) – and the early selection assumption that perceptual capacity is limited – suggested by Broadbent (1958). On this basis they stipulate that "irrelevant information will be excluded from processing only if the prioritised relevant processing exhausts all of the available capacity" (Lavie & Tsal, 1994, p. 185). This implies that if the available perceptual

capacity is not exhausted by the main stimuli, distractors will also be processed. In this account the locus of selection depends on the load of the incoming stimuli. Lavie (2000) also proposes a measure of perceptual load that, she argues, should be based on the number of stimuli and the perceptual operations necessary to elaborate them. A further contribution to the understanding of preattentive and attentive processes comes from Treisman's *feature integration theory* (Treisman, 1998; Treisman & Gelade, 1980) developed from her studies based on visual modality. She proposes that parallel preattentive processes simultaneously extract essential features of the visual stimuli (e.g. colour and orientation). The integration of the different features is then obtained applying serial attention to the location of each item. It is this serial process that binds all the elements together and produces the appropriate multidimensional percepts of objects with particular colours, orientations, etc.

In relation to system design, this theory tells us that certain searches on the visual field can be made very efficient if they only involve basic features, i.e. features that can be extracted preattentively (e.g. finding the only red object in a very complex scene). Many researchers in cognitive psychology have tried to identify such features. A critical review of this research is proposed in Wolfe (2001). One result is that "preattentive basic features are coarsely coded. [...] In orientation, the preattentive categories seem to be "steep", "shallow", "left" and "right". In size, the categories are probably merely "big" and "small"; in depth, "near" and "far" and so forth" (Wolfe, 2001, pp. 7–8). Pre-attentive visual processes seem to work at a higher abstraction level than early vision, i.e. there has been some sort of categorisation of what we actually see, such as the distinction between large and small. On the basis of this observation Wolfe (2001, pp. 7–8) proposes that pre-attentive vision comes after early vision (which we could loosely define as the physical layer of vision). In pre-attentive vision a scene is parsed in pre-attentive objects as bundles of basic features. It is only once attention is directed to a pre-attentive object that the relationships between basic features are bound to form the objects we recognise. In a successive phase, that Wolfe calls post-attentive, the object is recognised by linking it to a representation in memory that allows its identification. In a detailed review of research in visual attention Cave and Bichot (1999, p. 204) cover issues such as "location versus object-based selection, whether attention moves in an analog fashion, [...] whether attention works by facilitating targets or inhibiting distractors, and whether multiple non-contiguous regions of space can be selected simultaneously". More recently, VanRullen and his colleagues have conducted a set of experiments demonstrating that the simple dichotomy parallel/preattentive versus serial/attentive used to describe visual discrimination tasks may need to be replaced by a more complex model that takes into account several variables (VanRullen, Reddy, & Koch, 2004).

We believe that the results of research in visual discriminations will be a precious input to designers of attentive visual display and will allow them to design systems capable of guiding user attention in a flexible and context related manner. More generally, insights into human attention mechanisms will support the evaluation of which stimuli (in which context, at which times, and in which format) will be more easily detected, or ignored, by the user.

2.3. From cognition to perception

The classic theories of attention described in the previous section all recognise the importance of top-down processes in attention. However, they often focus on bottom-up processes. This may be due to the fact that in laboratory experiments target subjects

are requested to perform a certain task (e.g. a search in a visual space) and therefore the top-down processes are necessarily restricted to the imposed goal. Another stream of research tries to explain attentional processes within the larger frame of general cognitive processes (such as learning, formulating and pursuing intentions, achieving goals). These theories differ from the ones presented in the previous section in that they attempt to infer possible structures, architectures, and components, for the processes guiding human actions and cognitive activities. The models they propose delineate the role of attention in cognition and therefore serve both (1) to explain how supporting attentional processes one may support general cognitive processes, and (2) to understand how general cognitive processes may affect attentional processes. The models provide interesting insights into the factors that may influence the cognitive load imposed by certain external stimuli and by self-selected, or imposed tasks.

First we introduce two connectionist models aiming at explaining how humans may learn about the world. The third model focuses on how humans may control task switching. They have been selected as representative models of high-level cognitive processes; however, other models have also been extensively used (see for example ACT-R Research Group, 2002–2005) or simply proposed (Hewett, 2000).

2.3.1. Connectionist models of learning and attention: attention and intentionality

Grossberg's adaptive resonance theory (Grossberg, 1976a, 1976b, 1999) explains how attentional processes are integrated into more general processes guiding human behaviour. The two main questions addressed by this work are, (1) how we (humans) may keep learning about our changing environment, and (2) how this may lead to conscious experience. Grossberg proposes that the processes guiding this learning "include the learning of top-down expectations, the matching of these expectations against bottom-up data, the focusing of attention upon the expected clusters of information, and the development of resonant states between bottom-up and top-down processes as they reach an attentive consensus between what is expected and what is there in the outside world. [...] The process whereby the top-down expectation selectively amplifies some features while suppressing others helps to "focus attention" upon information that matches our momentary expectations. This focusing process helps to filter out the flood of sensory signals that would otherwise overwhelm us and to prevent them from destabilizing our previously learned memories. Learned top-down expectations [...] focusing attention [allow us to learn] quickly and stably without catastrophically forgetting [...] past knowledge" (Grossberg, 1999, pp. 1 & 10). Whenever a matching between the sensorial input and the top-down expectations is achieved the system enters a resonance state. This is due to the input activating the expectation, which in turn selects the input, and so on. Adaptive resonance theory maintains that only these resonant states reach consciousness.

One important aspect of ART, with respect to our enquiry about attention aware systems is its modelling of intentionality as guiding attention. In the ART framework intentions guide attention in two ways. First, intentions reflect expectations of events that may (or may not) occur. Second, intentions help monitoring sequences of events (often called plans by other authors) that should take place in order to satisfy behavioural goals. Grossberg explains that "Implicit in the concept of intentionality is the idea that we can get ready to experience an expected event so that when it finally occurs we can react to it more quickly and vigorously, and until it occurs, we are able to ignore other, less desired, events" (Grossberg, 1999, p. 12).

This has two immediate consequences for system design. First, since users' attention will be focussed on information that matches their momentary expectations, understanding users' intentions (both in the sense of behavioural goals, and in the sense of events likelihood) is essential in supporting cognitive attentional processes. Second, one way of directing and maintaining user focus is to act at the level of intention. There seem to be a very close relation between ART's top down expectations and what in cognitive psychology literature is called *priming*. Arvidson (2003, pp. 116–117) defines the *prime* as "any stimulus or condition that facilitates (facilitation effect) or interferes (interference effect) the processing of the target stimulus. [...] The prime is not what the subject is directed toward by the experimenter, but affects the target stimulus of attention anyway". Further research may be able to shed more light on if and how exactly priming effects are related to top down expectations and consequently on how they may be exploited in attention aware systems.

2.3.2. Connectionist models of learning and attention: learned attention and inattention

By concentrating on some apparently illogical human use of certain cues, such as the fact that sometimes relevant cues are ignored and irrelevant ones are attended to, Kruschke (2001) proposes a different model for high level cognitive processes. He explains the above phenomena in terms of learned attention (highlighting) and learned inattention (conditioned blocking). Kruschke stipulates that "both attentional shifting and associative learning are driven by the rational goal of rapid error reduction" (Kruschke, 2003, p. 171) so that, in order to reduce errors with respect to already learned cues, we learn to attend to certain cues and to ignore others. In particular, following the model proposed by Mackintosh (1975), attention to cues that have been learned to be relevant increases, whilst attention to cues that have been learned to be irrelevant decreases.

Practical examples of these types of behaviour would be that if one is used to regularly checking the email notification icon he/she may look for that icon even if not needed or even if another icon indicates that the user is working off-line; if one has learned to ignore certain links on a web page as not useful, he/she is likely to ignore them even when they will become useful.

It is particularly relevant for the purposes of our study that the order in which cues are presented is the determining factor in highlighting and conditioned blocking effects. Experiments by Kruschke and others (see references in Kruschke, 2001, 2003) demonstrate that, once people have learned that a cue A corresponds to an effect X, highlighting and conditioned blocking effects may be associated to the pairing of cue A with other cues resulting in effect X, or Y different from X. These observations are relevant to the design of interfaces. For instance, the fact that a user has learned a given mode A to perform a task could retard his learning of a quasimode[1] AB to achieve the same task (perhaps in a more efficient form, or in alternative version).

Systems capable of modelling users' learned attentions and inattentions could provide better customisation of presentation and attention aware system may exploit learned attentions and inattentions effects to facilitate user focus.

---

[1] "Using the Caps Lock key to type uppercase letters and holding the shift key to the same effect are significantly different. The first case establishes a mode; the second case [establishes a quasimode]" (Raskin, 2000, p. 55). Experiments have demonstrated that quasimode induce less errors than modes.

2.3.3. Executive-process interactive control (EPIC) architecture: attentional processes in multi-tasking

Classic theories and experiments in cognitive psychology have concentrated on the study of how humans may perform individual tasks, these being motor or cognitive tasks. This section reviews the work done to define the EPIC architecture (Kieras, Meyer, Ballas, & Lauber, 2000), a computation model of human multi-task performance, and other theories following the EPIC principles (Rubinstein, Meyer, & Evans, 2001). The stated goal of the EPIC project is to simulate human–system interaction in a test-bed modelling human performance and capable of rapidly assessing the quality of an interface.

Although the classic example of human multi-tasking, proposed in much of the academic and general literature, is that of driving while speaking on the telephone, it seems to us that multi-tasking has a much more general and frequent occurrence in human life. Even if we disregard the multi-tasking related to unconscious processes, such as breathing, digesting, and swallowing, while doing one's homework, multi-tasking at the level of conscious activity, such as drinking a cup of coffee while doing the homework are definitely normal occurrences in human life. Descriptions of multi-tasking are also obviously related to the granularity of the description of the task. The "doing one's homework" task can be divided into smaller tasks – such as writing, reading, reaching for a book – which can be considered as executed in sequence or in parallel. The task of writing itself includes at least two significantly different tasks, one at the cognitive level – such as one's task of selecting the appropriate meaning, syntax, and lexicon – and one at the spatial-motor level of physically writing the selected words. It is clear that an understanding of multi-tasking and the attentional processes that may guide the allocation of resources for the performance of such tasks is essential if we want to be able to support computer mediated human activity.

The debate among theorists is still open on whether "executive control processes" supervising the selection and scheduling of motor and cognitive tasks exist, and whether these processes may be distinguishable from the basic processes enabling the execution of each task. However, for our goals, this is not the most pressing question. What this theoretical research may contribute to our more pragmatic approach is to supply us with a set of parameters that may enable/disable, facilitate/hinder human multi-tasking. Our question is not how does multi-tasking take place, but rather do we know enough about the characteristics of processes controlling multi-tasking so that we can facilitate these processes? It is with this question in mind that we approach the research reported in this section.

Rubinstein and his colleagues (Rubinstein et al., 2001) propose that two distinguishable sets of processes control the execution of consecutive tasks: executive control processes, and task processes. Task processes control performance of the individual tasks and include "three principal stages, *stimulus identification*, *response selection*, and *movement production*" (Rubinstein et al., 2001, p. 770, author's emphasis). Executive control processes control task switching which "entails at least two functionally distinct stages of executive control, goal shifting and rule activation, which are separable from the basic perceptual-motor and cognitive processes used for performing individual tasks"(Rubinstein et al., 2001, p. 763). Following the model proposed by Rogers and Monsell (1995) the authors also indicate how these processes include endogenous control that in a top-down manner prepares for the next task, and exogenous control that is triggered by the onset of the next task stimulus and completes the preparation for the task. According to the authors, executive control processes, and task processes, although distinct, may have interleaved execution, in particular the *rule*

activation process, which sets the rules that will be used by the *response selection process*, may under certain conditions be triggered by a stimulus recognition. For example, say that the *goal shifting* process has selected the goal of performing a colour-discrimination task, then the *rule activation* process may wait to recognise a colour-stimulus before loading the colour-discrimination task rules (these rules may for example specify that if the colour is red than a certain key should be pressed). The authors assume that "if a switch occurs from one task to another, there is a pause between the end of stimulus identification and the beginning of response selection for the current task [...]. This pause is used by an executive control process whose operations enable the subsequent response selection stage to proceed correctly" (Rubinstein et al., 2001, p. 770). In support of this hypothesis, the authors report Jersild's (1927) experiments which demonstrate not only that increasing the complexity of a task increases the completion time and introducing task switching increases the mean completion time, but also that these two factors interact reliably. In fact the difference in performance time for task repetition and task alternation increases with the complexity of the tasks.

The above findings have at least two implications for system design. First there is always a cost associated with switching attention from one task to another; second, this cost is related to the complexity of the tasks involved. In order to design systems that guide user attention efficiently, the evaluation of the costs/benefits of attention switches must take into account the complexity of the task. This opens two questions: (1) what are the parameters that define task complexity? And (2) given some evaluation of task complexity, how can switching cost be evaluated? It seems likely that both general and user-related parameters will contribute to the evaluation of task complexity (intuitively we can define the level of complexity of a task both "in a general sense" and "for a specific person"). The function mapping task complexity to switching costs seems also likely to be related to individual cognitive characteristics. Cognitive models of the user could therefore play an essential role in attention aware systems.

Other experiments by Spector and Biederman (Spector & Biederman, 1976) reported by Rubinstein and his colleagues (2001) demonstrated that task switching times may be significantly reduced if visual cues are provided about the task to be performed next. In the settings of this experiment participants had to go through a list of numbers and alternatively add 3, and subtract 3. Performance improved when a +3, −3 cue was provided next to each number. Rubinstein et al. (2001) hypothesise that task cueing facilitates the executive control process devoted to choosing which task should be executed next. Although these experiments dealt with repetitiveness within task switching it seems likely that they can be generalised to task (and therefore attention) switching in general. If this was the case, systems capable of providing cues about the task to be performed next would reduce cognitive load for the users. Resuming the execution of an interrupted task is a particular case of task switching. Therefore, providing cues about the context of the interrupted work would support attentional processes in task resumption. For example, in a word processor, task cues may provide information about which part of a document was last edited, and about the context in which that editing took place (e.g., after opening a certain web page and reading a certain email.)

Another parameter affecting performance in task switching is the duration and regularity of the *response stimulus interval* (RSI). RSI is the time between the response to a stimulus and the presentation of the subsequent stimulus. Rubinstein and his colleagues hypothesise that endogenous processes initiate preparing for the next task only if the

RSI is predictable, otherwise, these operations are postponed until the onset of the next task stimulus. This hypothesis was based on the results of some experiments by Rogers and Monsell (1995) and by Allport et al. (Allport, Styles, & Hsieh, 1994) demonstrating that, under certain conditions, increasing the length of RSI decreases switching times costs only if the RSI is constant. These observations imply that increasing the time between attention switches does not per se reduce users' cognitive load. A system that aimed at supporting users' attentional processes by selectively delaying interruptions could therefore be beneficial only if the interruption times can be predicted by the user.

Finally, one more result showing how attention switches may affect performance is reported in Rogers and Monsell (1995). The authors demonstrate that in task alternation, stimuli that are irrelevant for the current task may induce higher response times (RT) and task switching time if they are related to the currently irrelevant task. Or, as the authors phrase it: "stimuli evoke associated task-sets" (Rogers & Monsell, 1995, p. 207). This implies that stimuli related to familiar and recent foci may cause greater disruption to the user's current activity (see also (Rafal & Henik, 1994) as reported in Section 2.2 of this paper). Therefore, an assessment of the context of the user activity should be added to the evaluation of the costs/benefits of interruptions. For example, if a user is extracting some data from an email to insert it in a spreadsheet, an email notification is more likely to disrupt the work of the user than a pop up reporting on issues unrelated to recent activities.

To summarise, the work reported above has the following implications for the design of attention aware systems. (1) There is always a cost associated in switching attention from a task to another one. (2) Higher task complexity implies higher costs for attention switch. (3) The cost of switching attention from one task to another can be reduced if cues are provided about the task to be performed next (at least in situations of task alternation). (4) Increasing the time between attention switches does not reduce the cost of the switch unless this time is constant (i.e. attention switches happen regularly), or at least predictable. (5) Stimuli related to familiar (and recent) tasks are more likely to act as distractors for current tasks.

3. Human attention: rhetorical, aesthetic and social aspects

There are many fundamental aspects of attention, of the mechanisms that humans use to manage it, and of their application in digital environments that require much further exploration. As will become obvious from the description below, this exploration would greatly benefit from a more interdisciplinary approach to the design of attention aware systems.

Earlier we have discussed the role of reactive and deliberative processes regulating attention control. A complete account of attentional mechanisms, however, must take into consideration two more factors that interplay with reactive and deliberative processes. These are social and aesthetic factors controlling attention. The reason why such factors should be considered separately (although interplaying with other processes) is that they may be guided by different motors. Reactive processes are largely guided by one's unconscious reaction to sensorial stimulus. Deliberative processes are largely guided by one's intention to achieve certain goals. Social factors influencing attention are guided by one's sense of relation to others. Aesthetic aspects affecting attention are guided by one's experience-based and culturally conditioned reflection on some external stimulus. Some examples may further clarify the different nature of these. In a room full of people, the

conversation of someone who is considered ''important'' (the company's CEO, a movie star) may attract more attention than the conversation of others. In the same room, if one notices that everyone is looking at the ceiling then his/her attentional focus will most likely switch to the ceiling. These behaviours are not due simply to a reactive or deliberative process, but rather represent a combination of these with the social aspect. When listening to a piece of music, one can be completely engrossed, focused on the experience. In this case it is the aesthetic aspect that prevails. The link between social and aesthetic aspects and their interplay with attentional mechanisms are, of course, culturally determined, as in Bourdieu's (1990) definition of 'habitus', ''the system of structured, structuring dispositions'', produced by conditioning, ''which is constituted in practice and is always oriented towards practical functions.'' (Bourdieu, 1990, p. 52) What is 'recognised' by the user as related to a pre-conditioned, cultural 'disposition' is more likely to appear worthy of attention. This factor should be considered not simply in terms of choice of design which satisfies these cultural and/or aesthetic expectations, but also in terms of whether information is conveyed and cued through sound, image, or verbal text and whether or how successfully the source of information or message is communicated to the user.

In *What is Thought?* Eric Baum (2004) suggests a computational model of thought processes in which he relates human thought structures directly to the environmental world structure, and attributes the development of human thinking to the success of evolutionary DNA in privileging the ability to select relevant possibilities quickly and efficiently from a vast pool of information and sensory input. If this is the case, we are programmed by environmental, physiological and cultural factors to 'pay attention' to certain kinds of information as a preface and key to appropriate selection. Possible mechanisms for activating this 'attention conditioning' are themselves in turn selected from as tools for effective communication discourse. Attracting attention is to persuade that the stimuli, information, or object is worth consideration for possible selection – for further attention, as it were.

There is an 'umbrella' discipline which deals with the art and science of persuading to attend, selecting appropriate mechanisms for arousing attention, and identifying how such attention processes may themselves be used to persuade, and that is the ancient discipline of rhetoric. LaGrandeur (2003), for example, has provided a background and justification for analysing the digital image in terms of classical Western rhetoric, and Fagerjord (2003) has made the point that the new and different abilities of computer technology can also be used to form new rhetorical techniques. From the time of Aristotle, rhetoricians have taught that different situations, applications, and audiences require different mechanisms for arousing and maintaining attention, and that knowing the audience and being aware of audience predispositions and expectations – in other words, being able to model the user successfully – is the most important prerequisite for successful communication. Just as a speaker might choose different vocabulary, sentence structure, turns of phrase, familiar examples (commonplaces), and frame the argument differently depending on the context and subject of the speech and whether an audience of students or an audience of professional business people were being addressed, so a system designed to target user attention must also consider these factors and adjust the attention mechanisms accordingly. For example, the discrete display of translation aids utilised by the iDict system (Hyrskykari, Majaranta, Aaltonen, & Räihä, 2000) takes into account the user need for minimal distraction; in a collaborative work environment where immediate attention to new information might be vital to the successful completion of a project, mechanisms

to attract user attention would of necessity have to weight interruption/distraction factors differently.

If we consider the Ann scenario, we might imagine that Ann herself embodies at least two different 'audiences', depending on whether she is acting in a private or a professional capacity. Her e-mail notification could thus be framed to appeal to these differences – e-mail from family and friends being differentiated from e-mail from her teacher or her employer's business partner through the manner of its presentation. Further, depending on the context, social aspects of attention might dictate whether Ann wishes to consult the professional e-mail before the e-mail from her friend (or the reverse!). In this case, difference in display, informed by Ann's cultural and aesthetic preferences (her 'habitus'), would aid her in selecting the appropriate focus of attention quickly and efficiently. Since Ann's immediate goal is to complete her assignment, when she has trouble finishing the preliminary exercise after a certain period of time has passed, she could be offered a clear explanation in simple language of the task, using a familiar example which she can relate to as illustration. Alternatively, if she finishes the preliminary exercise very rapidly, she could be offered a supplementary commentary exploring the subject in more depth (perhaps with the addition of tables, figures, statistics), so that her attention is maintained and she is encouraged to complete the actual assignment.

Such consideration of the role of rhetorical practice in attention processes is obviously not confined to language or to use and choice of image in terms of quantity and kind of visual display (colour, medium, movement), but also includes manipulation of strategies of sound, and user involvement in terms of real or displaced physical motion. As Kress and Van Leeuwen point out (2001, p. 30) "the success of 'rhetoric', of attracting attention and persuading, depends on 'the ability to select the discourses which are to be 'in play' on a particular occasion, in a particular 'text'". Kress and Van Leeuwen also suggest that one important element in terms of multi-modality is the issue of reader/viewer/user control and movement – different forms of control and movement become operational elements of the rhetoric/ attention mechanisms of different modal realisations (Kress & Van Leeuwen, 2001, pp. 132–133).

Recent developments in attentive user interfaces (for example (Vertegaal, Chen, Shell, & Mamuji, 2006)) and ubiquitous computing confirm that these aspects of rhetoric continue to be as crucial away from screen-based interfaces as on screen.

Elsewhere (Roda & Thomas, 2005) we have discussed the classic HCI debate over transparent versus proactive/adaptive interfaces and how a focus on user attention resolves this dichotomy. However, this new focus also brings a new question: what do we call interface? What is the user really paying attention to? We believe that we should call *interface* the collection of all the layers of representations that a human being will have to "go through" to reach "meaning". When using digital systems, users often have to negotiate at least three layers of interface. First there is the "outer application" level interface. This is the interface of the software (nowadays often a client software) that the user is running. For example, we call "outer application" level interfaces the interface of a Web browser, an email client, or a word processing application. Second there is the "inner application" level interface. These are software "objects" that can be seen (or run) within the outer application. Examples of inner application interfaces are Web pages, email messages, and word processing documents. In fact there may be several other layers of interface between the "outer application" and the "inner application" layers. For example a Web based email client imposes one more level of application interface to the user (in between

the browser and the email message). Finally, there is the content level interface (e.g. the main text of Web pages or email messages). It is dangerous conceptualising medium and rhetoric as consisting of different parts, no matter how more or less converging or diverging they appear in specific cases. Whilst it is convenient (and indeed necessary), for the sake of analysis and creation, to identify these parts and attempt to define each part's contributing characteristics to the whole effect, in doing so we must be constantly aware that we risk (1) privileging the contribution of one part at the expense of another and (2) forgetting that the user/audience does not respond to 'parts', but to their successful interplay.

In designing our interfaces we should remember that there is always more than one interface, and the interface produced and demanded by the appropriate presentation of the content – that is, the selection of language, image, sound, layout, links, possible inter-action and control – to direct user/audience attention is an integral and important aspect of designing for attention. This selection/presentation of content is of course always subject to the tools available for the creation of the technical interface, but in terms of rhetoric, in terms of persuasion and attention, the technical interface and the content interface appear indivisible. The interplay between the content interface and the technical interface is vital to the successful exploitation of attention processes – and just as the content interface must correspond to and attract attention to the appropriate message, so the technical interface should adapt as far as possible to further the attention/selection profile of that message. The discipline of interaction design seems well placed to address this aspect of attention aware systems.

4. Current approaches to attention aware systems

In Section 1, we have discussed how attention aware systems address the problem of information overload. However, since this problem is so ubiquitous, we should explore which environments/application domains should be prioritised in our enquiry about attention aware systems. In order to assign this priority we can identify those domains where improved management of attentional resources may have a significant impact on the user experience and/or performance. These include domains where: (1) attentional switches are very often solicited, or (2) where the users' lack of experience with the environment makes it harder for them to select the appropriate attentional focus, or (3) where an inappropriate selection of attentional focus may cause serious damage to the system, its users, or third parties, or (4) where the very reason for the system to exist is to attract the user attention. Systems relying highly on multi-user interaction such as virtual communities, and systems supporting cooperative work are examples of environments where attentional switches are often solicited. Online educational systems are examples of environments where the lack of knowledge and experience of users with the subject at hand makes it harder for them to select the appropriate attentional focus and may easily cause a loss of focus. Life critical systems are examples of environments where an inappropriate selection of attentional focus may cause serious damage to the system, its users, or third parties. Finally, performing a task or achieving a goal may not be the main target of the user who instead may turn to digital artefacts for their symbolic or affective value, entertainment, or pleasure in general – see for example Lowgren's arguments for Interactive Design versus classic HCI (Löwgren, 2002). These are the domains where capturing and maintaining user attention may be the ultimate goal of the system.

This section reviews the work done so far in attention aware systems. It should be noted, however, that attention has not often been prioritised as a specific subject of research in HCI (with some notable exceptions including the *Attentional User Interface* project at Microsoft research (Horvitz, Kadie, Paek, & Hovel, 2003)). As a consequence, much of the work relevant to the development of attention aware systems appears in the context of other research frames. This is especially the case as attention processes are related to, and necessary for, the successful accomplishment of many diverse activities. In the literature, attention aware systems have also been referred to as *Attentive User Interfaces* (Vertegaal, 2003); however, we prefer using the former name as it stresses the fact that issues related to attention are relevant to the design of the system as a whole rather than limited to the interface.

In order to analyse the issues related to the design of attention aware systems, we will focus on three aspects of attention management. (1) Detection of current user's attentional state. The system needs to establish what are the user's goals and current tasks, where is the user's attention focussed, and what is happening in the environment. (2) Detection and evaluation of possible alternative attentional state. The system establishes whether alternative foci are available, how important they may be for the user, and the cost effectiveness of possible focus switches. (3) Strategies for presentation of alternative states to the user (or maintenance of current focus). The system defines the strategies best suited to present the user with alternative foci. These strategies will depend on the relevance of the new foci for the user and his current attentional state.

These three aspects of attention management are very much inter-related and their separation is only meant to help this presentation. In general, supporting users' attentional processes will require, amongst others, knowledge of users' focus, goals, and current environment. This is to say that we can analyse whether attention is allocated efficiently only with respect to a goal and the environment the user is immersed in.[2]

4.1. Detecting current attentional state

Detection of current attentional state involves collecting information about users' current focus of attention, their current goals, and some relevant aspects of users' current environment. The mechanisms for detection of user attention that have been most often employed are based on the observation of sensory cues of users' current activity and of the environment; however others, non-sensory based, mechanisms also need to be employed to form a complete picture of the user's attentional state.

4.1.1. Sensory based mechanisms for the detection of attentional states

Amongst sensory-based mechanisms, *gaze tracking* is the prevalent approach. It has been applied to the detection of current attentional focus of users performing various tasks such as: participating in a virtual meeting (Vertegaal, 1999), reading an electronic document written in a foreign language (Hyrskykari et al., 2000), or pointing a cursor (Zhai, 2003). One of the major difficulties related to employing gaze-tracking in detecting attention is that, although it is generally recognised that gaze direction is a good indicator of attentional

---

[2] Whether this is the case also in situations of activities that are not "task-oriented" can be debated but we will not delve into this subject in this paper.

focus, it cannot be taken as a certain indicator. Recent cognitive psychology results in change blindness ("failure to see large changes that normally would be noticed easily" (Simons & Rensink, 2005, p. 16)) support the hypothesis that attention is symmetrically distributed around the fixation point (Simons & Rensink, 2005). However, Rayner observes that one may shift the point of visual attention without necessarily moving the eyes (Rayner, 1995). For this reason, gaze tracking is sometimes integrated with other mechanisms for the collection of sensory cues such as *gesture tracking* (Hinckley, Pierce, Sinclair, & Horvitz, 2000). Other examples of collection of multiple sensory cues are the system proposed by Stiefelhagen that tracks *head pose* and acoustic information about who is speaking in a meeting (Stiefelhagen, 2002), or the systems for spoken-language developed as part of the Attentional User Interface project (Horvitz et al., 2003). This type of environmental information may help to detect both current focus and possible alternative foci of attention for the user. These include both information already available to the user, (e.g. the information displayed on a screen on which the user is not currently focussing, people in the same room as the user), and information that is available to the system but not to the user (e.g. incoming emails, the content of a database, news feeds, online status of acquaintances).

One aspect of attentional state detection that has been largely explored using sensory cues is user's interruptibility. Based on the observation that human beings can very efficiently, and in presence of a very small number of cues, evaluate other's interruptibility, Hudson et al. (2003) explore "whether, and how, robust sensor-based predictions of interruptibility might be constructed, which sensors might be most useful to such predictions, and how simple such sensors might be" (Hudson et al., 2003, p. 257). They come to the conclusion that "robust results should be attainable from practical, relatively easy to implement sensors" (Hudson et al., 2003, p. 263) and that "speech detectors are the most promising sensors" (Hudson et al., 2003, p. 264). More recently, experiments have been conducted using sophisticated physiological cues. Chen and Vertegaal (2004) propose to "measure mental load using heart rate variability (HRV) signals, and motor activity using electroencephalogram (EEG) analysis" (Chen & Vertegaal, 2004, p. 1513). This information allows them to "distinguish between 4 attentional states of the user: at rest, moving, thinking and busy" (Chen & Vertegaal, 2004, p. 1513).

4.1.2. Non-sensory based mechanisms for detection of attentional states

As observed by Horvitz and his colleagues (Horvitz et al., 2003) detection of attention allocation may also take into account non-sensory cues including: users' scheduled activities (e.g. using online calendars), users' interaction with software and devices, and information about the users and their patterns of activity and attention. In any case, even when employing mechanisms capable of taking into account all these cues, a certain level of uncertainty about users' focus will always remain. For this reason, dealing with uncertainty remains an important aspect of users' attention detection. A set of tools included in the Interruption Workbench (Horvitz & Apacible, 2003) implement a promising approach allowing the capture of events in the user's environment that are later used to build statistical models of user's interruptibility in various situations. Users are asked to assign a dollar value – how much they would be willing to pay – to avoid being interrupted by a certain event in a certain situation. The model implemented, in a way, integrates top-down relationships between user's goals and attention. In fact, since the user knows what his goal may be when in a certain situation for which he would pay a certain amount to avoid disruption, the statistical model subsumes top-down relationships.

In general, the relationship between user's goals and user's focus has not been largely explored in current research. Some authors concentrate on high level user's goals, such as not wanting to be disturbed, or wanting to be informed about certain events (e.g. McCrickard & Chewar, 2003). However, as discussed in the first part of this paper, cognitive psychology has demonstrated that current, low level tasks and goals (such as completing one's assignment, or verifying one's availability on a certain date) are fundamental drivers of attention. As a consequence more specific mechanisms are needed to relate current and alternative foci to current user's tasks and goals, and to assess such tasks and goals. In order to perform this assessment, user's tasks and goals can either be declared by the user (e.g. Roda, Angehrn, Nabeth, & Razmerita, 2003), be implicitly defined by the system or the situation in which it is used, or be dynamically evaluated on the basis of the user activity. In this last case, task-modelling tools (e.g. Bailey, Adamczyk, Chang, & Chilson, 2006; Maulsby, 1997) may be used to infer possible user's tasks from observed actions. Methodologies for focus determination similar to those described above may also be used to collect information helping to infer the current tasks and goals. User's tasks and goals may be detected at various levels of details and form the reference against which it is possible to evaluate the appropriateness of the current focus.

4.2. Determining and evaluating possible alternative FOCI

Given the users' current attentional state, attention aware systems should evaluate if alternative foci should be presented to the user, and in which format. In the introductory scenario, for example, the system has been able to propose to Ann a short summary of a preparation exercise, and later, information about the online status of her teacher. In order to display such behaviour the system should be able to evaluate the appropriateness of Ann's current focus with respect to alternative possible foci given her goal of completing the exercise. The evaluation of relevant alternative foci should take into consideration the interplay between top-down and bottom-up processes in human attention. Selected changes in the environment initiate bottom-up processes that can be pruned by top-down processes guided by goals. A system integrating the two processes has been proposed by Hill in his work on the synthetic helicopter pilot (Hill, 1999).

4.2.1. Determining alternative foci

The need for a focus switch may result from several conditions including: the availability of a ''better'' focus for the current user's goal, a temporary or durable change in the user's goal, and a temporary or durable change in the environment. In the sample scenario above the alternative foci on the exercise summary and on the teacher's online status have been determined as a ''better value'' foci for achieving Ann's current goal. The request of Ann's brother instead has generated a set of possible alternative foci (including Ann's diary) due to a change in the current user's (Ann's) focus and environment and the consequent temporary change of goal. Note the difference between this external event, which was observed by both Ann and the system, and the event corresponding to the arrival of the message coming from a business partner of Ann's employer. In the latter case, the event was only observed by the system that may consider the message as a possible alternative focus. Research in attention aware systems has mostly concentrated on this latter aspect of focus shift: the system interrupting the user to notify him/her of events of which he/she could not be aware of otherwise (e.g. the arrival of an email, the activity

of a collaborative colleague, the changed value of a stock ticker). We believe that attention aware systems could fulfil two other functions. First, they could follow up on focus shifts initiated by the user such as the display of Ann's diary in the sample scenario. Second, they could dynamically update information presentation so that, as information becomes more relevant, it also becomes more noticeable. For example, one could think that a link to the exercise summary already existed on Ann's screen but, as the exercise became more relevant, its presentation was changed so that it became more noticeable. This system behaviour would effectively deal with phenomena such as learned inattention discussed earlier in the paper.

4.2.2. Evaluating alternative foci

As alternative foci are detected the system should evaluate whether they should be presented to the user and in which format. The latter issue (format of presentation), which is closely related to the former, will be discussed in the next section.

Deciding whether to present an alternative focus amounts to either (1) establishing that the user's goal has changed (and consequently the focus should be adjusted) or (2) establishing that the user should be interrupted in her current activity to be presented either with task related information or with information related to other tasks or goals. Human beings deal with interruption constantly in their every day activities and interruption management has been studied both in cognitive psychology and in management science. Seshadri and Shapira (2001), for example, propose a study of how managers may attend to a main long-term process whilst allocating short-term attentional resources to secondary but pressing tasks. They propose that managers, depending on their personal inclinations and their environment, apply attention allocation rules, and they study the efficiency of these rules.

Interruptions may have both positive and negative effects on the primary tasks. On the one hand interruptions are necessary to bring to one's attention events or information that may be useful for the primary task or for some higher-level goal. It has been argued that, in the case of simple primary tasks, interruptions may even facilitate task performance (Speier, Vessey, & Valacich, 2003). On the other hand, as widely reported in the literature, interruptions may generate stress (Bailey, Konstan, & Carlis, 2001; Zijlstra, Roe, Leonova, & Krediet, 1999) and hinder in various ways the performance of the primary task – see for example (Franke, Daniels, & McFarlane, 2002; McFarlane & Latorella, 2002; Nagata, 2003; Speier et al., 2003). As emphasised by McFarlane and Latorella (2002) the implementation of any semiautonomous or user multitasking technology entails interrupting the user from other activities they are performing. In their seminal paper these authors review theoretical and experimental research on human interruption management and propose strategies for the design of systems capable of supporting this management process. They propose four design solutions to coordinate user interruptions: "immediate, negotiated, mediated, and scheduled. Interruptions can be delivered at the soonest possible moment (immediate), or support can be given for the person to explicitly control when they will handle the interruption (negotiation). Another solution has an autonomous broker dynamically decide when best to interrupt the user (mediated), or to always hold all interruptions and deliver them at a prearranged time (scheduled)" (McFarlane & Latorella, 2002, p. 5). On the basis of a set of experiments comparing these interruption methodologies, McFarlane (2002) concludes that in most situations negotiation is the best choice.

We have seen earlier that collaborative, multi-user environments are amongst those that would most benefit from the development of attention aware systems. In these environments the activities and attention allocation of one user may strongly influence the performance of other users. This implies that, whilst a given focus may be optimal for the performance of one user, it may be sub-optimal for the performance of the group. As a consequence, in these situations strategies optimising individual attention allocation may not be sufficient and the evaluation of the relevance of alternative foci must take into account also the state of other users. Although most of the work on the evaluation of the cost/benefits of interruptions has been done taking the point of view of the user being interrupted, some analysis takes into account also the cost/benefit to the interrupter, and the joint cost/benefit (Hudson, Christensen, Kellogg, & Erickson, 2002; O'Conaill & Frohlich, 1995).

4.3. Presenting alternative foci or maintaining the current focus

As alternative foci are identified and evaluated, strategies for their presentation to the user must be defined. Information presentation is an essential aspect in the design of attention aware systems. Depending on how information is presented it may impact on the user activity at various levels (see also discussion in Section 3). For example, it may go completely unnoticed, it may smoothly integrate with the user's current task, or it may capture the user's attention and cause a temporary or durable focus switch. As noted by McCrickard and his colleagues (McCrickard, Catrambone, Chewar, & Stasko, 2003) in their work on notification systems, ''Some notification systems are designed to attract attention and compel other activities, thus by their very nature they must interrupt users from some primary task and minimize delay in attending to a notification. However, many systems are intended to preserve as much primary task attention and performance as possible. Understanding how and when to best accomplish both is important to this field'' (McCrickard et al., 2003, p. 549). They propose to measure the effects of visual notification with respect to (1) users' interruption caused by the reallocation of attention from a primary task to a notification, (2) users' reaction to a specific secondary information cue while performing a primary task, and (3) users' comprehension of information presented in secondary displays over a period of time. In McCrickard and Chewar (2003) the same parameters, *interruption*, *reaction*, and *comprehension*, together with *satisfaction*, are seen as general user goals that can be used both to measure the costs/benefits of attention switches, and to guide the system design. So that, for example, small sized in-place animation can be defined as best suited for goals of minimal attention reallocation (*low interruption*), immediate response (*high reaction*) and small knowledge gain (*low comprehension*). Although McCrickard and his colleagues concentrate on classic notification systems (i.e. systems that allow the user to monitor information related to secondary activities, see McCrickard, Czerwinski, & Bartram, 2003) their work promises to be applicable to more general attention aware systems that manage several types of information and evaluate if, when, and how to make it available to the user. Notification systems have been studied in a wide variety of application domains including messaging systems (Cutrell, Czerwinski, & Horvitz, 2001; Czerwinski, Cutrell, & Horvitz, 2000; Horvitz et al., 2003), alerting in military operations (Obermayer & Nugent, 2000), and shared document annotation (Brush, Bargeron, Gupta, & Grudin, 2001).

At least three aspects of information presentation have a significant impact on attention allocation, the presentation style (including modality), the timing of presentation, and the

amount of content presented. The latter may range from a notification of information availability (e.g. a flashing icon indicating the presence of email) to complex awareness mechanisms (e.g. awareness display in a distributed collaborative system), to a complete switch of context (e.g. opening of a new window with a new application). Presentation style, timing, and content are obviously related, both because certain styles are better suited for certain contents, and because the effects achievable (possibly in terms of interruption, reaction, comprehension, and satisfaction) will often depend on how well the new information is integrated in the context of the current activity. For example, Carroll and his colleagues (Carroll, Neale, Isenhour, Rosson, & McCrickard, 2003), focussing on the support of collaborative activities, make a compelling case for the integration of awareness information within the context of user's current activity.

4.3.1. Content selection

To our knowledge, few authors have directly addressed the problem of the adaptation of the message content to the attentional state of the user. However, several authors working at the READY project have presented encouraging results in this direction. READY is a natural language interface aiming at dynamically adapting to the user's time pressure, and working memory limitations. Dynamic Bayesian networks and influence diagrams are used "for modelling the user's resource limitations and making decisions about the system's behaviour" (Jameson, Schafer, Weis, Berthold, & Weyrath, 1999, p. 81). The two prototypes – one supplying instructions for car repair (Jameson et al., 1999, p. 81), and one for making a phone call in an airport (Bohnenberger, Brandherm, Grossmann-Hutter, Heckmann, & Wittig, 2002) – serve to explore methodologies for assessing users' resource limitation on the basis of their speech, and consequently bundling instructions in appropriately long sequences. Although, as the authors indicate, these prototypes are still experimental, we believe that they highlight issues that are extremely relevant to attention aware systems.

4.3.2. Modality selection

One aspect of attention that has been the subject of much research in psychology and neurology is visual attention. Visual attention is particularly relevant to attention aware systems since the current predominant modality for computer-to-human communication is visual. Bartram, Ware, and Calvert (2003, p. 515) note that "Key to a notification technique is how easily the notification is detected and identified" and, basing their work on findings in cognitive psychology, they propose the use of *moticons* (icons with motions) as an effective visual technique for information rich displays that minimise distraction.

A wide variety of visualisation systems and techniques address, more or less explicitly, attentional processes in the visual modality. Toet (2006) offers an extensive review of adaptive techniques for visual information presentation.

Modalities other than visual have been studied in interfaces for intelligent environments and ubiquitous computing where several channels are often employed in order to communicate with the user – see Abowd, Mynatt, and Rodden (2002) for a review. Whilst attentional processes in multimodal interaction have been less studied than visual-only processes both by psychologists and HCI researchers, some advances in this direction have been reported. Bearne, Jones, and Sapsford-Francis (1994, p. 105), for example, propose "some initial guidelines for the design of usable multimedia systems" and base their advice on psychological theories of multimodal interaction. They observe that auditory attention,

as well as visual attention, can be focused. Then they discuss how these different attention modalities may be integrated in a system by exploiting the human ability to combine certain types of activities (e.g. seeing and hearing), whilst avoiding interference between the different types of interaction. Arroyo and Selker (2003) study the effects of using different modalities for interruption in ambient displays concentrating on the effects of heat and light channels.

4.3.3. Timing selection

Another important element in the presentation of alternative foci is the time of presentation with respect to the execution of the current task. How often has someone walked into your office while you were writing something and you have said "Hold on, I'll just finish writing this sentence"? Although most tasks can be interrupted without major disruption, the exact timing of the focus switch is intuitively very important and the delay imposed for the new attentional focus is often acceptable. Oulasvirta and Salovaara (2004) propose a memory-based approach that would present interruptions during low working memory load and provide retrieval cues allowing the user to mentally restore the cognitive state necessary to resume the interrupted task. Adamczyk and Bailey (2004) and Bailey and Konstan (2006) have studied the impact of interruptions at different times in a task execution. Basing their work on event perception theories, the authors represent tasks as two level hierarchies composed of *coarse* events further split into *fine* events (for example, a coarse event would be the selection of the email application, which would then be further decomposed in selecting the email application, typing in the username, and typing in the password). The authors then measure the impact of interruptions as they occur at various points within these hierarchies and demonstrate that the best times for interruptions correspond to coarse breakpoints. The availability of such a hierarchical task model would allow a system to infer the best time for interruption. Other authors have proposed temporal strategies for interruption normally basing their strategies on a decomposition of task completion in some phases (e.g. planning, execution, and evaluation). A short review of this work can be found in the papers quoted above.

5. A research agenda for attention aware systems

This paper has supplied the reader with an overview of the many pieces of the puzzle that make up current research on attention aware systems, and perhaps has allowed the assembling of some of those pieces. However, the road to completing the picture seems still quite long and in this section we explore a few issues that should be addressed to make our reconstruction a bit easier. As in a real puzzle we will describe the remaining work in terms of three important questions: What is the frame of the picture, i.e. how can we define and delimit attention aware systems? How can the different pieces be assembled, i.e. how can the work in different disciplines, or within the same discipline but taking different approaches, be understood and integrated? Do we have all and only the necessary pieces of the puzzle, i.e. what research directions are most promising, and which still need to be explored?

5.1. Defining attention aware systems

We have defined attention aware systems as "systems capable of supporting human attentional processes". Vertegaal (2003, p. 32) has defined attentive user interfaces as

''computing interfaces that are sensitive to the user's attention''. Both these definitions, in most cases, are too general to allow us to compare different systems or theories that may go under this name, to integrate their results, or even to establish in some scientific way whether a system should have this name. We believe that establishing a more precise definition of this field of research should have a very high priority in our research agenda. So what types of definitions could be useful to create a secure frame for our puzzle? Amongst the many possible approaches, three seem to be most promising: functional, perceptual, and task oriented definitions. Before analysing these three alternatives it is important to distinguish between at least two approaches which have been implicit in current research.

The first approach, that we call *attentive dispatching*, is best represented by notification systems. Attentive dispatching systems aim at presenting the user with newly available information in a way that minimises disruption to activity whilst satisfying informational needs. The tacit assumptions backing attentive dispatching are that (1) the user's current focus should be maintained or at least restored after the dispatch has taken place (and it has possibly been acted upon), that (2) the information dispatched to users is not already available to them in some other form, and that (3) relevant alternative foci are generated only by the system, and user-generated (or environment-generated) interruptions are not considered in the design. Variants to the attentive dispatching approach include systems where either assumption (2) or (3) are dropped.

The second approach, which is the one taken in this paper, sees attention aware systems as a super-set of attentive dispatching systems. It assumes that: (1) support may be given to users in orienting their attention as well as in maintaining it on a current task. (2) Users may overlook important information that is already available and they may want to be helped in the selection of this information. (3) Attention switches initiated by the user or by events in the environment are just as relevant as the attention switches provoked by the system, and should be evaluated with respect to the user goals. These assumptions have two consequences. First, the system must have a more in-depth knowledge of the users' tasks, and goals (as in the example of the system recognising Ann's goal to complete an exercise). Second, the system must consider a wider range of possible attentional foci (as in the example of the system recognising the preparatory exercise, the online status of the teacher, or the arrival of Ann's brother as alternative foci either related or unrelated to the current goal). Keeping the difference between attentive dispatching and more general attention aware systems in mind, we will now analyse possible ways of defining this latter type of systems.

5.1.1. Functional definitions of attention aware systems

Functional definitions of attention aware systems would indicate what type of support the system should provide to the user. For example users could be supplied with facilities that allow them to easily resume interrupted work. They could be reminded about unfinished tasks and supplied with the contextual information necessary to resume them. They could be presented with information relevant to the current task that they may have overlooked. Users could have the choice to specify times or situations in which they are available for interruption. They could indicate the relevance of certain goals and whether they would accept suggestions for their achievement. They could specify who are the people who should have priority in communicating with them and under which conditions the system should notify them of partner's activities and messages. And many other such functions could be listed.

The fact that the functions may be application independent (as in the examples above) is both an advantage and disadvantage of functional definitions. It is not clear at this stage of the research whether these functions could be compared to one another once placed within the frame of a specific application. Also, the definition of a meta-level, application independent, module dealing with attentional processes seems unlikely if users' current goals are to be evaluated to weigh possible foci. One could think of separating the system attentional mechanisms on two levels loosely corresponding to attentive dispatching and motivational related components of attention aware systems. The core of the attentive dispatching component may work at the meta-level, whilst the motivational component would be application dependent. This suggestion arises from the observation that once the relevance of a message has been established we can, on most occasions, decide on the most appropriate time and format for delivery without any specific knowledge of our communicative partners' current goals. Often, simple information about their current activities rather than about their goals (i.e. knowledge about the fact that someone is writing, or speaking, or listening, rather than knowledge about the fact that one is writing a paper, or a resignation letter, or that one is trying to convince someone to let him do something) is necessary to assess interruptibility. Indeed, if we can identify those functions/components of attention aware system that are application independent we could build a core set of such functionalities that could be re-used by several designers.

Another commonly recognised disadvantage of functional definitions is that they tend to be more focussed on the system rather than on the user experience; the next possible type of definitions moves the imbalance the opposite way: toward the user.

5.1.2. Perceptual definitions of attention aware systems

Perceptual definitions would specify user perception parameters to be used to measure the level of support to attentional processes provided by the system. One could for example define parameters such as user perceived facility to concentrate on a specific task, or to monitor desired information, or to switch from one activity to another one, or to resume an interrupted activity, or to perform tasks in multiple alternative manners, etc. Measures of stress levels (including, for example, predictability) and pleasure could also be included.

The advantage of this type of definition is that it would allow designers to take into account the great variety of individual differences in attentional processes independently of the specific system functionality. The disadvantage would be that, since the same perceptual result may be obtained by systems with very different designs and functionalities, capitalisation on existing work might be more difficult to achieve.

5.1.3. Reference tasks and metrics definitions of attention aware systems

Finally, we could define attention aware systems in terms of the reference tasks (Whittaker, Terveen, & Nardi, 2000) that they support along with the metrics necessary for task evaluation. This would amount to specifying key users' activities supported by attention aware systems and the performance expected. Whittaker and his colleagues see the definition of reference tasks in HCI as the necessary step toward establishing a common research focus, building on existing achievements, and comparing different techniques. Although this proposal is attractive, some aspects of its application to attention aware systems need further analysis. Whittaker and his colleagues have already indicated some problematic issues arising for the definition of reference tasks in HCI. These include the fact that some HCI results may not be implemented as part of a system (e.g. methodologies for design or

usability studies), that interactive tasks may be significantly more complex to evaluate than non interactive ones, and that ''bringing humans into the evaluation (as users, subjects, judges) produces a more complicated, costly, and subjective process'' (Whittaker et al., 2000, p. 114). Furthermore, basing their work on activity theory (Kaptelinin, Kuutti, & Bannon, 1995), the authors define tasks as *a conscious action subordinate to an object*. For example, *extracting information* from a *voice mail message* is a task or action subordinate to the *voice mail message* object. Activity theory however offers a much richer framework that we would need to further exploit in order to define reference tasks that are sufficiently meaningful for the definition and evaluation of attention aware systems. In particular, issues of social interaction and exchange with a physical environment, which are central to current activity theory, need to be reintegrated in the definition of reference tasks. Activity theory has already been applied to the design and analysis of interactive tasks in multi-user environments (see for example (Cluts, 2003; Fjeld et al., 2002)); therefore we believe that an extension of reference tasks in this direction would be feasible. Some of these applications advance models that appear particularly well suited for attention oriented task specification. For example, the definitions given by Fjeld et al. (2002) of *goal oriented* and *exploratory* actions seem to enable modelling of both general tasks that are currently executed by the user, and monitoring tasks that are currently out of focus. Extending the definition of reference tasks and relative measures to a richer vocabulary (possibly taken from activity theory) would enable us to include high level tasks such as: cooperatively writing a document, or completing an exercise in the context of a virtual learning environment. These tasks would need a further decomposition for evaluation. Such decomposition could also be described within the framework. Whittaker and his colleagues base their proposal of reference tasks definition on earlier work of several authors and in particular they see the metrics used to evaluate reference tasks as corresponding to Newman's *critical parameters* (Newman, 1997). Such parameters allow the comparison of different systems on the same task by establishing how well they score on each critical parameter. For example, critical parameters relevant to attention aware systems for the task of *completing an exercise in the context of a virtual learning environment* could include time for finding relevant associated material, time for contacting relevant partners, and time for writing, all in presence of a given number of external events that the student may wish to monitor. As we have mentioned earlier in this paper, some authors have already proposed critical parameters for attention aware systems. In a recent paper (Chewar et al., 2004) Chewar and his colleagues analyse advantages and problems in the definition of critical parameters for the design of interactive systems. In particular, they propose parameters for notification systems that allow one to measure the appropriateness and *interruption level* of interruptions, the user reaction they produce, and the level of comprehension (situated awareness) they may achieve.

In our view definitions of attention aware systems in terms of reference tasks and critical parameters could have the advantage of integrating both aspects of functional definitions – reflected in the reference tasks specifications – and of perceptual definition, that may be reflected in the metrics, or critical parameters specifications. The major challenge is the definition of a formal and agreed upon specification mechanism for reference tasks and critical parameters. This mechanism should allow researchers to specify reference tasks and critical parameters in a way that would be immediately and univocally understandable by others. Only once this mechanism has been defined and has reached a wide enough acceptance could we look forward to a rich and fruitful discussion on what the

reference tasks and critical parameters should be for HCI applications in general and attention aware systems in particular.

5.2. How can the work on attention in different disciplines be integrated?

Although human attention has been the subject of study of many different disciplines (as this paper has tried to demonstrate), the interdisciplinary work is still too limited to create the necessary awareness amongst a wide range of scholars about the work done in disciplines different from their own. However, as explained in Section 1, the design of systems capable of supporting attentional mechanisms is of paramount importance for the users of information and communication technologies. The creation of shared ontologies, interdisciplinary labs, and shared research agendas between researchers in cognitive psychology, computer science, sociology, and other disciplines addressing attention, should be encouraged.

Although a very significant amount of research on human attention has been undertaken in psychology, several HCI researchers agree that the reported theories are often too far removed from the specific issues relevant to human computer interaction to be easily applied to this field of research (McCrickard et al., 2003) and that more focussed research in this direction is needed (Horvitz et al., 2003). In the section describing the work on human attention, for example, we have seen that some psychologists believe that it is theoretically possible to provide users with information in a format such that it will be noticed only if relevant to the user's current purpose. However, in order to consistently achieve this behaviour in interactive systems, interdisciplinary exploration would be needed to find out what knowledge about the *user's current purpose* a system needs to have in order to provide the information in the appropriate format. Such an exploration should also address issues of whether the format for presentation is dependent on individual users' characteristics and/or on the presentation environment. Another example of a promising interdisciplinary area of study is related to users' behaviour during the execution of tasks with respect to learned attention and inattention to cues and how systems may adapt to this human characteristic.

Whilst efforts for integrating results from research in cognitive psychology have at least been initiated, there appears to be an almost complete lack of intake of results from disciplines such as sociology and rhetoric. This could be due to the fact that most of the research in attention aware systems has concentrated on bottom-up control of attention rather than top-down. As a consequence, motivational issues related to attention, such as the ones that would be addressed in those disciplines, have been largely disregarded. In relation to this issue some promising research directions are discussed in the next section.

5.3. What are the most promising research directions?

Throughout the paper we have pointed out open questions and issues that would benefit from further exploration. In this final section we highlight some of the most promising amongst those research directions for attention aware systems.

5.3.1. Addressing top-down control of attention
Present research has concentrated on bottom-up/exogenous control of attention (attention drawn by external events) rather than top-down/endogenous control (attention driven by one's goals). This is the case with respect both to system design and to basic research in

cognitive psychology. For system design the connection between attentional foci and goals or motivations has been largely limited to very high-level goals (see Section 4.1.2). In cognitive psychology top-down control, although recognised as fundamental, seems less understood than bottom-up control. The main obstacle to this understanding is that top-down control can be explained only within the frame of fairly complete theories of cognition, such as those discussed in Section 2.3. It seems however that attention aware systems in their broader definition, i.e., as a superset of attentive dispatching, will indeed require addressing top-down control of attention at a deeper level than has been done so far. This remains an open and promising field of research that should address questions such as: (1) How can current goals be detected and represented? (2) How could the relevance of a given focus be evaluated with respect to tasks, goals, and environment? (3) How can the cost/benefits of attention switching be evaluated taking into account the user's goals?

5.3.2. Addressing social, aesthetic, and rhetorical aspects

In the first part of this paper we have proposed that social, aesthetic, and rhetorical aspects are important in the analysis of human attention. It seems that, even in relatively simple tasks, a further focus on these processes would be beneficial. In the context of email notification, for example, the costs/benefits of a certain focus switch can often be evaluated only with respect to social considerations – what position the sender holds in the user's social network?

As we have discussed earlier, current research has studied how various mechanisms may convey information to users in fashions that are more or less disrupting for the current task. Both the way in which meaning is communicated through the content of a message and the choice of medium, style, and tone to convey this content can have a significant impact on how disruptive it may be. Further research on mechanisms capable of identifying the profile and expectations of the audience, as well as their more immediate goals could then provide systems capable of selecting appropriate tools to customise both content and presentation to suit the user profile.

5.3.3. Addressing the relationship between attention and awareness

The role of attention also needs to be further explored in awareness mechanisms. In collaborative situations it is important to know the relationship between the tasks performed by the user and his partners in order to decide whether a focus shift would be appropriate. Some authors in cognitive psychology have discussed the relationship between attention and awareness as mechanisms for conscious, or accessible knowledge (Lamme, 2004; Simons & Rensink, 2005). The results of these studies indicate that, contrary to what has been assumed by most designers of awareness mechanisms, making information about the environment available to the users is not a sufficient condition to create awareness. Attention is a necessary condition for awareness, or more precisely "attention is necessary for conscious change perception" (Simons & Rensink, 2005, p. 17). More focussed research is necessary to ensure that awareness mechanisms actually achieve their objective in situations where the user's cognitive capability may be overloaded and he may therefore miss relevant information.

5.3.4. Addressing attention in collaborative environments

As discussed in Section 4.2.2, optimising individual attention allocation may not be the best strategy in situations when the system aims at supporting collaborative activities. In

these environments, a focus that may be sub-optimal for an individual may in fact be optimal for the community. Further research is needed in order to explore methodologies for (1) the evaluation of cost/benefits of foci with respect to collective goals and activities; (2) effectively guiding users' attention to foci that they may consider suboptimal (social and motivational issues are obviously very relevant).

5.3.5. Other research areas

Other important areas that would benefit from further research include: (1) the definition of user models that could support the development of attention aware systems. (2) The characterization of strategies for optimising the timing of focus shifts. (3) The development of systems that address the most promising application areas as defined in Section 4. Finally, (4) theories of attention in natural language and human communication (Argyle & Cook, 1976; Clark, 1996, pp. 274–282; Grosz & Sidner, 1990) may provide useful insights for the definition and design of attention aware systems and could be further explored with this objective in mind.

6. Conclusion

Attention aware systems will be crucial for the development of applications in a wide variety of domains including education, life critical systems (e.g. air traffic control), support to monitoring and diagnosis, knowledge management, simulation of human like characters, games, and e-commerce.

In this paper we have presented some of the most significant research on attention in several fields, with an emphasis on cognitive psychology and system design. We have pointed out the findings of research on the processes of human attention that promise to be most useful for system design. We have analysed existing research on attention aware systems with respect to three fundamental and inter-related aspects of attention management: detection of current user's attentional state, detection and evaluation of possible alternative attentional foci, and strategies for presentation to the user of these alternative foci. Finally, we have sketched a research agenda for attention aware systems. We believe that important elements of this agenda include achieving a more precise and shared understanding of the purposes that these systems should serve, and developing a wider interdisciplinary base for the design of these systems.

References

Abowd, G. D., Mynatt, E. D., & Rodden, T. (2002). The human experience. *IEEE Pervasive Computing, 1*(1), 48–57.

ACT-R Research Group. (2002–2005). Act-r. Retrieved 2003, from http://act-r.psy.cmu.edu/about/.

Adamczyk, P. D., & Bailey, B. P. (2004). If not now, when? The effects of interruption at different moments within task execution. In *Proceedings human factors in computing systems: CHI'04*, New York.

Allport, A. (1989). Visual attention. In M. Posner (Ed.), *Foundations of cognitive science* (pp. 631–681). Cambridge, MA: MIT Press.

Allport, A., Styles, E., & Hsieh, S. (1994). Shifting intentional set: exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance xv: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge: Cambridge University Press.

Arroyo, E., & Selker, T. (2003). Self-adaptive multimodal-interruption interfaces. In *Proceedings international conference on Intelligent User Interfaces (IUI'03)*.

Arvidson, P. S. (2003). A lexicon of attention: from cognitive science to phenomenology. *Phenomenology and the Cognitive Sciences, 2*(2), 99–132.

Baddeley, A., & Weiskrantz, L. (Eds.). (1993). *Attention: Selection, awareness and control. A tribute to donald broadbent*. Oxford: Clarendon Press University.

Bailey, B. P., Adamczyk, P., Chang, T. Y., & Chilson, N. A. (2006). A framework for specifying and monitoring user tasks. *Computers in Human Behavior, 22*(4), 709–732.

Bailey, B. P., & Konstan, J. A. (2006). On the need for attention aware systems: measuring the effects of interruption on task – performance, error rate, and affective state. *Computers in Human Behavior, 22*(4), 685–708.

Bailey, B. P., Konstan, J. A., & Carlis, J. V. (2001). The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In *Proceedings INTERACT'01*.

Bartram, L., Ware, C., & Calvert, T. (2003). Moticons: detection distraction and task. *International Journal of Human–Computer Studies, 58*(5), 515–545.

Baum, E. (2004). *What is thought*. Cambridge, MA: MIT Press.

Bearne, M., Jones, S., & Sapsford-Francis, J. (1994). Towards usability guidelines for multimedia systems. In *Proceedings second ACM international conference on multimedia*, San Francisco, CA, United States.

Bohnenberger, T., Brandherm, B., Grossmann-Hutter, B., Heckmann, D., & Wittig, F. (2002). Empirically grounded decision-theoretic adaptation to situation-dependent resource limitations. *Künstliche Intelligenz, 3*, 10–16.

Bourdieu, P. (1990). *The logic of practice*. Stanford: Stanford University Press (R. Nice, Trans.).

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

Brush, B., Bargeron, D., Gupta, A., & Grudin, J. (2001). Notification for shared annotation of digital documents. Technical Report No. MSR-TR-2001-87, Microsoft Research Microsoft Corporation.

Carroll, J. M., Neale, D. C., Isenhour, P. L., Rosson, M. B., & McCrickard, D. S. (2003). Notification and awareness: synchronizing task-oriented collaborative activity. *International Journal of Human–Computer Studies, 58*(5), 605–632.

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: beyond a spotlight model. *Psychonomic Bulletin & Review, 6*(2), 204–223.

Chen, D., & Vertegaal, R. (2004). Using mental load for managing interruptions in physiologically attentive user interfaces. In *Extended abstracts of the 2004 conference on human factors and computing systems* (pp. 1513–1516). Vienna, Austria: ACM Press.

Chewar, C. M., McCrickard, D. S., & A.G., S. (2004). Unpacking critical parameters for interface design: evaluating notification systems with the irc framework. *Proceedings conference on designing interactive systems: Processes, practices, methods and techniques (DIS2004)*.

Chun, M. M., & Wolfe, J. (2001). Visual attention. In E. B. Goldstein (Ed.), *Blackwell's handbook of perception* (pp. 272–310). Oxford, UK: Blackwell.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Cluts, M. M. (2003). The evolution of artifacts in cooperative work: constructing meaning through activity. In *Proceedings international ACM SIGGROUP conference on Supporting group work*, Sanibel Island, FL, USA.

Cutrell, E., Czerwinski, M., & Horvitz, E. (2001). Notification, disruption, and memory: effects of messaging interruptions on memory and performance. In *Proceedings Interact 2001 IFIP conference on human–computer interaction*, Tokyo, Japan.

Czerwinski, M., Cutrell, E., & Horvitz, E. (2000). Instant messaging: effects of relevance and time. In Proceedings HCI 2000 – 14th British HCI group annual conference.

Deutsch, J., & Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological Review, 70*, 80–90.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92*, 53–78.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87*(272–300).

Fagerjord, A. (2003). Rhetorical convergence: studying web media. In Liestol, Morrison, & Rasmussen (Eds.), *Digital media revisited* (pp. 293–325). Cambridge, MA: MIT Press.

Fei Fei, L., VanRullen, R., Koch, C., & Perona, P. (2002). Natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America, 99*(14), 9596–9601.

Fjeld, M., Lauche, K., Bichsel, M., Voorhorst, F., Krueger, H., & Rautergerg, M. (2002). Physical and virtual tools: activity theory applied to the design of groupware. *CSCW, 11*(1–2).

Franke, J. L., Daniels, J. J., & McFarlane, D. C. (2002). Recovering context after interruption. In *Proceedings 24th annual meeting of the Cognitive Science Society (CogSci 2002)*.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding. I. Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23*, 121–134.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding. Ii. Feedback, expectation, olfaction, and illusions. *Biological Cybernetics, 23*, 187–202.

Grossberg, S. (1999). The link between brain learning, attention, and consciousness. *Consciousness and Cognition, 8*(1), 1–44.

Grosz, B. J., & Sidner, C. L. (1990). Plans for discourse. In P. R. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 417–444). Cambridge, MA: MIT Press.

Hewett, M. (2000). Computational perceptual attention. Unpublished PhD Dissertation, University of Texas at Austin, Austin, TX.

Heylighen, F. (2004). Complexity and information overload in society: why increasing efficiency leads to decreasing control. *Technological Forecasting and Social Change* (submitted) Available from http://pcp.vub.ac.be/Papers/PapersFH2.html.

Hill, R. (1999). Modeling perceptual attention in virtual humans. In *Proceedings 8th conference on computer generated forces and behavioral representation*, Orlando, FL.

Hinckley, K., Pierce, J., Sinclair, M., & Horvitz, E. (2000). Sensing techniques for mobile interaction. In *Proceedings ACM UIST 2000 symposium on user interface software and technology*, San Diego, CA.

Horvitz, E., & Apacible, J. (2003). Learning and reasoning about interruption. In *Proceedings 5th international conference on multimodal interfaces*, Vancouver, BC, Canada.

Horvitz, E., Kadie, C., Paek, T., & Hovel, D. (2003). Models of attention in computing and communication: From principles to applications. *Communications of the ACM, 46*(3), 52–59.

Hudson, J. M., Christensen, J., Kellogg, W. A., & Erickson, T. (2002). 'i'd be overwhelmed, but it's just one more thing to do: availability and interruption in research management. In *Proceedings SIGCHI conference on human factors in computing systems*, Minneapolis, MN, USA.

Hudson, S. E., Fogarty, J., Atkeson, C. G., Avrahami, D., Forlizzi, J., & Kiesler, S., et al. (2003). Predicting human interruptibility with sensors: a wizard of oz feasibility study. *Proceedings CHI 2003*, Ft. Lauderdale, FL, USA.

Hyrskykari, A., Majaranta, P., Aaltonen, A., & Räihä, K.-J. (2000). Design issues of idict: A gaze-assisted translation aid. In *Proceedings ETRA 2000, eye tracking research and applications symposium*, Palm Beach Gardens, FL.

Hyrskykari, A., Majaranta, P., & Räihä, K. -J. (2003). Proactive response to eye movements. *Proceedings INTERACT 2003*.

Jameson, A., Schafer, R., Weis, T., Berthold, A., & Weyrath, T. (1999). Making systems sensitive to the user's time and working memory constraints. In *Proceedings of the 4th international conference on intelligent user interfaces* (pp. 79–86). Los Angeles, CA, United States: ACM Press.

Jersild, A. (1927). Mental set and shift. *Archives of psychology (Whole No. 89) – Referenced in Rubinstein, Meyer, and Evans 2001*.

Kaptelinin, V., Kuutti, K., & Bannon, L.J. (1995). Activity theory: basic concepts and applications. In *Proceedings human–computer interaction, 5th international conference, EWHCI'95*, Moscow, Russia.

Kieras, D. E., Meyer, D. E., Ballas, J. A., & Lauber, E. J. (2000). Modern computational perspectives on executive mental processes and cognitive control: where to from here? In S. Monsell & J. Driver (Eds.). *Control of cognitive processes: Attention and performance* (Vol. xviii, pp. 681–712). Cambridge, MA: MIT Press.

Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse*. Arnold and Oxford University Press.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45*(6), 812–863.

Kruschke, J. K. (2003). Attention in learning. *Current Directions Psychological Science, 12*, 171–175.

LaGrandeur, K. (2003). Digital images and classical persuasion. In Hocks & Kendrick (Eds.), *Eloquent images: Word & image in the age of new media* (pp. 117–136). Cambridge, MA: MIT Press.

Lamme, V. A. F. (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Networks, 17*(5–6), 861–872.

Lavie, N. (2000). Selective attention and cognitive control: dissociating attentional functions through different types of loads. In S. Monsell & J. Driver (Eds.). *Attention and performance, Vol. xviii, control of cognitive processes* (pp. 175–197). Cambridge, MA: MIT Press.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics, 56*(2), 183–197.

Lewis, J. (1970). Semantic processing of unattended messages using dichotic listening. *Journal of Experimental Psychology, 85*, 225–228.

Löwgren, J. (2002). Just how far beyond hci is interaction design? Retrieved 20/11/04, 2004, from http://www.boxesandarrows.com/archives/just_how_far_beyond_hci_is_interaction_design.php.

MacKay, D. (1973). Aspects of the theory of comprehension, memory and attention. *Quarterly Journal of Experimental Psychology, 25*, 22–40.

Mackintosh, N. J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological Review, 82*, 276–298.

Maulsby, D. (1997). Inductive task modeling for user interface customization. In *Proceedings IUI'97: 2nd international conference on intelligent user interfaces*, Orlando, FL, United States.

McCrickard, D. S., Catrambone, R., Chewar, C. M., & Stasko, J. T. (2003). & Establishing tradeoffs that leverage attention for utility: Empirically evaluating information display in notification systems. *International Journal of Human–Computer Studies, 58*(5), 547–582.

McCrickard, D. S., & Chewar, C. M. (2003). Attuning notification design to user goals and attention costs. *Communications of the ACM, 46*(3), 67–72.

McCrickard, D. S., Czerwinski, M., & Bartram, L. (2003). Introduction: design and evaluation of notification user interfaces. *International Journal of Human–Computer Studies, 58*(5), 509–514.

McFarlane, D. C. (2002). Comparison of four primary methods for coordinating the interruption of people in human–computer interaction. *Human–Computer Interaction, 17*(1), 63–139.

McFarlane, D. C., & Latorella, K. A. (2002). The scope and importance of human interruption in human–computer interaction design. *Human–Computer Interaction, 17*(1), 1–62.

Nagata, S. F. (2003). Multitasking and interruptions during mobile web tasks. In *Proceedings 47th annual meeting of the human factors and ergonomics society*.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Croft.

Newman, W. M. (1997). Better or just different? On the benefits of designing interactive systems in terms of critical parameters. In *Proceedings designing interactive systems: Processes, practices, methods, and techniques*, Amsterdam, The Netherlands.

Norman, D. (1969). *Memory and attention: An introduction to human information processing*. New York: Wiley.

O'Conaill, B., & Frohlich, D. (1995). Timespace in the workplace: Dealing with interruptions. *Proceedings CHI'95 conference companion*, Denver, CO, United States.

Obermayer, R. W., & Nugent, W. A. (2000). Human–computer interaction for alert warning and attention allocation systems of the multi-modal watchstation. *Proceedings SPIE 2000*. The International Society for Optical Engineering.

Oulasvirta, A., & Salovaara, A. (2004). A cognitive meta-analysis of design approaches to interruptions in intelligent environments. In *Extended abstracts of the 2004 conference on human factors and computing systems* (pp. 1155–1158). Vienna, Austria: ACM Press.

Parasuraman, R., & Davis, D. R. (Eds.). (1984). *Varieties of attention*. San Diego: Academic Press.

Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Posner, M. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Posner, M. (1982). Cumulative development of attention theory. *American Psychologist, 37*, 168–179, Quoted in Parasuraman, R. (2000). Attention as selection. *The American Journal of Psychology, 113*(1), 135.

Rafal, R., & Henik, A. (1994). The neurobiology of inhibition: integrating controlled and automatic processes. In D. Dagenbach & T. Carr (Eds.), *Inhibitory processes in attention, memory, and language*. New York: Academic Press.

Raskin, J. (2000). *The humane interface: New directions for designing interactive systems*. Reading, MA: Addison-Wesley.

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes and application* (pp. 3–21). New York: Elsevier.

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research, 40*(10–12), 1469–1487.

Roda, C., Angehrn, A., Nabeth, T., & Razmerita, L. (2003). Using conversational agents to support the adoption of knowledge sharing practices. *nteracting with Computers, 15*(1), 57–89 (Special issue on intelligence & interaction in community-based systems).

Roda, C., & Thomas, J. (2005). Attention aware systems. In: C. Ghaoui (Ed.), *Encyclopaedia of hci*. IDEA Group 55–62.

Rogers, R., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207–231.

Rubinstein, J. S., Meyer, D. E., & Evans, J. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27*(4), 763–797.

Seshadri, S., & Shapira, Z. (2001). Managerial allocation of time and effort: the effects of interruptions. *Management Science, 47*(5), 647–662.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16–20.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89*, 669–679.

Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences, 34*(4), 771–797.

Stiefelhagen, R. (2002). Tracking focus of attention in meetings. *Proceedings 4th IEEE international conference on multimodal interfaces*, Pittsburgh, PA.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Tipper, S. P. (1985). The negative priming effect: inhibitory effects of ignored primes. *Quarterly Journal of Experimental Psychology A, 37*, 571–590.

Toet, A. (2006). Gaze directed displays as an enabling technology for attention aware systems. *Computers in Human Behavior, 22*(4), 615–647.

Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12*, 242–248.

Treisman, A. (1969). Strategies and models of selective attention. *Psychological Review, 76*, 282–299.

Treisman, A. (1998). Features and objects: the fourteenth bartlett memorial lecture. *Quarterly Journal of Experimental Psychology A, 40*, 201–237 (referenced in Driver, 2001).

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136 (referenced in Chun & Wolfe, 2001).

VanRullen, R., Reddy, L., & Koch, C. (2004). Visual search and dual-tasks reveal two distinct attentional resources. *Journal of Cognitive Neuroscience, 16*(1), 4–14.

Vertegaal, R. (1999). The gaze groupware system: mediating joint attention in multiparty communication and collaboration. In *Proceedings SIGCHI conference on human factors in computing systems*, Pittsburgh, PA, United States.

Vertegaal, R. (2003). Introduction. *Communications of the ACM, 46*(3), 30–33.

Vertegaal, R., Chen, D., Shell, J., & Mamuji, A. (2006). Designing for augmented attention: towards a framework for attentive user interfaces. *Computers in Human Behavior, 22*(4), 771–789.

Whittaker, S., Terveen, L., & Nardi, B. (2000). Let's stop pushing the envelope and start addressing it: a reference task agenda for hci. *Human Computer Interaction, 15*, 75–106.

Wolfe, J. M. (2001). The level of attention: mediating between the stimulus and perception. In L. Harris (Ed.). *Levels of perception: A festschrift for ian howard*. Berlin: Springer. Accessed online (15.2.05) draft of 8/2/01, 12:06 PM at http://search.bwh.harvard.edu/pdf/Festschrift.pdf.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp. 223–256). London, UK: University College London Press.

Zhai, S. (2003). What's in the eyes for attentive input. *Communications of the ACM, 46*(3), 34–39.

Zijlstra, F. R. H., Roe, R. A., Leonova, A. B., & Krediet, I. (1999). Temporal factors in mental work: effects of interrupted activities. *Journal of Occupational and Organizational Psychology, 72*, 163–185.